ABSTRACT

STRAUB, CHRISTOPHER THOMAS. Metabolic Engineering of Extreme Thermophiles for Conversion of Renewable Feedstocks to Industrial Chemicals (Under the direction of Dr. Robert M. Kelly.)

Extremely thermophilic microorganisms (\( T_{\text{opt}} \geq 70^\circ \text{C} \)) challenge assumptions about the origins and limits of life. The diverse and specialized biological machinery, mechanisms, and systems utilized by these microorganisms ensure that these bacteria and archaea thrive in extreme thermal environments. These facets can be exploited for contributions to biotechnology in the pursuit of renewable, sustainable, and environmentally compatible solutions to needs for energy and materials.

*Caldicellulosiruptor bescii* is an anaerobic bacterium that was isolated from thermal hot springs. It grows optimally at 78°C (172°F) on plant matter that has washed into the hot springs in which it resides. *C. bescii* natively deconstructs and ferments the cellulose and hemi-cellulose primarily to acetate, CO\(_2\) and H\(_2\). *C. bescii* has previously been shown to utilize approximately 30% of the available carbohydrate content in biomass, limited by the complex network of lignocellulosic biomass. However, by implementation of a mild sodium hydroxide pretreatment, *C. bescii* metabolized 70% of the carbohydrate content of pretreated switchgrass. Wild-type and transgenic black cottonwood (*Populus trichocarpa*) were also evaluated as feedstocks for *C. bescii* conversion. With milled wild-type poplar, *C. bescii* converted only 25% of the carbohydrate content, in contrast to nearly 90% of the carbohydrate content of a low lignin transgenic line created by down-regulation of the coumarate-3-hydroxylase (C3H3) gene. Furthermore, when entire stem samples of the transgenic line were utilized without milling, *C. bescii* solubilized over 50% of the feedstock, which has important biotechnological and economic consequences.

The native *C. bescii* fermentation products have little economic value. By re-engineering the organism’s metabolism, the fermentation can be directed toward chemicals of value. Thermophilic enzymes capable of producing acetone at 70°C *in vitro* were identified, expressed, and characterized. These enzymes were expressed in a metabolically engineered strain of *C. bescii*, producing acetone at titers up to 4.5 mM (0.26 g/L), thereby providing a platform strain for further optimization. An engineering analysis illustrated the process benefits of producing volatile compounds, such as acetone or ethanol in extreme thermophiles, exploiting product volatility to facilitate separation.
*Pyrococcus furiosus* is found in ocean volcanic vents, rich in sulfur, hydrogen gas, carbon monoxide, and carbon dioxide. Optimally growing at 100°C (212°F), *P. furiosus* metabolizes the most abundant organic compounds found in the ocean - laminarin from seaweed, chitin from the exoskeletons of arthropods, and peptides from deceased animals to name a few. As with *C. bescii*, *P. furiosus* has adapted to thrive by producing primarily acetate, H\(_2\) and CO\(_2\). Its central glycolytic metabolism differs somewhat from the canonical Embden-Meyerhof pathway which affects its suitability as a metabolic engineering platform for production of bio-based chemicals. Engineered strains were generated that utilized either ferredoxin-dependent glyceraldehyde-3-phosphate oxidoreductase (Gapor) or non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (Gapn) for a key step in central glycolysis to examine how this effects the resulting fermentation profile. In addition, an NADPH-dependent primary alcohol dehydrogenase was identified in *Caldanaerobacter subterraneus* and further characterized. Insertion of this gene alone allowed *P. furiosus* to produce ethanol at high yield at 80°C. Its insertion into various strains demonstrated the effects of modification of central glycolysis on a heterologous pathway. Insights into the glycolytic pathway in engineered strains further elucidated the potential for *P. furiosus* as a metabolic engineering platform.

These two extreme thermophiles provide a strategic set of native capabilities to degrade and metabolize nature’s renewable resources. Metabolic engineering to significantly alter their fermentation profiles has resulted in both fundamental understandings of key factors affecting their performance and demonstrated improvements toward this goal. Further work to optimize pathways to improve yield and titers and to maintain metabolic activity are the two major challenges remaining for these two extreme thermophiles to reach their biotechnological potential.
DEDICATION

To my high school chemistry teacher, Mrs. Eleanor Herbstritt, and my high school physics teacher, Mr. James Lallman.
BIOGRAPHY

Christopher Thomas Straub was born and raised in Saint Marys, Pennsylvania, as were both of his parents, all four of his grandparents, and all eight of his great-grandparents. He attended Catholic schools in Saint Marys and graduated from Elk County Catholic High School in 2003. A child of two Penn State graduates and a lifelong fan of the Nittany Lions, Chris pursued his undergraduate studies at The Pennsylvania State University – University Park where he obtained his degree (BS) in Chemical Engineering. Following graduation, he moved to Delaware where he worked for the E.I. du Pont de Nemours Company in manufacturing and technical roles in the fluorine chemicals and polymers division.

After a seven year career at Dupont, Chris decided to pursue a PhD and began his studies at the North Carolina State University in August 2014. He spent the following years engineering extreme thermophiles in order to advance the field of renewable biofuels and biochemicals under the advisement of Dr. Robert M. Kelly. Following receipt of his PhD, Chris plans to return to industry with a focus on process development to bring innovative ideas to industrial scale.
ACKNOWLEDGEMENTS

There are many people to thank for their help along the way. I would like to thank my family and friends for their support and encouragement. I would like to extend my sincere gratitude to all of the undergraduates who worked with me during my time here. Your assistance was incredibly valuable and the mentorship experience was rewarding. I owe a great deal of gratitude to my graduate advisor, Dr. Robert M. Kelly, and all of the members of the Kelly Lab at NC State.

For all those who helped me along the way, no one put more trust in me than the taxpayers of the United States. More than $500,000 of taxpayer funds went toward my tuition, stipend, student health insurance, lab supplies, travel, and other expenses. I will work each and every day to demonstrate that not only was I a worthwhile venture, but also that the public funding of science is a high-return investment for our citizens, industries, markets, and the entirety of our economy.
TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. ix
LIST OF FIGURES ................................................................................................................ x
CHAPTER 1: Extremely Thermophilic Energy Metabolisms: Biotechnological Prospects . 1
Abstract ................................................................................................................................. 2
Introduction ............................................................................................................................ 2
Genetics ................................................................................................................................. 3
Carbon Dioxide Fixation ..................................................................................................... 3
Carbon Monoxide ............................................................................................................... 4
Alcohols ............................................................................................................................... 5
Impact of environmental factors on host metabolism ...................................................... 6
Chemolithotrophy ................................................................................................................ 7
Hydrogen/Hydrogenases/Bifurcation ................................................................................ 8
Lignocellulose & Biomass Degradation ............................................................................ 10
Conclusions ......................................................................................................................... 10
Acknowledgments ............................................................................................................... 11
References ............................................................................................................................ 12
CHAPTER 2: Biotechnology of Extremely Thermophilic Archaea ......................... 23
Abstract ................................................................................................................................. 24
Introduction ............................................................................................................................ 25
Biotechnological potential of extremely thermophilic archaea ....................................... 26
H₂ and C1 metabolism in extremely thermophilic archaea .............................................. 27
  Molecular Hydrogen (H₂) ............................................................................................... 28
  Carbon monoxide (CO) and Formate (HCOO⁻) ......................................................... 30
  Carbon dioxide (CO₂) .................................................................................................. 31
  Methane (CH₄) .............................................................................................................. 34
Sulfur ................................................................................................................................. 36
  Sulfur reduction ............................................................................................................ 36
  Sulfur oxidation ............................................................................................................ 39
Metals ................................................................................................................................. 40
  Metal biooxidation ........................................................................................................ 41
  Biomining ...................................................................................................................... 42
Enzymes for Molecular Biology ....................................................................................... 44
  DNA Polymerases ....................................................................................................... 45
  Ligases ............................................................................................................................ 47
  Inteins ............................................................................................................................ 48
Enzymes for Biotransformations ....................................................................................... 49
  Protein expression ....................................................................................................... 50
  Glycoside hydrolases ................................................................................................. 50
CHAPTER 5: Quantitative Fermentation of Unpretreated Transgenic Poplar by
*Caldicellulosiruptor bescii* .................................................................................. 163
Abstract .................................................................................................................... 164
Introduction .............................................................................................................. 165
Results ..................................................................................................................... 165
Discussion ............................................................................................................... 168
Methods and Materials .......................................................................................... 169
References ............................................................................................................... 174

CHAPTER 6: A synthetic enzymatic pathway for extremely thermophilic acetone production based on the unexpectedly thermostable acetoacetate decarboxylase from *Clostridium acetobutylicum* .................................................................................. 176
Abstract .................................................................................................................... 177
Introduction .............................................................................................................. 178
Methods and Materials .......................................................................................... 180
Results ..................................................................................................................... 184
Extremely Thermophilic Acetone Pathway Candidates .......................................... 184
Recombinant Production of Acetone Pathway Enzymes ........................................ 184
Biochemical Characterization of Extremely Thermophilic Acetone Pathway Enzymes 186
Acetone Production *in vitro* by the Extremely Thermophilic Enzyme Pathway ...... 188
Discussion ............................................................................................................... 189
Acknowledgements ................................................................................................. 193
References ............................................................................................................... 202

CHAPTER 7: Bio-acetone production from microcrystalline cellulose by the metabolically engineered extreme thermophile *Caldicellulosiruptor bescii* .................. 206
Abstract .................................................................................................................... 207
Introduction .............................................................................................................. 208
Results and Discussion .......................................................................................... 210
*C. bescii* Acetone Tolerance .................................................................................. 210
Metabolic Acetone Conversion to Isopropanol in *C. bescii* ...................................... 211
Processing Advantages of Acetone Production in Extreme Thermophiles ............. 212
Biosynthetic Pathway for Acetone Production ......................................................... 213
Driving Acetone Production in Engineered *C. bescii*: Acetate Addition to Medium ... 215
Optimizing Bio-Acetone Production in *C. bescii* .................................................. 216
Conclusion ............................................................................................................... 217
Methods and Materials .......................................................................................... 218
Acknowledgements ................................................................................................. 219
References ............................................................................................................... 224
CHAPTER 8: Analysis of a Modification in *Pyrococcus furiosus* Central Metabolism and Its Implications for Metabolic Engineering ................................................................. 229

Abstract .................................................................................................................................................. 230
Introduction ............................................................................................................................................... 231
Results and Discussion .......................................................................................................................... 232
  Gapor or Gapn Alone Can Support Glycolytic Growth in *P. furiosus* .................................... 233
  Utilizing a Redox Balanced Ethanol Pathway in *P. furiosus* ......................................................... 234
  Characterization of an extremely thermophilic primary alcohol dehydrogenase ........... 234
  *In vivo* AOR Pathway with *C. subterraneus* AdhA ................................................................. 235
  Gapn and Gapor Effects on Ethanol Pathway in *P. furiosus* ....................................................... 236
Conclusions ............................................................................................................................................... 237
Methods and Materials ......................................................................................................................... 238
References ............................................................................................................................................. 247
LIST OF FIGURES

Figure 1-1. Four carbon fixation cycles found in thermophilic organisms .................................. 18
Figure 1-2. Three routes from pyruvate to ethanol ...................................................................... 19
Figure 1-3. Iron as an electron donor and acceptor by acidothermophiles ................................. 20
Figure 1-4. Electron acceptors in fermentative metabolisms ...................................................... 21
Figure 1-5. In vitro route from carbohydrates to hydrogen at near theoretical yield ............... 22
Figure 2-1. In vitro synthetic pathway for hydrogen production .................................................. 74
Figure 2-2. CO₂ fixation cycles including those found exclusively in extremely thermophilic archaea ......................................................................................................................... 75
Figure 2-3. Sulfur and iron oxidation pathways in extremely thermophilic archaeal species applicable to biomining ........................................................................................................... 76
Figure 2-4. Metabolic Engineering Temperature Shift Strategy Demonstrated in P. furiosus . 77
Figure 3-1. Alkali Pretreated Switchgrass Fermentations (5 g/L) with C. bescii ......................... 131
Figure 3-2. Bioreactor-scale Lignocellulose Fermentation by C. bescii ........................................ 132
Figure 3-S1. Configuration of Continuous Purged Bioreactor for C. bescii Lignocellulose Fermentation ........................................................................................................................... 133
Figure 3-S2. Acetate Productivity for Gas Sparged Bioreactor Fermentation, with and without Continuous Liquid Purging ........................................................................................................... 134
Figure 3-S3. Composition of Feedstock (switchgrass) and Residual Material from Gas Sparged Bioreactor Fermentation, with and without Continuous Liquid Purging 135
Figure 4-1. Monolignol synthesis pathway present in angiosperms.......................................... 150
Figure 4-2. C. bescii fermentation production from poplar lines .................................................. 151
Figure 4-3. S/G Ratio Effect on C. bescii Fermentation Products .................................................. 152
Figure 4-4. Fermentation-Growth Factor of Trangenics Compared to Wild Type for C. bescii Fermentation .......................................................................................................................... 153
Figure 5-1. Solubilization and conversion of wild-type and low lignin transgenic poplar lines by ethanol-producing C. bescii over 7 days at 65°C ............................................................... 172
Figure 5-2. Images of wild type and lignin-modified transgenic poplar stems (#54 and #80) with and without solubilization by C. bescii ................................................................. 173
Figure 6-1. Vapor-liquid phase envelopes for acetone and ethanol at various pressures of “bio-reactive distillation” operation ................................................................. 194
Figure 6-2. SDS-PAGE and Blue-native PAGE of acetone enzymes ............................................. 195
Figure 6-3. Thermal stability of thermophilic acetone production enzymes .............................. 196
Figure 6-4. Alignment of conserved active site residues from Family I CoA transferases ..... 197
Figure 6-5. In vitro function of the full acetone pathway ............................................................. 198
Figure 6-6. Three enzyme pathway to acetone production ........................................................... 199
Figure 7-1. Conversion of lignocellulose to acetone, carbon dioxide, and hydrogen in metabolically engineered C. bescii ................................................................. 220

x
Figure 7-2. Analysis of potential in situ separation of acetone from fermentation with an extreme thermophile at T = 70°C ................................................................. 221
Figure 7-3. Conversion of microcrystalline cellulose to acetone by metabolically engineered *C. bescii* ........................................................................................................... 222
Figure 7-4. Transcription of heterologously expressed genes for synthetic acetone pathway in *C. bescii* relative to gapdh................................................................. 223
Figure 8-1. Central glycolytic pathway comparison........................................... 241
Figure 8-2. Energy Conservation in *P. furiosus* (A), AOR pathway to ethanol in *P. furiosus* (B).............................................................................................................. 242
Figure 8-3. Potential futile cycling during gluconeogenesis (growth on pyruvate) when gapn is expressed at high levels................................................................. 243
Figure 8-4. Alcohol specificity of *C. subterraneus* alcohol dehydrogenase .................. 244
CHAPTER 1: Extremely Thermophilic Energy Metabolisms: Biotechnological Prospects

Christopher T. Straub¹, Benjamin M. Zeldes¹, Gerrit J. Schut²,
Michael W.W. Adams², and Robert M. Kelly¹

¹Department of Chemical and Biomolecular Engineering, North Carolina State University,
Raleigh, NC 27695-7905

²Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA
300602

Abstract

New strategies for metabolic engineering of extremely thermophilic microorganisms to produce bio-based fuels and chemicals could leverage pathways and physiological features resident in extreme thermophiles for improved outcomes. Furthermore, very recent advances in genetics tools for these microorganisms make it possible for them to serve as metabolic engineering hosts. The key is to develop approaches that utilize the unique metabolic characteristics of high temperature microorganisms for alternative routes to desirable biotechnological products. This review considers recent developments in extreme thermophile biology as they relate to new horizons for energy biotechnology.

Introduction

By the dawn of the 21st century, the metabolic basis for microbial life could be probed through the strategic application of modern (molecular genetics, functional genomics) and classical (laboratory pure culture) methods. Such endeavors were largely focused on what are now referred to as ‘model organisms’. Escherichia coli (gram-negative), Bacillus subtilis (gram-positive) and Saccharomyces cerevisiae (eukaryote), among other mesophilic microorganisms, were examined with increasingly sophisticated experimental and computational tools to the point that prospects for synthetic life forms and minimal genomes could be considered [1]. While studies of model microorganisms laid the foundation for understanding the complex machinery and mechanisms of life, they only hinted at the diversity and unique niches within which life exists and develops. Now, as the circle of ‘model’ microorganisms expands with molecular genetic tools becoming more widespread [2], so too does the potential of microbial biotechnology. Extremophiles epitomize this vast potential, particularly those prospering at the upper thermal limits for life. The lens of modern biology facilitates a more discerning look at these microorganisms for new metabolic pathways and physiological features that can be exploited for technological benefits, especially as this relates to reducing the reliance upon fossil fuels to meet the energy demands of modern society. In this review, we consider recent developments in the biology and biotechnology of extreme thermophiles, microorganisms that grow optimally (T_{opt}) at and above 70°C, as these relate to bioenergy.
Genetics

To capture the full potential of extremely thermophilic microorganisms for technologically relevant applications, molecular genetics tools are needed. These are important not only for understanding the intrinsic bases of their metabolism and physiology, but to establish them as metabolic engineering platforms [3]. To this end, over the past decade, genetic systems for several extreme thermophiles have been developed and utilized in multiple laboratories, including for *Sulfolobus sp.* (*T*_{opt} = \sim 80°C) [4-6], *Thermococcus kodakarensis* (*T*_{opt} = 85°C) [7], *Pyroccocus furiosus* (*T*_{opt} = 100°C) [8,9], *Thermus thermophilus* (*T*_{opt} = 72°C) [10], *Thermoanaerobacter mathranii* (*T*_{opt} = 75°C) [11] and *Caldicellulosiruptor bescii* (*T*_{opt} = 78°C) [12]. The COM1 strain of the hyperthermophilic archaeon *Pyrococcus furiosus*, which was isolated in the laboratory, exhibits natural competence with chromosomal insertion achieved for gene clusters as large as 17 kb [13]. Recently, a high temperature kanamycin resistance based selection system (originally developed in *Thermus thermophiles* [14]) was adapted for use in the extremely thermophilic bacterium *Caldicellulosiruptor bescii*, significantly improving genetic manipulations with this bacterium and opening the door for metabolic engineering efforts [15]. Genetic tools for the extremely thermoacidophilic archaeon *Metallosphaera sedula* (*T*_{opt} = 75°C) [16] and the hyperthermophilic bacterium *Thermotoga sp.* (*T*_{opt} = \sim 80°C) [17,18] have also been reported but are in their early stages and not yet widely utilized. As genetic tools for extreme thermophiles become more versatile and reliable, new methods for genome editing (CRISPR) and promoter design can be incorporated into metabolic engineering efforts, thereby accelerating development of these microorganisms for industrial biotechnology.

Carbon Dioxide Fixation

Two of the six presently known CO₂ fixation pathways are found exclusively in extremely thermoacidophilic archaea. *Metallosphaera sedula* utilizes the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle [19], while *Ignicoccus hospitalis* (*T*_{opt} = 90°C) contains a related dicarboxylate/4-hydroxybutyrate cycle [20]. These cycles, closely related to both the reverse TCA cycle found in many thermophiles and the 3-hydroxypropionate cycle first identified in the thermophilic green non-sulfur bacterium *Chloroflexus aurantiacus* (*T*_{opt} = 55°C), provide opportunities for products related to pathway intermediates (see Figure 1-1) and the potential to incorporate CO₂. A reaction kinetic model was recently reported for the 3HB/4HB cycle as it
appears in the Sulfolobales which was used to evaluate metabolic engineering options for using all or parts of the cycle to produce bio-based chemicals [21].

Fixation of CO₂ (either from the atmosphere or from industrial emissions) to create renewable chemicals is a global, aspirational goal to address and reduce greenhouse gas emissions. Plants and algae as renewable feedstocks rely upon the Calvin-Benson-Bassham (CBB) carbon fixation cycle. However, the other known fixation pathways present opportunities to incorporate CO₂ directly into fuels and chemicals, bypassing carbohydrates as intermediates. Incorporation and reduction of CO₂ requires significant energy input, which is provided through chemolithoautotrophic growth in certain extreme thermophiles. The addition of CO₂ can be targeted for products containing highly oxidized functional groups (e.g., succinic acid) to minimize energy and redox inputs. Along these lines, recombinant *P. furiosus* strains have been created that express five genes encoding three enzymes from the *M. sedula* 3HP/4HB cycle to produce 3-hydroxypropionate from CO₂ and maltose [22,23]. While the incorporation of CO₂ occurs at the expense of ATP and is followed by two NADPH-dependent steps, this aligns with traditional Embden-Meyerhof-Parnas glycolysis, which generates two ATP and four reducing equivalents from glucose. Thus, if reductases are present to transfer electrons to NADPH, a carbon neutral, energy neutral, and electron neutral pathway results.

**Carbon Monoxide**

The utilization of CO (typically from a syngas stream) by microorganisms has been considered for some time as a route to bio-based products. While certain anaerobic microorganisms can utilize CO via the Wood-Ljungdahl pathway (reductive acetyl-CoA), the ability to engineer in cofactors necessary for this pathway adds complexity to the metabolic engineering effort. However, other options exist to utilize CO. *P. furiosus* has recently been engineered with a 16-gene cluster encoding CO dehydrogenase/hydrogenase from *Thermococcus onnurineus*, allowing the generation of an ion-gradient via the oxidation of CO to CO₂ coupled to the evolution of H₂ [24]. *P. furiosus* can conserve energy from the ion gradient via an ATP synthase, thus permitting the oxidation of CO to act as the primary energy source. This pathway does not fix carbon and, thus, a complex carbon source (maltose or peptides) was necessary to provide for biosynthesis. Since *P. furiosus* typically produces acetate as a metabolic by-product from these complex carbon sources, the demonstration of acetate consumption during grown on CO indicates
the reverse reaction with regeneration of acetyl-CoA. In another interesting study, a two-step process with *T. onnurineus* and *Thermoanaerobacter kivui* (*T*<sub>opt</sub> 65°C) converted a CO-based gas mixture, typical of steel mill waste streams, to acetate. In the first step, *T. onnurineus* converted 70% of the CO to CO₂ and H₂, which was then fed to *Thermoanaerobacter kivui* (*T*<sub>opt</sub> 65°C), obtaining acetate (~0.5 g/L) [25]. The two-stage process, utilizing a carboxytrophic thermophilic archaeon and an autoacetogenic thermophilic bacterium, beneficially exploits their native capabilities toward transferring waste syngas-like streams into chemicals. Further genetic engineering efforts may permit more interesting products to be generated in such processes.

**Alcohols**

Many ethanoLOGenic organisms, such as *Saccharomyces cerevisiae*, obtain 2 ATP per glucose through traditional EMP glycolysis with the resulting pyruvate being decarboxylated via pyruvate decarboxylase complex (PDC) to acetaldehyde that is subsequently reduced by NADH to ethanol (Figure 1-2). However, a thermophilic PDC, while of great potential importance, has not yet been reported. Thus most thermophilic native ethanol-producing organisms follow a three-step pathway from pyruvate to ethanol. The moderately thermophilic *Clostridium thermocellum* (*T*<sub>opt</sub> 60°C) has achieved an ethanol titer of 38 g/L by removal of three genes that redirected carbon flux away from the organism’s other native products, acetate and lactate [26].

The production of ethanol from pyruvate is energetically unfavorable when compared to acetate production, since it eliminates the ATP-yielding hydrolysis of acetyl-CoA. By utilizing a unique aldehyde oxidoreductase (AOR) from *P. furiosus*, acetate can be reduced to ethanol without ATP expenditure. This enzyme is capable of reducing organic acids to aldehydes, utilizing electrons from the low potential electron carrier ferredoxin. Through a NADPH-dependent alcohol dehydrogenase, these aldehydes are subsequently converted to alcohol with high efficiency. This pathway can be extended to other metabolic engineering efforts to convert organic acids to alcohols, since both the AOR and the ADH demonstrated broad substrate activity [27]. In theory, this pathway allows an ethanol-producing organism with EMP central glycolysis to obtain 4 ATP from glucose to ethanol, thereby doubling the energy yield of most ethanologens. *P. furiosus* was also engineered to produce n-butanol, utilizing a template for this alcohol pathway from *Clostridium acetobutylicum* [23]. By recruiting seven enzymes from four different thermophilic species, n-butanol (~60 mg/L) production was demonstrated. Complementary to this effort, a
reaction kinetic model provided insights into optimal enzyme concentrations and promoter engineering for improved n-butanol production and selectivity [28].

The *Caldicellulosiruptor* species possess a unique ability to access and consume the carbohydrate content of lignocellulosic biomass, which is metabolized to acetate, lactate and hydrogen. *C. bescii* was engineered to eliminate lactate formation while producing ethanol, through the insertion of an NADH-dependent bi-functional acetaldehyde/alcohol dehydrogenase from *C. thermocellum* for operation at 65°C. Growth on untreated switchgrass yielded 0.6 g/L ethanol, demonstrating direct conversion of lignocellulose to product without pre-treatment nor exogenous enzyme additions [29]. The NAD(P)H-dependent acetaldehyde/alcohol dehydrogenases from *Thermoanaerobacter pseudethanolicus* were similarly engineered into *C. bescii* to more closely match enzyme thermoactivity with the optimal temperature of the host. While the operating temperature (75°C) may have improved the host metabolism, the comparatively lower ethanol concentrations (0.1 g/L) were likely a result of the limited availability of NAD(P)H [30].

Directing carbon flux away from organic acid production is critical for yield considerations in metabolic engineering efforts. Furthermore, organic acids can be toxic at low concentrations. By-products, such as acetate, can significantly inhibit growth, yet complete removal of the pathways for these compounds may be infeasible because of bioenergetics issues. To address these problems, detoxification by metabolite conversion has been demonstrated in *Thermoanaerobacterium saccharolyticum* (*T*<sub>opt</sub> 60°C). By insertion of a three-enzyme heterologous pathway, the low levels of acetate could be converted to acetone, thereby permitting three-fold improvement in ethanol concentrations to 45 g/L, an industrially relevant titer [31].

**Impact of environmental factors on host metabolism**

Microbial metabolic activity can be significantly influenced by environmental factors, creating a strategic opportunity for metabolic engineering. Extreme thermophiles could function as hosts for more moderately thermophilic enzymes and pathways, with the host metabolism operating in the ‘background’ while the heterologous pathways proceed. *P. furiosus* typically produces CO₂, acetate, and hydrogen from hexoses. Yet, insertion of a lactate dehydrogenase from *C. bescii*, an organism with optimal growth temperatures some 20°C below that of *P. furiosus*, enabled temperature-dependent gene regulation by insertion of the *ldh* gene downstream of the
cold shock promoter. Lactate was not detected near the optimal temperature for *P. furiosus* (98°C), yet lactate (0.3 g/L) was produced near the optimal temperature of the heterologous enzyme (72°C) [32]. In the aforementioned efforts with *P. furiosus* to produce 3HP, temperatures nearly 30°C below its optimal growth temperature were used in the production phase to minimize the background metabolism of the host [33]. In another study, metabolic profiling of *P. furiosus* revealed acetoin production (0.6 g/L) during growth at sub-optimal temperatures (70-80°C), yet this compound was absent at its optimal growth temperature of 100°C. [34]. Thus, these results with *P. furiosus* illustrate the potential for thermally-based gene regulation.

In addition to the temperature effects observed in *P. furiosus*, other environmental changes can impact core metabolism, such as variations in growth substrate in the metabolically versatile *Sulfolobus solfataricus* [35]. Toward that end, a genome-based metabolic network was constructed for *S. solfataricus* [36], which was expanded and validated by comparing metabolite, transcript, and enzyme activity levels during growth on glucose and fucose [37]. The results highlighted substantial differences between archaeal and bacterial core metabolism. A reversal of the ribulose-monophosphate pathway (involved in formaldehyde uptake in methanotrophs) was already known as an alternative to the pentose phosphate pathway in *S. solfataricus*, and in many other archaea [38]. This latest effort suggests that *S. solfataricus* uses a portion of the 3HP/4HB cycle in a manner similar to the glyoxylate shunt. *S. solfataricus* is able to obtain less biomass from growth on fucose than on glucose. To make up for this, it appears cells growing on fucose up-regulate enzymes of the 3HP/4HB cycle (resulting in a concomitant increase in chemical intermediates) in an attempt to replenish TCA cycle intermediates while minimizing loss of carbon during growth on this less nutritive sugar. These differences between archaea and bacteria will be important to understand, as archaea are considered further as metabolic engineering hosts.

**Chemolithotrophy**

Most metabolic engineering efforts are centered upon the fermentation of saccharides, such that energy and reducing equivalents are obtained through central metabolism and pyruvate is converted to a desired metabolite. However, there exist other options for cellular bioenergetics. Thermoacidophiles inhabit environments that are often devoid of organic carbon and oxidize metals and sulfur to provide the reducing potential and energy for autotrophy. By oxidizing Fe(II) to Fe(III), organisms, such as *Metallosphaera sedula*, can generate energy and reducing
equivalents required to grow autotrophically. Since the thermodynamics of carbon fixation require many Fe(II) atoms in order to fix a single carbon, relying solely on Fe(II) oxidation presents a challenge to conceiving a viable biotechnological process with reduced metal as the energy source.

Metal species can also act as important electron sinks in certain environments where electron acceptors, such as oxygen and sulfur, are absent. The thermoacidophilic archaeon, *Ferroglobus placidus* (T_{opt} 85°C), has been shown to utilize benzene as the electron donor and its oxidation is coupled to the reduction of Fe(III) to Fe(II) [39]. While the pathways of aromatic degradation were not completely elucidated, evidence was provided to confirm benzene consumption and reduction of Fe(III) as the sole source of energy metabolism. By combining both iron oxidation and reduction in a cycle (Figure 1-3), the potential exists to utilize complex aromatic substances, such as lignin, as a substrate for production chemicals [40].

In addition to iron oxidation, many extreme thermoacidophiles grow by the oxidation of metal sulfides (i.e., pyrite and chalcopyrite) [41,42] and elemental sulfur [43]. Growth on metallic ores is strategic for acidophiles, as the low pH solubilizes the metal substrate. On the other hand, elemental sulfur is sparingly soluble, and low pH prevents abiotic formation of more soluble sulfur compounds, such as thiosulfate and polysulfides [44]. An acidophilic solution to sulfur oxidation, first identified in thermoacidophilic archaea [45], is cytoplasmic sulfur oxygenase reductase (SOR), which has also been found recently in acidophilic bacteria [46]. In archaea, SOR seems to be essential for growth on elemental sulfur, although species that lack SOR may be able to grow on other sulfur compounds.

**Hydrogen/Hydrogenases/Bifurcation**

Many microorganisms, including extreme thermophiles, generate hydrogen gas to recycle reducing equivalents generated during oxidative metabolism. The thermodynamics of this reaction at higher temperatures can be advantageous. A temperature change from mesophilic conditions (37°C) to hyperthermophilic environments (100°C) increases the entropy term of the Gibbs free energy equation (\(\Delta G = \Delta H - T\cdot\Delta S\)) by 20%. Given that the majority of biological reactions operate close to equilibrium, this can transform a reaction that may be endergonic at mesophilic conditions to one that is exergonic and favorable at thermophilic conditions. Such is the case with many hydrogenases, since H\(_2\) production from reducing equivalents becomes much more favorable at elevated temperatures [47]. An excellent example is the case of formate oxidation to hydrogen by
the archaeon *Thermococcus onnurineus* \((T_{\text{opt}} \, 85^\circ\text{C})\). At ambient temperatures the reduction potentials of the formate/\(\text{CO}_2\) and \(\text{H}^+/\text{H}_2\) couples are almost identical but at \(85^\circ\text{C}\) hydrogen production is favored and the organism is able to conserve energy for growth [48].

In fermentative metabolism, the standard EMP pathway generates both NADH and reduced ferredoxin. The low potential of the ferredoxin \((E_m\) typically near \(-500\) mV) endows the capacity to generate \(\text{H}_2\) gas \((\text{H}_2/2\text{H}^+ = -420\) mV\) directly by a ferredoxin-dependent hydrogenase. However, it is not energetically favorable for NADH \((E_m = -320\) mV\) to be used as an electron donor for \(\text{H}_2\) production. Thus, NADH oxidation must be coupled to the production of reduced carbon products, such as lactate or ethanol from pyruvate. The necessity for an electron acceptor mandates that pyruvate, or a downstream product from pyruvate (e.g. acetaldehyde), be dedicated to this purpose. This prevents the ATP formation from pyruvate to acetate, resulting in a less energetically favorable metabolism (Figure 1-4).

In some microorganisms this problem is overcome by a so-called bifurcating hydrogenase, which was first discovered in the extremely thermophilic bacterium *Thermotoga maritima* [49]. This enzyme uses reduced ferredoxin and NADH simultaneously to produce \(\text{H}_2\). In essence, the enzyme uses the free energy of ferredoxin oxidation to drive the endergonic oxidation of NADH. As shown in Figure 1-4, this has very important consequences in fermentation as now all of the reductant generated from glucose oxidation can be used to produce \(\text{H}_2\) at what is known as the Thauer limit (4 \(\text{H}_2/\text{glucose}\)); the maximum amount of free energy is conserved (4 \(\text{ATP}/\text{glucose}\)). It is estimated that about one-third of all ferredoxin-dependent hydrogenases are bifurcating enzymes [50].

The complete oxidation of glucose to \(\text{CO}_2\) theoretically yields 12 mol \(\text{H}_2/mol\) glucose, but fermentative pathways yield a maximum of 4 mol \(\text{H}_2/mol\) glucose. Due to the necessary energy conservation in the form of ATP, acetate rather than \(\text{CO}_2\) is generated. The Thauer limit clearly presents a barrier to biohydrogen production processes [51]. To address this limitation, an *in vitro* pathway was constructed by recruiting thirteen enzymes from six thermophilic organisms, focusing on oxidative and non-oxidative components of the pentose phosphate pathway (Figure 1-5) [52]. This achieved nearly the theoretical yield of 12 mol \(\text{H}_2/mol\) glucose (sucrose), or three times the Thauer limit.
Lignocellulose & Biomass Degradation

Efforts to transform native biomass, such as grasses and trees, into fuels and chemicals is hindered by the complex lignocellulosic plant structure. Most efforts to transform these feedstocks into commercial products rely upon chemical and thermal pre-treatment, followed by addition of exogenous enzyme cocktails to produce fermentable hexoses and pentoses. As a further complication, many fermentative organisms, such as the moderate thermophile *Clostridium thermocellum*, utilize only hexose sugars [53].

Several species in the genus *Caldicellulosiruptor* not only degrade lignocellulosic biomass and access its carbohydrate content but also co-ferment hexoses and pentoses from complex polysaccharides [53]. Unique to the genus are a suite of glycoside hydrolases, polysaccharide lyases, and carbohydrate esterases [54]. A number of these enzymes have been characterized in detail to determine substrate specificity, kinetics, and importance to the overall complex process of accessing the carbohydrate content of lignocellulosic feedstocks [55]. One such multi-domain enzyme has been demonstrated as the most potent cellulolytic enzyme identified to date, with seven times higher activity compared to the standard exo- and endo- cellulase mixture [56].

Marine-based substrates for bio-based fuels and chemicals, such as algae and related seaweeds, are much less recalcitrant since they lack lignin and crystalline cellulose. The marine bacterium *Defluviitalea phaphyphila* (*T*<sub>opt</sub> 60°C) produced ethanol (10 g/L) at yields approaching 50% of theoretical from brown algae [57,58]. The ability to metabolize alginate, laminarin, and mannitol from the algae showed that marine-based photosynthetic carbon sources could be renewable feedstocks, potentially avoiding land use conflicts.

Conclusions

Not only do extreme thermophiles provide a source of unique metabolic pathways, carbon fixation pathways, and ability to metabolize unique substrates, there are also additional benefits related to thermodynamics at higher temperatures. As such, extremely thermophilic microorganisms can be valuable resources for addressing the current metabolic engineering barriers. By deliberate and structured design of pathways inspired by these non-model organisms, there now exists new opportunities for commercially relevant processes. Many of these metabolic features have been validated as a ‘proof-of-concept’, but it remains to be seen how extreme thermophiles will fit into the emerging industrial biotechnology sector.
Acknowledgments

This work was supported by grants to RMK and MWWA by the US Department of Energy Research ARPA-E Electrofuels Program (DE-AR0000081), the US National Science Foundation (CBET-1264052, CBET-1264053), the US Air Force Office of Scientific Research (AFOSR) (FA9550-13-1-0236), and the BioEnergy Science Center (BESC), a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.
References


Figure 1-1. Four carbon fixation cycles found in thermophilic organisms  Reverse TCA Cycle (yellow), 3-Hydroxypropionate/4-Hydroxybutyrate (blue), dicarboxylate/4-hydroxybutyrate (green), and 3-hydroxypropionate bicycle (orange).

*AH₂ indicates any or unknown electron carrier. A represents its oxidized counterpart.
**CoA inclusion and dissociation reactions not explicitly shown.
**Figure 1-2. Three routes from pyruvate to ethanol**  Four enzyme, ATP-generating pathway utilizing single step reduction of acetate to acetaldehyde (1-4 - yellow), two enzyme pathway utilizing direct decarboxylation of pyruvate to acetaldehyde (5 & 4 - blue), and three enzyme pathway with reduced ferredoxin generated during pyruvate to acetyl-CoA step followed by consumption of NADH during conversion to acetaldehyde. (1, 6, & 4 - green)

Acetyl-CoA Synthetase (ACS)[1], Acetyl-CoA Dehydrogenase (ACD), Aldehyde Oxidoreductase (AOR)[3], Alcohol Dehydrogenase (AdhA) [4], Pyruvate Decarboxylase Complex (PDC) [5], and Acetaldehyde Dehydrogenase (ALDH) [6].
Figure 1-3. Iron as an electron donor and acceptor by acidothermophiles  Aerobic oxidation of Fe$^{2+}$ by *M. sedula* allows the fixation of carbon dioxide. Anaerobic reduction of Fe$^{3+}$ by *F. placidus* permits oxidation of reduced carbon in benzene (39).
**Figure 1-4. Electron acceptors in fermentative metabolisms** (Left) Electrons from reduced electron carriers require high potential electron acceptor in the form of pyruvate or an aldehyde, resulting in energy yield of 2 ATP per glucose. (Right) Bifurcating hydrogenase allows energy yield of 4 ATP per glucose by generation of hydrogen from reduced electron carriers. (49)
Figure 1-5. *In vitro route from carbohydrates to hydrogen at near theoretical yield* (24 mol H₂ per dihexose) utilizing eleven thermophilic enzymes recruited from four different species (52).

*Two mesophilic and two thermophilic enzymes, all from different species, were utilized for conversion of sucrose to glucose-6-phosphate (not shown).*
CHAPTER 2: Biotechnology of Extremely Thermophilic Archaea

Christopher T. Straub¹, James A. Counts¹, Diep M. N. Nguyen², Chang-Hao Wu², Benjamin M. Zeldes¹, James R. Crosby¹, Jonathan M. Conway¹, Jonathan K. Otten¹, Gina L. Lipscomb², Gerrit J. Schut², Michael W.W. Adams², and Robert M. Kelly¹

¹Department of Chemical and Biomolecular Engineering
North Carolina State University, Raleigh, NC 27695-7905

²Department of Biochemistry and Molecular Biology
University of Georgia, Athens, GA 30602

Published in: FEMS Microbiol Rev. 42, no. 5 (Sep 2018): 543-578.
Abstract

Although the extremely thermophilic archaea (T_{opt} ≥ 70°C) may be the most primitive extant forms of life, they have been studied to a limited extent relative to mesophilic microorganisms. Many of these organisms have unique biochemical and physiological characteristics with important biotechnological implications. These include methanogens that generate methane, fermentative anaerobes that produce hydrogen gas with high efficiency, and acidophiles that can mobilize base, precious and strategic metals from mineral ores. Extremely thermophilic archaea also have been a valuable source of thermoactive, thermostable biocatalysts, but their use as cellular systems has been limited because of the general lack of facile genetics tools. This situation has changed recently, however, thereby providing an important avenue for understanding their metabolic and physiological details and also opening up opportunities for metabolic engineering efforts. Along these lines, extreme thermophilic archaea have recently been engineered to produce a variety of alcohols and industrial chemicals, in some cases incorporating CO_{2} into the final product. There are barriers and challenges to them reaching their full potential as industrial microorganisms but, if these can be overcome, a new dimension for biotechnology will be forthcoming that strategically exploits biology at high temperatures.
Introduction

Nearly four billion years ago, on an earth still cooling with a thin oxygen-free atmosphere, microbial life arose (Olsen et al., 1994). While debate still surrounds the details of primordial biology, extremely thermophilic archaea are “living fossils” and provide a glimpse into this critical period in evolution (Whitfield, 2004). Often overshadowed by their prokaryotic cousins, the *Bacteria*, in terms of both public perception and scientific study, their biochemical and physiological features offer intriguing opportunities for biotechnology. These are directly related to their proposed primitive beginnings: the ability to inhabit and thrive at extreme temperature and pH, along with metabolizing simple but technologically important compounds, such as hydrogen gas and C1 chemicals (methane, carbon dioxide, carbon monoxide), present in volcanic vents in the Hadean Ocean. As another hint of their prehistoric nature, extremely thermophilic archaea typically have small genomes (~2 Mb), potentially simplifying systems biology analysis and subsequent metabolic engineering efforts.

To date, archaea have lagged behind bacteria and eukaryotes on the industrial biotechnology stage, yet they have been important in several ways, most notably for their role in copper biomining (Wheaton et al., 2015) and for their high fidelity thermostable DNA polymerases in the Polymerase Chain Reaction (PCR) (Pavlov et al., 2004). Recent research breakthroughs have demonstrated that facile genetic systems developed for several archaea can be used for engineering non-native chemical production, in some cases from inorganic substrates, such as carbon dioxide (Keller et al., 2013, Hawkins et al., 2015, Zeldes et al., 2015, Loder et al., 2016). In applications where the most widely used mesophilic, metabolic engineering platforms, such as *Escherichia coli* and *Saccharomyces cerevisiae*, have failed or performed poorly, extremely thermophilic archaea offer a compelling option to overcome certain bioprocessing problems, such as contamination and phage infection.

A return to the extremely thermophilic archaea, typically located at the roots of phylogenetic trees, brings with it opportunity. A narrative that probably began billions of years ago in undersea volcanic vents enters the post-genomics era with momentum, promise, and significance. Herein, we chronicle the contributions associated with extremely thermophilic archaea to current biotechnology and peer into the future to preview the promise of these microorganisms and their associated biomolecules.
The key to any biotechnological uses of the extremely thermophilic archaea is the isolation of these microorganisms from unusual environments. Following Woese’s seminal proposal that the archaea constitute a third domain of life (Woese & Fox, 1977), there was a dramatic increase in the isolation of new extremely thermophilic archaea, led by the pioneering efforts of Stetter (Stetter, 1996), Jannasch (Jorgensen et al., 1992) and other intrepid microbiologists. These archaea were obtained from a variety of globally diverse, natural thermal features, ranging from deep sea hydrothermal vents to shallow terrestrial hot springs. By 2000, the pace of finding new genera of extremely thermophilic archaea slowed considerably, in part due to some convergence in the physiological characteristics of microorganisms being isolated from disparate natural biotopes. In fact, had genome sequencing technology been more widely available sooner, it would have been clear that many ‘newly’ identified archaea were in fact very closely related to previous finds. Since 2000, there have been relatively few reports of extremely thermophilic archaea with definitively new properties. However, numerous biotechnological opportunities exist based on those archaea that have already been described and future discoveries of new archaea should be expected. In fact, as we become more knowledgeable about extreme thermophile physiology and metabolism, and as ‘omics’ tools and systems biology approaches become even more powerful, return trips to thermal biotopes previously examined could identify important archaea that were overlooked previously. In addition, re-examination of extreme thermophiles already ‘in captivity’ may reveal previously undiscovered features that have biotechnological importance.

**Biotechnological potential of extremely thermophilic archaea**

Table 2-1 lists representative extremely thermophilic archaea with biotechnological potential as sources of important enzymes or as a consequence of their metabolic and physiological features. Several have facile genetic systems, making them even more significant as promising metabolic engineering platforms. More detailed discussion of these, and other extremely thermophilic archaea, can be found in the following sections. Nevertheless, it would be useful to first provide some context about the approaches that have been employed and could be taken to exploit the potential of these microorganisms for biotechnology.

In the 1990s, interest in extremely thermophilic archaea was fueled by the potential of their intrinsically thermophilic enzymes as biocatalysts that would be robust in the face of biologically unfavorable industrial conditions (Adams, 1994, Adams & Kelly, 1995, Adams et al., 1995,
Adams & Kelly, 1998). When it became clear that the genes encoding these enzymes could be overexpressed and the resulting polypeptides folded into their functional form in mesophilic microbial hosts (Zwickl et al., 1990), biotechnology ventures arose. Diversa Corporation was one example of such an enterprise that leveraged innovative methods for high throughput screening from gene expression libraries to access biodiversity for a range of enzyme-based applications (Palackal et al., 2004, Solbak et al., 2005). However, one potential concern for large-scale applications of archaeal enzymes then, that remains today, is the challenge of obtaining the high overexpression levels needed to meet industrial demand on economic scale, which is typically multi-grams of protein per liter.

In addition to being a source of thermostable and thermoactive biocatalysts, several of the archaea listed in Table 2-1 have industrial potential as sources of novel biosynthetic pathways for producing bio-based fuels and chemicals that could be engineered into mesophilic microorganisms with highly developed genetic systems, such as Escherichia coli or Saccharomyces cerevisiae. However, the incompatibility of the host temperature range with the enzymes from the extreme thermophile is problematic. This may not be an insurmountable obstacle in all cases. Some extremely thermophilic enzymes are active in mesophilic temperature ranges, but clearly this will be a technological barrier if sub-optimal temperatures are required for their application. The solution currently being pursued is to establish extreme thermophiles as metabolic engineering platforms through the development of molecular genetics tools for these microorganisms. There are many promising advances in this regard (Adams & Kelly, 2017), such that certain extremely thermophilic archaea can already be metabolically engineered to produce industrial chemicals and fuels. These archaea have great promise as new platforms which can advantageously exploit elevated temperatures as well as their unique biological characteristics (Zeldes et al., 2015, Counts et al., 2017, Loder et al., 2017).

**H₂ and C₁ metabolism in extremely thermophilic archaea**

From a biotechnological perspective, one of the most promising characteristics of extremely thermophilic archaea is their ability to use and produce simple chemicals, including C₁ compounds (CO, formate, CO₂, CH₄), and molecular hydrogen (H₂). As mentioned, this ability may map back to their primordial origins when these chemicals were the primary carbon and
energy sources available. Now, as we understand more about the underlying metabolism involving these chemicals, avenues to biotechnology are emerging.

**Molecular Hydrogen (H₂)**

Hydrogen has potential as a renewable and carbon-neutral energy carrier. With the availability of the commercial fuel cell applications, the demand for renewable hydrogen gas is expected to increase dramatically in the near future. However, the current methods of industrial H₂ production still heavily rely on fossil fuels, so alternative methods for renewable hydrogen production are desirable. For bio-hydrogen production, an increase in temperature from 37°C to 100°C enhances the entropy by 20%, as determined from the Gibbs free energy relationship. Therefore, elevated temperature benefits H₂ production (Verhaart *et al.*, 2010). Extensive studies on H₂ metabolism have focused on *Pyrococcus furiosus*, *Thermococcus onnurineus* NA1 and *T. kodakarensis* KOD1, aided by the availability of genetic systems (Kim *et al.*, 2010, Lipscomb *et al.*, 2011, Hileman & Santangelo, 2012). These archaea have proven to be prolific H₂ producers. For example, in continuous culture, *P. furiosus* grown on maltose and *T. kodakarensis* grown on pyruvate, demonstrate similar specific H₂ production rates of 0.16 g g⁻¹ h⁻¹ and 0.12 g g⁻¹ h⁻¹ respectively (Schicho *et al.*, 1993, Kanai *et al.*, 2005). Similarly the CO-dependent H₂ production from *T. onnurineus* NA1 achieved 0.19 g g⁻¹ h⁻¹ with continuous CO feed (Kim *et al.*, 2013). Improvements in H₂ production by these and related archaea will come with improved understanding of their metabolism to inform metabolic engineering efforts.

Hydrogenases, which catalyze the reversible reaction between protons, an electron source, and hydrogen, can be classified into three types based on the metal content in the catalytic site: [NiFe], [FeFe] and mononuclear Fe hydrogenases but, for reasons that are not clear, the [FeFe]-enzymes have yet to be found in any of the archaea (Vignais & Billoud, 2007, Sondergaard *et al.*, 2016). The NiFe-enzymes are ubiquitous and the extremely thermophilic archaean *P. furiosus* possesses three such enzymes. One is membrane-bound hydrogenase (MBH) and two are in the cytoplasm (SHI and SHII). Most members of the Thermococcales contain genes encoding at least one MBH and one SH (Schut *et al.*, 2013). The physiological function of MBH is to generate hydrogen from the reducing equivalents generated by glycolysis as well as to produce a H⁺/Na⁺ gradient for energy conservation (Sapra *et al.*, 2003). The deletion of the genes encoding MBH in *P. furiosus* abolished H₂ production and eliminated growth in the absence of elemental sulfur as
an external electron acceptor, while in *T. kodakarensis* the overexpression of MBH leads to increased H\textsubscript{2} production (Schut *et al.*, 2012, Kanai *et al.*, 2015). The predicted function of SH is to oxidize the H\textsubscript{2} that is produced and regenerate reduced nicotinamide cofactors, but the deletion of SH in *P. furiosus* did not affect growth, indicating that alternative systems are present to provide NADPH for biosynthesis (Lipscomb *et al.*, 2011). In *P. furiosus*, no significant change in H\textsubscript{2} production rate was observed when the genes encoding both SHI and SHII were deleted, while a 10% increase in H\textsubscript{2} production rate was obtained when the SH was knocked out in *T. kodakarensis* (Kanai *et al.*, 2011, Schut *et al.*, 2012). The type of glycosidic linkage in the carbon sources also affects H\textsubscript{2} production in *P. furiosus*, where cellobiose-grown cultures have 50%-higher specific H\textsubscript{2} generation rates compared to maltose-grown cultures (Chou *et al.*, 2007). Efforts have been reported to utilize extremely thermophilic archaea for H\textsubscript{2} production from waste materials, such as chitin, agricultural waste, and feather meal (Balint *et al.*, 2005, Hensley *et al.*, 2016, Aslam *et al.*, 2017), which can be re-visited if metabolic engineering strategies to improve the conversion of sugars and peptides become available.

*In vivo*, biological fermentation for H\textsubscript{2} production is limited by the so-called ‘Thauer Limit’, where the maximum theoretical yield is 4 H\textsubscript{2} per glucose. Nevertheless, this could be overcome by cell-free *in vitro* synthetic systems that have a theoretical yield of up to 12 H\textsubscript{2}/glucose, with half of the hydrogen atoms provided by water. The full oxidation of carbon to CO\textsubscript{2} allows this to be a thermodynamically feasible and redox balanced pathway. In addition, *in vitro* synthetic systems have an advantage of producing H\textsubscript{2} as the primary product compared to biological fermentation systems. Because of its specificity for NADP\textsuperscript{+} and intrinsic stability, soluble hydrogenase I (SHI) from *P. furiosus* has been used in a wide range of applications, including the production of reduced nicotinamide cofactors and of hydrogen gas (Wu *et al.*, 2015). Among these applications, H\textsubscript{2} production from sugars by an extensively studied synthetic pathway is one of the most promising methods for bio-hydrogen production currently. This *in vitro* synthetic system was originally designed to use glucose-6-phosphate (G6P) as the starting substrate to generate NADPH in a reconstructed pentose phosphate pathway using mesophilic enzymes from various sources, with H\textsubscript{2} produced by SHI using the formed NADPH (Woodward *et al.*, 2000). This pathway was further modified to include a phosphorylation step to generate G6P from sugars (Zhang *et al.*, 2007). Over the last decade, this pathway has been engineered to use different enzymes to generate monosaccharides from diverse sugars as the energy source, including
cellulosic materials (Ye et al., 2009), xylose (Martin del Campo et al., 2013), sucrose (Myung et al., 2014), corn stover (Rollin et al., 2015), and xylooligosaccharides (Moustafa et al., 2016). Recently, instead of directly using NADPH for H₂ production by SHI, a synthetic electron mediator, benzyl viologen (BV), was included in the pathway, where NADPH was used by NADPH rubredoxin oxidoreductase from *P. furiosus* to reduce BV that was further used by SHI for H₂ production (*Figure 2-1*) (Kim et al., 2016, Kim et al., 2017). Used in combination with enzymes from other thermophiles, the operating temperature was elevated to 50°C and the H₂ productivity was enhanced more than 200-fold, to 0.18 g L⁻¹ h⁻¹, using starch as the energy source for *in vitro* H₂ production (Zhang et al., 2007, Kim et al., 2017). This pathway also achieved the maximum yield of 12 H₂ per glucose. A recently developed hybrid photocatalytic system consisting of nanocrystalline CdSe/CdS dot-in-rod coupled with SHI from *P. furiosus* also demonstrated that H₂ could be produced in a high efficiency using light as the energy source (Chica et al., 2017).

**Carbon monoxide (CO) and Formate (HCOO⁻)**

Despite its advantages as a renewable energy carrier, the issues of storage and delivery of H₂ hinder potential applications, drawing interest to compounds such as formate as a potential hydrogen storage chemical (Joo, 2008, Enthaler et al., 2010). Two H₂ production membrane complexes have been identified in the genus of *Thermococcus onnurineus*: formate hydrogen lyase (FHL) and carbon monoxide dehydrogenase (CODH) (Takacs et al., 2008, Kim et al., 2010, Kim et al., 2013, Kozhevnikova et al., 2016). FHL and CODH are part of a modular family of MBH-type complexes that are involved in energy conservation by Na⁺-dependent H₂-respiration (Schut et al., 2016). They have homologous subunits to MBH in *P. furiosus*, including the Na⁺/H⁺ transporter (Mrp) and the membrane-bound hydrogenase (Mbh). FHL contains subunits of formate dehydrogenase for conversion of formate to CO₂, while CODH contains subunits of CO dehydrogenase that catalyze CO oxidation to CO₂ (Lipscomb et al., 2014, Schut et al., 2016). Production of H₂ from *T. onnurineus* with substrates CO, formate or starch achieved modest specific production rates but yielded 98%, 100% and 78% of theoretical conversion, respectively (Bae et al., 2012). Growth conditions of *T. onnurineus* have been optimized for H₂ production using formate as the energy source, and rates increased to 0.48 g L⁻¹ h⁻¹ (Bae et al., 2015). When the cell density of *T. onnurineus* increased, using formate as the energy source, a linear increase
of volumetric H\textsubscript{2} production rate with cell density was reported, where the productivity reached 5.70 g L\textsuperscript{-1} h\textsuperscript{-1} with a specific rate of 0.82 g g\textsuperscript{-1} h\textsuperscript{-1} (Lim \textit{et al.}, 2012). The H\textsubscript{2} productivity in \textit{T. onnurineus} was further improved by over-expressing an F\textsubscript{420}-reducing class hydrogenase and the FHL complex to achieve a specific productivity of 1.01 g g\textsuperscript{-1} h\textsuperscript{-1}, which is one of the highest productivities reported to date (Rittmann \textit{et al.}, 2015). Three \textit{fhl} operons were identified in the genome of \textit{T. onnurineus}, with the \textit{fhl2} operon being essential for growth with formate as the energy source (Kim \textit{et al.}, 2010). The operon encoding the 18-subunit \textit{fhl2} complex was heterologously expressed in \textit{P. furiosus} and the amount of H\textsubscript{2} production during the growth in the presence of formate and sugars was more than two-fold higher than the parent strain (Lipscomb \textit{et al.}, 2014).

The expression of CODH in \textit{T. onnurineus} has been manipulated to also increase the H\textsubscript{2} production rate from CO. The operon encoding CODH was over-expressed in \textit{T. onnurineus} under the control of a strong promoter, and the H\textsubscript{2} production increased by 3.8-fold over the wild-type, reaching 0.27 g L\textsuperscript{-1} h\textsuperscript{-1}. The specific rate was also 1.8-fold higher when the cells were grown in the presence of CO (Kim \textit{et al.}, 2013). \textit{T. onnurineus} can also utilize CO in waste gas from the steel industry for H\textsubscript{2} production, although a reduction of the maximum rate by 30\% was observed due to the relatively low content of CO in the waste gas (Kim \textit{et al.}, 2013). Acetate production from steel mill waste gas has also been reported by using \textit{T. onnurineus} in the first stage to which Linz-Donawitz converter gas (56\% CO as the sole carbon source) was fed, followed by a bacterial homoacetogen (\textit{Thermoanaerobacter kivui}) (Kim \textit{et al.}, 2016). Besides manipulating the expression of CODH, simply over-expressing a putative transcriptional regulator, TON1015, resulted in a 5-fold improvement in H\textsubscript{2} productivity to 0.37 g L\textsuperscript{-1} h\textsuperscript{-1} in \textit{T. onnurineus} compared to the wild-type (Rittmann \textit{et al.}, 2015). The 17-gene cluster encoding CODH in \textit{T. onnurineus} has been heterologously expressed in \textit{P. furiosus}, and when grown in the presence of CO, H\textsubscript{2} production was enhanced by almost an order of magnitude compared to the growth in the absence of CO. The recombinant strain also had the ability to utilize CO as a respiratory energy source for growth (Schut \textit{et al.}, 2016).

**Carbon dioxide (CO\textsubscript{2})**

Direct incorporation of CO\textsubscript{2} into bio-based fuels and chemicals is an aspirational goal of clean energy systems. This presents an important biotechnological opportunity for extremely
thermophilic archaea capable of using CO₂ as a carbon source. Of the six known carbon dioxide fixation cycles, two are found exclusively in extremely thermophilic acidophiles (Berg et al., 2007) (Figure 2-2). In contrast to photosynthetically powered carbon fixation characteristic of the Calvin-Benson-Bassham (CBB) cycle, thermoacidophilic archaea obtain energy to fix CO₂ chemolithoautrophically, via aerobic oxidation of sulfur (i.e., Sulfolobus metallicus (Huber & Stetter, 1991)), oxidation of ferric iron (Fe²⁺) (i.e., Metallosphaera sedula (Huber et al., 1989)) or anaerobic reduction of sulfur with hydrogen (i.e., Ignococcus hospitalis (Paper et al., 2007)). While these pathways may appear to have a common root, most evidence points to convergent evolution, which involves substantially different enzymes to perform similar reactions (Braakman & Smith, 2012).

The 3-hydroxypropionate (3-HP) bi-cycle, found exclusively in the green non-sulfur bacteria family Chloroflexaceae, relies on two branches with overlap through a central chain from acetyl-CoA through propionyl-CoA (Strauss & Fuchs, 1993) (see Figure 2-2). The central branch is responsible for sequestering two carbons: one branch produces a two-carbon glycoxylate, while the second branch assimilates glycoxylate, producing pyruvate. Both branches generate two carbons as acetyl-CoA, heading the central trunk of the pathway. At present, primary interest in utilizing this pathway has centered upon mesophilic hosts, such as E. coli, which have produced 3-HP titers above 40 g L⁻¹ (Liu et al., 2016). Although this pathway contains fourteen enzymes (Strauss & Fuchs, 1993), its bioenergetic and redox requirements provide an advantage over related cycles (DC/4-HB and 3-HP/4-HB), as discussed below.

In the Dicarboxylate Cycle/4-Hydroxybutyrate (DC/4-HB) cycle (see Figure 2-2), the section of the pathway from succinyl-CoA to acetyl-CoA is nearly identical to the 3-HP/4-HB cycle with its most unique feature being the three-step carbon fixing pathway from acetyl-CoA to oxaloacetate (Huber et al., 2008). In two steps, the DC/4-HB cycle sequesters two carbons at the expense of only one ATP and one reducing equivalent, nearly matching the efficiency of the reverse TCA cycle (rTCA). Thus, the DC/4-HB has been proposed as a potential thermophilic pathway to produce succinate, requiring two ATP and three reducing equivalents in addition to a source of acetyl-CoA to form succinate. The phosphoenolpyruvate (PEP) carboxylase, the central enzyme in this pathway, could be used in conjunction with other synthetic pathways due to this energy conservation (Bar-Even et al., 2010).
As with rTCA and DC/4-HB cycles, acetyl-CoA exits the 3-HP/4-HB cycle to meet cellular needs. As it requires 4 ATP and four reducing equivalents to fix a single carbon dioxide, it is the least energy efficient of the six carbon dioxide fixation pathways. This inefficiency is due to the lack of energy conservation during CoA transfers. Two of the four ATPs consumed in the cycle are required to fix carbonate, but the cycle also requires ATP at two separate points to transfer a free CoA onto a carboxylic acid. Given that it does not contain any intermediate that is only found in the cycle and its status as the least efficient carbon fixation pathway, at this point the biotechnological importance of the 3-HP/4-HB cycle stems from the three enzyme route from acetyl-CoA to 3-HP. Utilizing acetyl-CoA carboxylase and two successive NADPH dependent reductases, 3-HP production has been demonstrated in a recombinant extremely thermophilic host, *P. furiosus*, which grows optimally near 100°C (Keller et al., 2013, Hawkins et al., 2015, Lian et al., 2016). By insertion of the three enzymes from *Metallosphaera sedula*, *P. furious* was able to produce up to 0.5 g L⁻¹ 3-HP at 72°C, a non-optimal temperature for the host, but the optimal temperature for the pathway (Lian et al., 2016).

There are other features in extremely thermophilic archaea that relate to CO₂ fixation in mesophilic organisms. The CBB cycle is responsible for the majority of the earth’s carbon sequestration from the environment. The cycle depends upon Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) for the carboxylase activity, which is not only a relatively slow enzyme, but also suffers from catalyzing a competitive oxygenase reaction. In order to remove the bottleneck in CO₂ fixation processes, significant efforts to enhance both the kinetics and specificity of Rubisco have been undertaken. While there are no known archaea utilizing the CBB cycle, Type III Rubisco enzymes are found exclusively in archaea, possibly hinting at the evolutionary heritage of this important protein (Sato et al., 2007). *T. kodakarensis* contains such a version of this Rubisco that is neither affected by, nor inactivated by, oxygen (Ezaki et al., 1999). Additionally, this enzyme has been shown to function in vivo when cloned into a mesophilic bacterium, *Rhodopseudomonas palustris*, demonstrating its potential to operate at ambient temperatures (Nishitani et al., 2010). More recent efforts have attempted to understand these properties as well as explore mutations for optimization and insertion into photosynthetic mesophilic plants, bacteria, or algae to enhance carbon fixation (Fujihashi et al., 2016).

Analysis of the existing CO₂ fixation cycles has inspired the design of synthetic carbon fixation pathways by combining enzymes from a variety of sources toward achieving a kinetically
and thermodynamically favorable route for sequestering carbon into commercial products. In many cases, the carboxylation step - integration of carbon dioxide (or bicarbonate – HCO$_3^-$) is the rate-limiting step. A survey carboxylase enzymes responsible for the fixation of CO$_2$ or carbonate (HCO$_3^-$) suggested that an optimized synthetic pathway should contain phosphoenolpyruvate carboxylase (from DC/4-HB) and/or pyruvate carboxylase (in production of oxaloacetate) which possess the highest activity and affinity for the carbon species of those carboxylating enzymes yet reported (Bar-Even et al., 2010).

The thermodynamics of many biological reactions are at, or near, equilibrium and a subtle shift in pH, temperature, or other environmental factors can affect the favorability of certain pathways. These modifications in the environmental conditions in which the enzymes function could result in reactions with lower potential electron carriers or energy conserving steps. The archaeal carbon fixation cycles provide an enzymatic toolkit to construct hybrid pathways that function over a wide range of temperature and pH. Since many of these enzymes retain significant activity at lower temperatures, their potential use is not limited to thermophile metabolic engineering, but may find use in photosynthetic plants and algae (Wilson et al., 2016).

**Methane (CH$_4$)**

Methane is a valuable energy source and the biological production of this hydrocarbon offers biotechnological opportunities ranging from treating agricultural and domestic wastewater as well as the conversion of H$_2$ into an energy carrier that is compatible with the existing natural gas infrastructure (Jentsch et al., 2014, Puyol et al., 2016, Lecker et al., 2017). All methane-generating organisms, including those responsible for methane generation in the human gut, are archaea. Methane production in pure culture can proceed through acetoclastic methanogenesis where acetate is converted to CO$_2$ and CH$_4$, methylotrophic methanogenesis where methylated compounds are converted to methyl-S-CoM and either reduced by H$_2$ or disproportionated to CO$_2$ and CH$_4$, or through hydrogenotrophic methanogenesis where H$_2$ reduces CO$_2$ to CH$_4$ (Costa & Leigh, 2014). Acetoclastic and methylotrophic methanogenesis is performed only by the *Methanosarcinales* and is common in anaerobic digester communities as they catalyze the last step for treatment of agricultural and domestic waste (Ferry, 1992, Fournier & Gogarten, 2008). However, the *Methanosarcinales* grow at temperatures less than 70°C and have lower growth rates.
compared to extremely thermophilic methanogens, which limits acetoclastic and methylotrophic methanogenesis to mesophilic and thermophilic archaea (De Vrieze et al., 2012).

While CO₂ reduction by hydrogen is thermodynamically feasible, this reaction becomes less thermodynamically favorable at higher temperatures (Thauer et al., 2008). Autotrophic methanogens attract biotechnological interest as they are able to use H₂ to upgrade CO₂ to methane from waste gases (Simon, 2015). While not as efficient as using H₂ directly as a fuel, it has the potential for reducing the effects of methane and CO₂ emissions, since the methane combustion emissions are immediately re-captured. Several high temperature autotrophic, methanogenic archaea have been characterized including Methano(caldo)coccus jannaschii (T_{opt} 80°C), Methanothermus fervidus (T_{opt} 83°C), Methanocaldococcus vilosus (T_{opt} 80°C), and Methanopyrus kandleri (T_{opt} 98°C). Despite the thermodynamic challenges of methanogenesis, these archaea can grow rapidly, in some cases with doubling times as under 1 h (Stetter et al., 1981, Jones et al., 1983, Kurr et al., 1991, Bellack et al., 2011). These organisms were isolated from thermal features and likely survive on the combination of H₂ from thermal fluids and H₂ production from heterotrophs (Ver Eecke et al., 2012, Topçuoğlu et al., 2016). Autotrophic methanogenesis and growth rates can be uncoupled due to factors such as temperature and hydrogen availability (Tsao et al., 1994, De Poorter et al., 2007, Ver Eecke et al., 2012). As a result, methane generation is dependent on these factors, along with reactor conditions including gas flow rates and agitation speeds (Rittmann, 2015).

Biogas upgrading to improve energy content has been limited to mesophilic and moderately thermophilic systems, as few studies have considered more thermophilic methanogens. To optimize H₂ mass transfer and methane production, several reactor configurations have been employed including trickle bed and up-flow reactors (Lee et al., 2012, Burkhardt et al., 2015). Although these studies were conducted with inocula from mesophilic anaerobic digesters, they were nonetheless able to produce high concentrations and productivities of methane. Reactor configurations that address mass transfer limitations for methane production have not been explored for more thermophilic archaea, although such approaches could provide enhanced volumetric productivities, considering the rapid growth rates of some extremely thermophilic methanogens.
Sulfur

Sulfur is an essential element in biology, existing in oxidation states ranging from -2 (H₂S) to +6 (H₂SO₄), thereby contributing to redox reactions as well as a constituent of metal-sulfur clusters in electron carriers and catalytic sites in enzymes. Sulfur is also important in the chemical industry, with nearly 70 million metric tons produced worldwide in 2016 almost exclusively recovered as a byproduct during fossil fuel processing. Sulfur is used industrially, primarily in the form of sulfuric acid, with the majority being used to produce fertilizers, but with large amounts also utilized in the processing of petrochemicals and metallic ores (Ober, 2002).

While normally present in trace amounts in nature, sulfur can become enriched in extreme environments, such as those inhabited by extremely thermophilic archaea. The high temperatures and pressures characteristic of hydrothermal features separate solutes based on volatility. Less volatile salts are concentrated in the liquid phase and result in discharges of brackish water, while nearby vapor-dominated discharges contain mostly CO₂ and H₂S, resulting in sulfur-dominated acid pools as H₂S oxidizes to sulfuric acid (Nordstrom et al., 2005). Deep-sea hydrothermal systems also release substantial amounts of reduced sulfur compounds in the form of metal sulfides and hydrogen sulfide (Hannington, 1995). Sulfides are also present in the earth’s crust as metallic sulfide ores. In these specialized environments, sulfur becomes not just an important nutrient to microorganisms, but a key source of energy. Depending upon the reducing environment, sulfur compounds can serve as electron acceptors or donors for both heterotrophic and chemolithotrophic microorganisms. While the ability to use sulfur in this way is not unique to archaea, their prevalence in extreme environments means that they have adapted a variety of approaches to take advantage of sulfur where it is plentiful. The relationship between high-temperature organisms and sulfur has been known for some time (Amend & Shock, 2001). Sulfur metabolism may also have played a key role in the early evolution of life - the “Iron-Sulfur World” hypothesis suggests that early life appeared under thermoacidophilic conditions, with hydrothermal H₂S serving as the electron donor (Wächtershauser, 1988), although this theory is controversial (Bada & Lazcano, 2002).

Sulfur reduction

Elemental sulfur (S⁰) is used as an electron acceptor in lieu of oxygen by a number of anaerobic thermophilic archaea, both autotrophs and heterotrophs. The reduced sulfur compounds
that result can precipitate dissolved metal ions, which has applications in biotechnology and bioremediation. Perhaps the most common energy metabolism at extremely high temperatures is the reduction of $S^0$ by heterotrophs, where electrons derived from peptides and sugars are transferred to $S^0$, generating $H_2S$ and oxidized organic fermentation products (acetate, alanine) in the process. Members of the archaeal order Thermococcales are the best-studied examples of this metabolism. As discussed earlier, some Thermococcales also grow in the absence of $S^0$, generating $H_2$ from reduced ferredoxin (Fd$_{red}$), but this growth mode is only possible on sugars, since peptide oxidation proceeds through the less reduced NADPH electron carrier, from which $H_2$ production is thermodynamically unfavorable (Schut et al., 2014). The redox-responsive transcription factor responsible for the switch from $H_2$ to $H_2S$ production in $P. furoisus$ ($T_{opt}$~100°C), SurR, has been characterized in great detail (Lipscomb et al., 2009, Yang et al., 2010, Lipscomb et al., 2017). Direct physical contact between cells and $S^0$ was found to be unnecessary for $P. furoisus$ sulfur reduction, suggesting the true substrate is soluble polysulfide, which forms abiotically under standard growth conditions (Blumentals et al., 1990). Gaps remain in the current understanding of $S^0$ reduction in Thermococcales, but some of the key players have been identified. A membrane-bound oxidoreductase complex (MBX), with homology to the MBH complex mentioned previously in connection with hydrogen production, is up-regulated during growth on $S^0$. However, membrane fractions exhibit no sulfur reduction activity and it was proposed that, given its high homology to MBH, MBX is responsible for energy conservation through unknown mechanisms, while a cytoplasmic NADPH sulfur oxidoreductase (NSR) is directly involved in $H_2S$ production (Schut et al., 2007). While knockout strains confirmed that MBX was essential for growth on $S^0$, strains lacking NSR grew similarly to the parent strain and continued to produce $H_2S$ (Bridger et al., 2011, Santangelo et al., 2011). The exact mechanism by which $S^0$ is reduced is therefore not clear at present. While Pyrococcus and Thermococcus species are the best studied, heterotrophic $S^0$ reduction is also observed in the thermophilic archaeal genera Acidilobus, Caldisphaera, Caldivirga, Desulfurococcus, Hyperthermus, Palaecoccus, Pyrobaculum, Staphylothermus, Stetteria, Thermophilium, Thermoplasma, Thermoproteus and Vulcanisaeta (Kletzin, 2006). Thermoproteus tenax ($T_{opt}$~80°C) has been reported to have the unusual ability to oxidize organic compounds completely to CO$_2$ using $S^0$ as the electron acceptor (Selig & Schonheit, 1994).
Autotrophic growth via $S^0$ reduction is carried out by some members of the extremely thermophilic archaea where hydrogen gas is the source of electrons rather than organic compounds. However, the mechanism of heterotrophic $S^0$ reduction is completely different from that of the autotrophs. The model system examined for this mode of autotrophic growth among archaea is the thermoacidophilic crenarchaeon *Acidianus ambivalens* ($T_{\text{opt}} \approx 80^\circ C, \text{pH}_{\text{opt}} \approx 2$). This microorganism uses two closely associated membrane proteins, a sulfur reductase (SR) and a hydrogenase (Hyd), to generate an electrochemical membrane gradient, similar to the mechanism found in the mesophilic bacterium *Wolinella succinogenes* (Laska et al., 2003). Members of the hyperthermophilic crenarchaeal genus *Pyrodictium* ($T_{\text{opt}} \approx 100^\circ C$) contain similar membrane complexes and are also capable of autotrophic $S^0$ reduction (Dirmeier, Keller, Frey, Huber, & Stetter, 1998), as are *Ignicoccus* species (Huber et al., 2000).

Dissimilatory sulfate reduction (DSR), the anaerobic reduction of sulfate to $H_2S$ by reduction of $H_2$ or organic compounds, is an energy yielding process, in contrast to assimilatory sulfate reduction, which encompasses the uptake and incorporation of sulfur into biomolecules. The first confirmed archaeal sulfate reducers were in the hyperthermophilic genus *Archaeoglobus* ($T_{\text{opt}} \approx 85^\circ C$) (Stetter et al., 1987), which share gene homologs for sulfate adenyltransferase (Sat), adenylyl sulfate reductase (AprAB), and sulfite reductase (DsrABD) with bacteria (Klenk et al., 1997); recent genome sequence data suggests that the archaeon *T. tenax* also has the requisite genes for sulfate reduction (Siebers et al., 2011). The Dsr A, B and C subunits from *A. fulgidus* were integral in recent work elucidating the energy conserving mechanism of dissimilatory sulfate reduction (Santos et al., 2015).

Biotechnological applications of sulfur-based reductive metabolisms face a variety of challenges: the bioenergetic yield is low compared to oxidative pathways, the product ($H_2S$) is toxic and currently plentiful in natural gas and petrochemical sources, and the feedstocks ($H_2$ and organics) are likely to be more valuable than the product, except in specialized cases. One place where sulfur reducers show great promise is for bioremediation of water contaminated with heavy metals, since the biologically produced $H_2S$ would react with soluble metal ions to produce insoluble metal sulfides (Florentino et al., 2016). Sulfur-reducing microorganisms have been commercialized for sulfate removal (SULFATEQ™) and heavy metals removal (THIOTEQ™ Metal) from wastewater by Paques (Paques, Muyzer & Stams, 2008). Biologically generated $H_2S$ for metal precipitation can be produced safely on-site from elemental sulfur, and individual metals
can be recovered by tuning the reactor pH (Sanchez-Andrea et al., 2016). Many sulfur-reducing species also exhibit other growth modes (\textit{P. furiosus} reducing protons to H\textsubscript{2}, \textit{A. ambivalens} oxidizing S\textsuperscript{0}), which may have additional industrial applications.

**Sulfur oxidation**

Autotrophic growth by the aerobic oxidation of S\textsuperscript{0} was a defining characteristic of the thermoacidophilic crenarchaeal order Sulfolobales (Brock et al., 1972). This physiological feature is no longer considered universal for the order, but is still widespread (Albers, 2014). As with S\textsuperscript{0} reduction, the model for archaeal S\textsuperscript{0} oxidation is \textit{A. ambivalens}, from which a variety of relevant enzymes have been characterized; unfortunately, a genome sequence for this archaeon is not currently available. Unlike the neutral and reduced growth conditions of S\textsuperscript{0} reducers, such as \textit{P. furiosus}, oxygenated acidic environments do not favor the formation of soluble polysulfides. As a result, organisms like \textit{A. ambivalens} must contend with S\textsuperscript{0} in its inert and minimally soluble elemental form. The key enzyme for oxidation of S\textsuperscript{0} is the cytoplasmic sulfur oxygenase reductase (SOR), though species that lack it may still oxidize other inorganic sulfur compounds. SOR disproportionates inert S\textsuperscript{0} into more reactive HSO\textsubscript{3}\textsuperscript{2-} and H\textsubscript{2}S in the presence of oxygen (Kletzin, 1989). These sulfur compounds are then oxidized through a series of energy conserving steps, either involving membrane-associated or cytoplasmic enzymes (Zimmermann et al., 1999). Membrane proteins that feed into the quinone pool (for eventual conversion to an electrochemical gradient via a terminal oxidase) include thiosulfate:quinone oxidoreductase (TQO) (Muller et al., 2004), sulfide:quinone oxidoreductase (Brito et al., 2009) and sulfite:acceptor oxidoreductase (SAOR, no gene for which has yet been identified). The cytoplasmic pathway proceeds through an adenosine-5\textsuperscript{'}-phosphosulfate (APS) intermediate via APS reductase (APSR), APS:phosphate adenyltransferase (APAT), and adenylate kinase (AK) (see Figure 2-3 and Table 2-2).

Autotrophic oxidation of H\textsubscript{2}S has been reported in another member of the Sulfolobales, \textit{Sulfolobus metallicus} (Morales et al., 2011, Morales et al., 2012), possibly proceeding through a homolog of the SQR identified in \textit{A. ambivalens}. This suggests potential applications in bioremediation, removing H\textsubscript{2}S from gas and water streams before they are released into the environment. THIOPAQ, a commercial product for biological removal of H\textsubscript{2}S from gas streams (relying on a mesophilic bacterium) has been developed by Paques and applied to both biogas and fossil fuel gas streams (Janssen, 2001). A variety of bioreactor designs have been developed for
H$_2$S removal. While most rely on mesophilic bacteria, the use of thermophilic archaea in these systems would be advantageous, since the waste streams involved are often hot (>50 °C) (Li et al., 2015).

Sulfur oxidizers also play an important role in bioleaching, since sulfur forms a passivating layer on mineral ore surfaces, blocking access of metal oxidizers and inhibiting ore solubilization. Therefore, sulfur oxidizers are often beneficial contributors to microbial consortia in ore leaching, since they remove the sulfur barrier, while producing sulfuric acid to aid in solubilizing metal ions (more details in metals section). A similar role for sulfur oxidizers in the removal of sulfur impurities from coal was investigated (Peeples & Kelly, 1993), although as coal power plants are phased out in favor of natural gas, the need for this sulfur removal could decline. Bioleaching is discussed in more detail in the next section.

The end product of sulfur oxidation is sulfuric acid, which as mentioned above is a key industrial chemical. The aqueous solutions produced biologically, even by acidophiles adapted to pH 1 and below (McCarthy et al., 2015, Ai et al., 2016), are too dilute for most industrial uses, but may have applications in lignocellulosic biomass pre-treatment and ore bioleaching operations. One advantage of biological sulfuric acid production is the ability to capture energy during its production, which can be utilized by sulfur-oxidizing autotrophs for carbon fixation. The resulting biomass, once separated from the acid stream, could serve as a protein supplement, or be converted to biogas by anaerobic digestion. If genetic systems were available, these autotrophs could even be engineered to produce a specific desired chemical from CO$_2$ and sulfur.

**Metals**

Metals are essential in biological systems for many reasons, but in most cases they are found at the catalytic centers of enzymes. There are several interesting aspects of metals in extremely thermophilic archaea, including the use of tungsten as a metal center in central metabolic enzymes, such as the glyceraldehyde-3-phosphate oxidoreductase (Mukund & Adams, 1996) and the aldehyde oxidoreductase (AOR) found in *P. furiosus* (Mukund & Adams, 1990), and the unexpected range of metals identified in the metalloproteome (Cvetkovic et al., 2010). From a biotechnological perspective, the ability of extremely thermoacidophilic archaea to oxidize metals is the basis for biohydrometallurgy applications related to the recovery of base, precious and strategic metals from mineral ores.
**Metal biooxidation**

Metal bio-oxidation is a particularly rare metabolic feature, limited mostly to acidophiles, whose environmental conditions favor the solubility of metal species, although some neutrophiles have been implicated in a similar process. The biochemical mechanism of bio-oxidation was first studied in the 1980’s, when a redox-active protein structure was spectroscopically identified from iron-grown cells (Ingledew, 1982, Barr *et al*., 1990, Blake *et al*., 1993). The mesophilic bacterium central to this study, *Acidithiobacillus ferrooxidans* (originally *Thiobacillus ferrooxidans*), is the most well-characterized of the metal mobilizers, with an established, working model for iron-mediated metal bio-oxidation (Quatrini *et al*., 2009, Bonnefoy, 2010). Membrane proteins shuttle electrons from extracellular ferrous iron via a cytochrome-c containing protein through a bifurcating pathway to either a type-aa3 terminal oxidase or an NAD+ reductase complex. The latter provides reducing potential, while the former provides energy via a membrane associated ATPase, and may help to alleviate acid stress in the cytosol, which is maintained at neutral pH (Quatrini *et al*., 2009). The two processes are driven by a modest electronic potential (from iron oxidation), resulting in the need for a large amount of substrate turnover and efficient fluxes to both pathways to maintain the vital pH homeostasis and provide energy for other cellular processes (Bonnefoy & Holmes, 2012).

In contrast, much less is known about iron bio-oxidation under extremely thermophilic conditions (Wheaton *et al*., 2015, Counts *et al*., 2017). To date, the organisms that have been isolated from extremely thermoacidophilic environments have been limited to the crenarchaeal order Sulfolobales, with emphasis on the genera *Sulfolobus, Acidianus,* and *Metallophashaera.* Most of our understanding of biooxidation at high temperatures has been limited to transcriptomic and comparative genomic data from only a few of these members, including *S. metallicus* (Bathe & Norris, 2007), *Metallophashaera yellowstonensis* (Kozubal *et al*., 2011), and *M. sedula* (Auernik & Kelly, 2008). The central proteins for the process are encoded primarily in the *fox* stimulon, a well-conserved set of genes that demonstrate varying levels of transcriptomic activation in the presence of different iron substrates (Bathe & Norris, 2007, Auernik & Kelly, 2008, Auernik *et al*., 2008) (see **Figure 2-3**). Particularly responsive among these genes are *foxC* and *foxD*, which likely form a membrane-bound complex with a cytochrome b domain that could presumably interact with iron as an initial electron acceptor. Additionally, *foxG* contains putative iron-sulfur domains, which could help with the flow of electrons to the multi-copper oxidase proteins. These genes are the
only ones without demonstrated synteny within the order, but appear to be conserved among the metal-mobilizers (Wheaton et al., 2015) and highly responsive in the species *M. yellowstonensis* as well as in environmental metal-transcriptomic samples from hot spring iron mats (Kozubal et al., 2011). In the final steps of the proposed pathway, electrons are either passed to a quinol pool, likely ending in a NAD\(^+\) reductase complex, or utilized to push protons from the interior to the exterior of the cell via FoxAB (against the natural gradient).

While the pathway can still be classified as hypothetical, the synteny of this stimulon among identified metal-mobilizers, and its absence among non-metal-mobilizers, provides further support for its function in iron biooxidation (Wheaton et al., 2015). Even though the proposed mechanism relies on a similar scheme to that of *A. ferrooxidans*, there are key differences which highlight the vastly different evolutionary paths of these organisms, including but not limited to differences in membrane structures (archaea lack a periplasm), the presence of cytochrome b (as opposed to cytochrome c), and a possible bifurcating membrane soluble multi-copper oxidase (Ilbert & Bonnefoy, 2013). The additional feature of thermal stability makes these proteins desirable targets for biotechnological innovations in the biomining field. The future of biomining was once described as “hot” (Rawlings, 2002), a designation that highlights the potential gains to be made from development of biomining technologies at higher temperatures.

**Biomining**

In the past century, significant focus was placed on the role of microbes in our natural geochemical environments. As a consequence, the field of biohydrometallurgy has matured to encompass all of the processes by which microbes interact with minerals. In particular, the role of microbes in mobilizing metals from metal-bearing ores has inspired the study of biomining, which involves a surprisingly large portion of the mineral processing industry. Recent estimates suggest that up to 20% of the copper and 5% of the gold produced on earth relies on the intentional application of microbes (Johnson, 2014, Watling, 2016). These two minerals also represent two distinct applications of biomining. The first is bioleaching, which is the mobilization of target metal species by the microbe (e.g. copper, iron, nickel, etc.), whereas bio-oxidation occurs when a microbe interacts with occluding minerals that hinder the mobilization of other metal species (pyrite occlusion of gold or silver) (Brierley & Brierley, 2013). A further distinction is the type of technology used in biomining: either heap or stirred-tank reactors. The former is the lower cost
option that often involves stacking of large, milled, low-quality ore in beds, which can be operated in ambient conditions with aeration and percolated fluids, while the latter is a more capital intensive option, often reserved for high-value product streams where recovery is a primary focus (Schippers et al., 2014). Both have been used successfully in industrial applications but have unique advantages, depending on the dynamics of the mineral system. In particular, stirred-tank bio-oxidation achieved early success as a means for improving gold yields via microbial pretreatment. It dates back to the 1980’s and remains a common practice (Kaksonen et al., 2014).

In this case, microbes oxidize sulfur species (the details of which are provided in the previous section) in the bulk ore material in order to improve dissolution using lixiviants downstream (e.g., cyanide) (van Aswegen et al., 2007). While this process was originally operated with mesophilic microbes, extremely thermophilic archaea can enhance the dissolution of sulfidic ores and consequently improve downstream yields (Lindström et al., 2003, van Aswegen et al., 2007, Astudillo & Acevedo, 2009, Ciftci & Akcil, 2013).

Copper has also been another focus of stirred-tank reactor technology, particularly with the adaptation of the BIOX process for a high-temperature copper extraction technology, BIOCOP (Batty & Rorke, 2006). There have been numerous demonstrations of increased dissolution of copper sulfide ores, considered recalcitrant, with thermophiles in comparison to mesophiles (Sandström & Petersson, 1997, Le Roux & Wakerley, 1998, Dew et al., 2000, d'Hughes & Foucher, 2001, Gericke et al., 2001, du Plessis et al., 2007, Norris et al., 2012, Li et al., 2014). Due to the rising value of copper, interest in thermophilic biomining ramped up in the early 2000’s, based on potential advantages at high temperatures, ranging from enhanced kinetics and favorable redox potentials (Ahonen & Tuovinen, 1990, Gericke et al., 2010) to minimizing adverse effects resulting from passivation or electron interfacial changes (Khoshkhoo et al., 2014, Crundwell, 2015). As the quality of available ores trends toward lower grades and mixed sulfides, extreme thermophiles have been investigated for their niche advantages, such as co-solubilization of other useful metals (e.g. nickel or zinc) (Sandström & Petersson, 1997, Konishi et al., 1998, Langwaldt, 2007, Gericke & Govender, 2011, Vukovic et al., 2014, Norris, 2017) or even other high-value metals, such as silver (Norris et al., 2017). Another distinct advantage is their ability to avoid the dissolution of microbiially-deleterious metal species, such as molybdenum (Romano et al., 2001, Abdollahi et al., 2014). Finally, these same organisms have shown promise in detoxifying arsenic species via oxidation in arsenopyritic sludge or in the presence of arsenic-containing copper ores.
(e.g. enargite), which are not conducive to traditional pyrometallurgy due to the chance of creating toxic metal dusts (Hita et al., 2008, Takatsugi et al., 2011).

For all of their advantages, several problems have been identified with using extremely thermophilic organisms in biomining operations (Donati et al., 2016). Specifically, gas-liquid mass transfer can be diminished at high temperatures; this could lead to limited yield in both heaps and stirred-tank reactors (du Plessis et al., 2007). Several studies have examined the role of O\textsubscript{2} and CO\textsubscript{2} on bioleaching with thermophiles, suggesting that strategies are needed to either enrich gas feed streams or improve gas transfer of both gas species (De Kock et al., 2004, Astudillo & Acevedo, 2009). For many of these organisms, bio-oxidation is achieved in chemolithoautotrophic environments, where CO\textsubscript{2} is required to build biomass and O\textsubscript{2} is essential for maintaining pH homeostasis and electronic potential (Wheaton et al., 2015). Another issue has been the inability of extreme thermophiles (in particular S. metallicus) to function at high pulp densities, which are necessary to boost the economics of extremely thermophilic biomining. Several reasons for this have been suggested, such as sheer stress and metal toxicity (du Plessis et al., 2007), while increased CO\textsubscript{2} in feed streams, serial adaptation of cultures, and particle size control have been suggested as solutions to this issue (Astudillo & Acevedo, 2008, Astudillo & Acevedo, 2009, Jones et al., 2012).

In summary, extremely thermophilic archaea have some distinct advantages in the field of biomining, ranging from selective dissolution of high-value minerals, inherent pathways for detoxification, and robust ability to handle large fluctuations in temperature resulting from low-grade gauge materials. These observations warrant further investigation, which could lead to improved processes for existing mesophilic and moderately thermophilic industrial processes or new potential applications of thermophilic metal-mobilizers. Further, these extremely thermoacidophilic archaea deserve more attention as they come from a domain that has only become a focus of microbiologists in the previous half-century. They contain a plethora of enzymes and pathways for which we have limited knowledge, and the number of new isolates is continually growing as geothermal sites throughout the world are explored.

**Enzymes for Molecular Biology**

The earliest commercial use of enzymes from extremely thermophilic archaea involved molecular biological applications, specifically for the polymerase chain reaction (PCR). DNA
polymerases from *Thermococcus litoralis* (Vent™ Polymerase) (Perler *et al.*, 1992) and *Pyrococcus furiosus* (Pfu™ Polymerase) (Lundberg *et al.*, 1991) found wide use as higher fidelity alternatives to the bacterial Taq Polymerase, but there are also more recent applications based on polymerases, ligases and inteins from extremely thermophilic archaea (see Table 2-3).

**DNA Polymerases**

Polymerases (pols) are prevalent in nature and are classified into seven families: A, B, C, D, X, Y and reverse transcriptase (RT) (Burgers *et al.*, 2001). Family A and B pols are the best studied; most archaeal pols fall into the B, D and Y families. The development of polymerase chain reaction (PCR) (Mullis *et al.*, 1986), a powerful technique to amplify a targeted DNA sequence, was a pivotal step in the emergence of molecular biotechnology and led to many other advanced methods, including DNA and RNA sequencing. Thermal cycling is key to PCR, and this was the reason why the thermostable DNA polymerase from *Thermus aquaticus* (Chien *et al.*, 1976) (or Taq polymerase) has been widely used for this purpose. Better DNA pol options than Taq pol have been sought, including those with the ability to proof-read (or possess 3’-5’ exonuclease activity), and to extend mismatched primers for longer DNA templates. To this end, as mentioned above, the archaeal polymerase from *T. litoralis* (Mattila *et al.*, 1991), and others from the genera *Pyrococcus* and *Thermococcus* were used for their more precise proof-reading ability. DNA pol from *Thermococcus kodakaraensis* KOD1 has a low error rate, high processivity and high extension rate, thus making longer DNA target amplification more accurate, up to 6 kb (Takagi *et al.*, 1997, Hashimoto *et al.*, 1999, Hashimoto *et al.*, 2001). By addition of a small amount of an archaeal family B pol, with 3’-5’ exonuclease activity, to the Taq pol reaction, improvements in PCR were obtained. A DNA pol from *Thermococcus barophilus* Ch5, when used in a mixture with Taq DNA pol, amplified 25 kb DNA templates at lower error rates than Pfu pol (Kwon *et al.*, 2016). Through mutagenesis, a low-fidelity DNA pol mutant from *P. furiosus* found use in error-prone PCR, a technique used to introduce random mutations into a gene, while retaining thermal stability and activity (Biles & Connolly, 2004).

Along these lines, the low-fidelity Dpo4 DNA pol from *S. solfataricus* was used in error-prone PCR and/or amplification of damaged DNA (McDonald *et al.*, 2006). There have been efforts to engineer the current archaeal pols to improve processivity by fusing a DNA-binding domain to one end of DNA pol, referred to as ‘chimeric DNA pol’. A non-specific DNA binding
protein (Sso7d) from *S. solfataricus* was fused to the C-terminus of *Pfu* DNA pol; the resulting fusion (Phusion™) DNA pol was 10 times more processive with 6-fold higher fidelity than the wild-type enzyme (Wang *et al*., 2004). Given its high performance and accuracy, Phusion™ DNA pol is useful for high-fidelity PCR and site-directed mutagenesis. The chimeric pol between a helix-hairpin-helix (HhH) DNA binding motif and *Pfu* DNA pol, or ‘Pfu-HhH DNA’ pol, has increased polymerization rates, increased thermal stability, and is highly resistant to salt and inhibitors, such as phenol, blood and DNA intercalating dyes (Pavlov *et al*., 2002). The chimeric Pfu-KOD1 DNA pols were made by fusing the high fidelity KOD1 DNA pol with the thermostable *Pfu* DNA pol. As a result, this fusion DNA pol has all the desired characteristics of the precursors: high fidelity, high performance and high stability at elevated temperature (Elshawadfy *et al*., 2014).

For more advanced applications, such as DNA sequencing, archaeal DNA pols were not used initially, due to their poor ddNTP/dNTP selectivity and their intrinsic 3’-5’ exonuclease activity. However, several archaeal pols were engineered for efficient usage of ddNTPs and loss of, or reduced, 3’-5’ exonuclease activity, such as variants of the DNA pols from *T. litoralis*, *P. furiosus*, *Thermococcus* JDF-3 and *Thermococcus* 9°N-7 (Gardner & Jack, 1999, Evans *et al*., 2000, Arezi *et al*., 2002). Additionally, the *Themococcus* JDF-3 DNA pol is significantly more processive than *Pfu* and *Vent* DNA pols (Arezi *et al*., 2002). This DNA pol was further engineered to incorporate fluorescently tagged dNTPs. Another variant of the archaeal DNA pol from *Thermococcus* 9°N-7, commercially available as Therminator™ DNA pol, has enhanced ability to incorporate modified nucleotides, e.g. dideoxynucleotides, ribonucleotides and acyclonucleotides (Gardner & Jack, 2002). This Therminator™ DNA pol was further engineered by amino acid substitutions to make it more versatile and applicable to use for the sequence-by-synthesis (SBS) method (Ju *et al*., 2006). Next-generation sequencing (Metzker, 2010, Mardis, 2011) requires pols with higher efficiency. To this end, a variant of *T. kodakarensis* DNA pol was engineered for one-step RT-PCR, a technique not widely used with archaeal pols due to their low cDNA synthesis activity. This variant is stable up to 100°C and is sensitive enough, compared to the retroviral reverse transcriptase, to be used in the current RT-PCR conditions (Okano *et al*., 2017). In general, archaeal DNA pols are thermostable and exhibit high-fidelity, such that they will be continue to be targets for use in next-generation DNA sequencing (Pascal *et al*., 2006).
Ligases

DNA and RNA ligases catalyze the formation of the phosphodiester bonds between the opposing 5' phosphate and the 3' hydroxyl termini of nucleic acids, and are an essential enzyme for biological processes, e.g. DNA replication, DNA recombination, genes rearrangement and DNA/RNA repairs. To date, approximately 25 archaeal ligases have been characterized (Tomkinson et al., 2006), with structures of these enzymes reported from A. fulgidus (Kim et al., 2009), P. furiosus (Nishida et al., 2006, Tanabe et al., 2014), S. solfataricus (Pascal et al., 2006), Thermococcus sibiricus (Petrova et al., 2012), Thermococcus sp. 1519 (Petrova et al., 2012), and S. zilligii (Supangat et al., 2010). Ligases are used for cloning, plasmid/fosmid library construction, Gibson assembly, and plasmid-based site-directed mutagenesis for re-circularizing linear DNA. Alternatives to Taq ligase for Gibson Assembly (Gibson et al., 2009) were sought that are capable of ligating double-stranded, cohesive- and blunt-ended fragments. This effort identified ligases from Aeropyrum pernix (Jeon & Ishikawa, 2003), S. marinus (Seo et al., 2007), Thermococcus sp.1519 (Smagin et al., 2008) and T. fumicolans (Rolland et al., 2004) that were able to join cohesive-ended fragments, and ligases from S. marinus and T. fumicolans that can ligate blunt-ended fragments. Thermococcus sp. 1519 ligase connects long cohesive ends (about 12-nucleotide overhangs), but not shorter cohesive ends (4-nucleotide) (Smagin et al., 2008). Since this enzyme is also most active at 60-70°C, it is a good candidate to replace Taq ligase in Gibson Assembly. In another application involving the ligase chain reaction (LCR) for detection of single nucleotide polymorphisms, hyperthermophilic ligases are used, since they are stable to survive the denaturing step at 95°C in the cycling protocol. For example, the DNA ligase from S. marinus has a half-life of almost 3 h at 100°C and can also catalyze both cohesive- and blunt-ended fragments (Seo et al., 2007). To date, only the DNA ligase from P. furiosus has been the focus of developing enhanced activity (Tanabe et al., 2012, Tanabe et al., 2014). Based on its structure, this ligase was modified by mutagenesis, particularly at the C-terminal helix that interacts with the oligonucleotide binding domain (OBD) and the adenylation domain (AdD), to enhance activity and a broadening of its temperature range from 20°C to 80°C (Tanabe et al., 2012). The mutant achieved maximum amplification of ligated DNA product after only 3 cycles, whereas the wild-type enzyme required 10 cycles (Tanabe et al., 2012). This result demonstrated that a rational design approach, involving the C-terminal helix with the OBD and the AdD, could be generalized to other archaeal ligase systems.
RNA ligases that are involved in RNA repair, splicing and editing are also important in molecular biotechnology. In fact, T4 RNA ligases I and II are the essential components for a subset of rapid amplification of cDNA, RNA labeling, and more recently, the preparation of miRNA sequencing libraries (Chambers & Patrick, 2015). The first characterized archaeal RNA ligase from *Pyrococcus abyssi* (Brooks et al., 2008) was active on single-stranded RNA substrates only. The RNA ligase *Methanothermobacter thermautotrophicus* (Torchia et al., 2008) can adenylate both single-stranded RNA and DNA, making it useful for 5’-adenylating single-stranded DNA adapters for construction of miRNA sequencing library. Usually, T4 DNA ligase is used for this 5’ adenylation step (Chiuman & Li, 2002). However, the adenylated product yield is low and the process is expensive. In contrast, the RNA ligase from *M. thermautotrophicus* is more efficient (Sriskanda et al., 2000) and currently available commercially. Furthermore, it was also shown that a single point mutation (K97A) completely stripped its adenylation activity, while retaining the ability to form phosphodiester bonds (Zhelkovsky & McReynolds, 2012). Another archaeal RNA ligase, from *T. kodakarensis* (KOD), has a template dependency that is better than that of *T. thermautotrophicus* RNA ligase (Zhang & Tripathi, 2017). KOD RNA ligase also possesses mismatch specificity, which is useful for RNA sequencing, since reduced background ligation will improve the fidelity of sequencing reads significantly and is tolerant of blood protein contaminant. In fact, KOD RNA ligase retains ligation activity in the presence of up to 5% human serum (Zhang & Tripathi, 2017). Even though no structural information is available for KOD RNA ligase, it has great promise for use in RNA detection (as it is demonstrated to detect Ebola RNA transcripts), RNA modification and sequencing.

**Inteins**

Inteins (*internvening protein*) are genetic elements with intervening sequence that are capable of self-splicing post-translationally and many also contain exonuclease components capable of invading DNA (Mills *et al*., 2014, Shah & Muir, 2014). It is a naturally occurring biochemical process that mediates post-translation conversion of a precursor polypeptide into a functional protein. Although inteins are found in all three domains of life, as well as in viruses and phages, they are present with the highest frequency in archaea (Novikova *et al*., 2014). Additionally, most intein-containing proteins are involved in DNA-related processing, e.g. polymerases, helicases and topoisomerases (Novikova *et al*., 2014). Initially, inteins were used for
development of self-cleaving affinity tags to use in protein expression, releasing the tag-less target upon intein splicing (Chong et al., 1998). Ideally for this application, the intein will cause minimal cleavage during protein expression, but undergo cleavage rapidly and with high specificity to release the tag-less protein once purification is completed. Understanding the mechanisms of intein function is essential. For example, the cleavage of RadA intein from *P. horikoshii*, was shown to be highly specific and rapid with addition of single-stranded DNA (Topilina et al., 2015). Inteins found in the MoaA precursor protein in *P. abysii* and a radical S-adenosylmethionine domain protein of *Archaeoglobus profundus* are trapped in their precursor proteins by formation of disulfide bond with the internal cysteine residues involved in the splicing mechanism; as such, cleavage will not occur until a thiol is added (Callahan et al., 2013). This insight inspired development of artificial inteins that utilize the di-sulfide bond to prevent premature cleavage. The split PI-Pfu intein from *P. furiosus*, though not used in protein purification, was the basis for a biosensor that can be expressed inside living cells. This intein was used to cyclize a green fluorescent protein *in vivo*, which is more stable than its linear counterpart (Iwai et al., 2001). With the development of genetics systems for extremely thermophilic archaea, the prospect of understanding intein function has improved, which will help to unlock novel applications of inteins for biotechnology.

**Enzymes for Biotransformations**

Once genome sequences of extremely thermophilic archaea became available in the mid-1990s (e.g., for *M. jannaschii* and *P. furiosus*) (Bult et al., 1996, Robb et al., 2001), the opportunity arose to identify thermophilic homologs of industrially relevant, mesophilic enzymes, already in industrial use, from mesophilic sources arose (Adams & Kelly, 1998). As a consequence, there were many reports focused on potential applications based on recombinant versions of these enzymes produced in mesophilic hosts (Adams et al., 1995). However, to date, there have not been many large-scale commercial processes using enzymes from extremely thermophilic archaea. There are several barriers to overcome, including the challenge of replacing existing enzymes entrenched in mature bioprocesses. The advent of metabolic engineering with extremely thermophilic archaea and new emerging concepts for high-temperature bioprocessing could provide further opportunities for commercial applications of enzymes from extremely thermophilic archaea. (see Table 2-4).
Protein expression

Proper folding and assembly of proteins from extreme thermophiles are presumably best performed in the native organism or in a closely related one. This had not been possible prior to the development of genetic systems for extremely thermophilic archaea, but there have been significant developments in this area (see section on Genetics). For example, a subunit of a chitinase from *T. kodakarensis* was overexpressed leading to an increase in chitinase activity over the protein when produced in a mesophilic host (Takemasa *et al.*, 2011). A His$_6$-tagged *P. furiosus* RNA polymerase was overexpressed with a gluconeogenic promoter, demonstrating inducible protein expression (Waege *et al.*, 2010). An NADPH-dependent cytoplasmic [NiFe]-hydrogenase was overexpressed in *P. furiosus*, and with the addition of a His$_9$-tag, the protein was purified to homogeneity by a single affinity chromatography step (Chandrayan *et al.*, 2015). With improvements in molecular genetics tools for extremely thermophilic archaea, the day may come when these microorganisms are utilized as overexpression hosts for proteins that require high temperatures for proper folding and function.

Glycoside hydrolases

Glycoside hydrolases (GHs) from extremely thermophilic archaea were among the first enzymes from these microorganisms to attract biotechnological interest, primarily due to their intrinsic thermostability, a desirable feature for many applications. For example, with an eye towards starch processing, many α-amylases have been characterized from *Thermococcus*, *Pyrococcus*, and *Sulfolobus* species as possible replacements for the widely used α-amylase from *Bacillus licheniformis* (Brown *et al.*, 1990, Chung *et al.*, 1995, Dong *et al.*, 1997, Jorgensen *et al.*, 1997). While many of these processes operate at pH 4.5-5.5, the pH optimum of *B. licheniformis* α-amylase is around 6, requiring pH adjustment (Chung *et al.*, 1995, Shaw *et al.*, 1999). However, a pH-stable and thermostable α-amylase was developed by using a gene shuffling approach based on the genes encoding three such enzymes from primary enrichments of deep-sea vent samples, which had greater than 85% sequence homology to those found in the Thermococcales (Richardson *et al.*, 2002). One identified variant had improved saccharification properties at pH 4.5 and 90°C for starch hydrolysis, which compared favorably to the *Bacillus licheniformis* α-amylase. Ultimately, this became the basis for FuelZyme®, which was one of the first large-scale biotechnological products, other than molecular biological enzymes, based on extremely
thermophilic archaea. Amylolytic enzymes continue to be an active focus of both discovery and improvement. Multi-functional hydrolases, such as the type III pullulanase from *T. kodakarensis*, which has a hydrolytic activity to both α-1,4 and α-1,6 bonds, has promise as an industrial enzyme since its thermostability and pH optimum are consistent with traditional starch processing conditions (Ahmad et al., 2014). A key limitation for any industrial enzyme candidate is the inability to obtain multi-gram per liter expression levels, and this has been a concern for archaeal enzymes. To address this, efforts for improving expression of an α-amylase from *P. furiosus* in mesophilic organisms, such as *Bacillus subtilis* (Jorgensen et al., 1997), *E. coli* (Wang et al., 2007) and *Bacillus amyloliquifaciens* (Wang et al., 2016), have been reported.

There has been interest in hydrolyzing polysaccharides other than starch at elevated temperatures. Microcrystalline cellulose degradation at high temperatures by enzymes is a long-sought property for lignocellulose processing (Blumer-Schuette et al., 2008). An archaeal consortium growing on filter paper at 90°C was isolated and the associated metagenome encoded a multi-domain glycoside hydrolase with a family GH5 domain, that is similar to an enzyme encoded in the extremely thermophilic bacterium, *Caldicellulosiruptor saccharolyticus* (e.g., Csac_2528) (Graham et al., 2011). However, the metagenome had no GH48 or Carbohydrate Binding Modules from family 3 (CBM3) domains, normally associated with crystalline cellulose degradation, and filter paper dissolution occurred relatively slowly at 90°C.

Chitin is the second most abundant polysaccharide after cellulose, but its use as a feedstock is limited by the inability to hydrolyze it into simple sugars. Given the marine biotopes of many extremely thermophilic archaea, it is not surprising that several chitinases have been characterized from these microorganisms (Tanaka et al., 1999, Gao et al., 2003, Tanaka et al., 2003). Recently, the first thermophilic chitinase able to hydrolyze the reducing end of chitin was reported in *Thermococcus chitonophagus*, potentially expanding the opportunities for using this material (Horiuchi et al., 2016). *Thermococcus* sp. strain 2319X1, a hyperthermophilic archaeon isolated from a primary enrichment from a tidal vent off the coast of Kunashir Island, can grow on xylan as the sole carbon source, although the enzymatic basis for this capability is not yet clear. However, a novel multi-domain glycoside hydrolase was identified in this species, containing GH5-GH12-GH12-CBM2-CBM2 domains, that had low activity on xylan but was highly active on mixed β-1,3 and β-1,4 glucans (Gavrilov et al., 2016).
Other applications involving glycoside hydrolases involve *in vivo* biocatalysis. Recently, a *P. furiosus* based microbial fuel cell demonstrated electricity production on maltose at a power density of 225 mW m$^{-2}$ (Sekar *et al.*, 2017). While the power density is low, this proof-of-concept could potentially lead to exploiting the glycoside hydrolase inventory from extremely thermophilic archaia to generate electricity from complex polysaccharides.

**Emerging applications of β-glucosidases**

β-glucosidases catalyze the hydrolysis of β-1,4 linkages in glucans and have been well characterized in *Pyrococcus, Thermococcus, and Sulfolobus* species (Cady *et al.*, 2001, Lebbink *et al.*, 2001, Moracci *et al.*, 2001). While these enzymes have applications in biomass processing, applications to food processing and biofuel production have emerged. The CelB from *P. furiosus* produced in *Pichia pastoris* hydrolyzed 90% of the lactose in milk within 30 minutes at 65 °C (Li *et al.*, 2013). Recently, exo-β-glucosidases have been shown to have broad specificity for glycoconjugated substrates that could be converted into products that improve human health (Shin *et al.*, 2013, Shin *et al.*, 2015, Shin *et al.*, 2017). Production of flavonoids and ginsenosides from citrus and ginseng extracts, respectively, has been demonstrated using the β-glucosidases from *P. furiosus* and *S. solfataricus*. To overcome kinetic limitations, semi-rational design of the β-glucosidase from *S. solfataricus* identified a W361F mutation that improved activity on ginseng extracts for Compound K conversion by over 4-fold (Shin *et al.*, 2017). Furthermore, this mutant also had a higher activity on flavonoids, indicating that promiscuity of this engineered hydrolase can be exploited for several nutritional products (Shin *et al.*, 2017). Steryl glucosides (SGs) are insoluble contaminants made as a byproduct of biodiesel that can cause it to fail quality tests. The β-glucosidase from *Thermococcus litoralis* removed nearly all of the SGs within 2 hours (Peiru *et al.*, 2015), suggesting that enzymatic treatment, instead of distillation, could be used to reduce contaminant levels during biodiesel purification.

**Wastewater treatment and organophosphorus decontamination**

Carbamate kinases have gained interest for their ability to sequester ammonia and carbon dioxide from wastewater as carbamate phosphate, which can serve as a building block molecule for chemical synthesis. Carbamate kinases from *Themococcus barophilus* and *T. sibiricus* had high activity in alkaline pH, indicating their potential applications for enhanced CO$_2$ sequestration in ammonia (Hennessy *et al.*, 2017). Ferritin serves as an iron storage protein that is found in all three domains of life. Ferritin from *P. furiosus* has been demonstrated as a cage for iron nanoparticle formation, which has a high affinity for phosphate and arsenate (Sevcenco *et al.*, 2015). Furthermore, ferritin from *Archaeoglobus flugidus* was shown to act as a cage for GFP fusion proteins (Tetter & Hilvert, 2017), which holds promise for pharmaceutical applications as well as immobilization of polishing enzymes in water treatment. In addition, *P. furiosus* was recently shown to contain a completely new type of iron-storage protein termed IssA (for iron-sulfur storage protein A) (Vaccaro *et al.*, 2017). IssA contains iron and sulfur in the form of thioferrate, an inorganic anionic polymer previously unknown in biology. IssA forms nanoparticles reaching 300 nm in diameter and is the largest natural metalloprotein complex known. IssA nanoparticles are visible by electron microscopy as electron-dense bodies in the *P. furiosus* cytoplasm. Purified nanoparticles appear to be generated from 20 nm units containing ~6,400 Fe atoms and ~170 IssA monomers. Whether IssA has any biotechnological relevance remains to be seen.

Other proteins from extremely thermophilic archaea with biotechnological importance include SsoPox, a bifunctional lactonase and phosphotriesterase isolated from *S. solfataricus* that has been shown to have broad specificity (Hiblot *et al.*, 2012, Hiblot *et al.*, 2012). While the wild-type phosphotriesterase activity is low (Merone *et al.*, 2005), several mutants with improved activity have been developed (Merone *et al.*, 2010, Del Giudice *et al.*, 2016, Remy *et al.*, 2016). SsoPox triple mutant C258L/I261F/W263A had a 300-fold increased activity on paraoxon, a highly toxic metabolite of the organophosphate insecticide Parathion, compared to wild-type SsoPox, although the $T_{opt}$ of the mutant decreased from 80°C to 65°C (Del Giudice *et al.*, 2016). This mutant was able to quickly hydrolyze paraoxon from cotton tissues and the surface of apples without detergents, which makes it a promising enzyme for detoxification of agricultural products and chemical warfare agents. A quadruple mutant of SSo Pox was shown to have over 2,000-fold higher activity on parathion compared to wild-type, furthering the substrates on which mutagenized enzymes are active (Jacquet *et al.*, 2017). Additionally, mutagenized versions of
SsoPox have been shown to retain their activity after sterilization, exposure to solvents, and immobilization, which can expand the possible applications and processing conditions this enzyme can tolerate (Remy et al., 2016). Furthermore, the W236I mutant was shown to have its lactonase activity restored by exposure to supernatants from biofilm-forming bacteria after heat inactivation. This could expand the use of SsoPox into biofilm disruption in pipelines and as a potential antimicrobial agent (Remy et al., 2016).

**Crop stress protection**

As global temperatures increase, agricultural products must be engineered to mitigate drought and heat stressors to maintain yields. Heat stress induces the formation of reactive oxygen species (ROS). While ROS function as signaling molecules, they become toxic to cells in high concentrations due to the inability to manage the associated oxidative stress (Mittler, 2002). A new type of oxidative stress was discovered in *P. furiosus*, which was subsequently found in virtually all anaerobes, involving the enzyme superoxide reductase (SOR) (Jenney et al., 1999). In contrast to superoxide dismutase, SOR converts O$_2^-$ into H$_2$O$_2$ without generating oxygen, an important feature for anaerobic organisms. The gene encoding the *P. furiosus* enzyme (PfSOR) has been overexpressed in tobacco cell culture (Im et al., 2005), model and non-model plant systems (Im et al., 2009, Geng et al., 2016), and silkworms (Jiang et al., 2017). Progressively increasing heat stress, up to 40°C for 24 hours, on *Cornus Canadensis* with PfSOR did not adversely affect growth or recovery from heat stress. Furthermore, there was less proline accumulation and lipid peroxidation in transgenic plants, indicating PfSOR can enhance ROS scavenging in non-model plants (Geng et al., 2016). Overexpression of PfSOR in *Bombyx mori* resulted in a 30% decrease in mortality after a 44-hour exposure to heat stress of 35°C without disruption to the cocoon production rates. Additionally, the transgenic silkworms with PfSOR showed delayed mortality when subjected to starvation conditions (Jiang et al., 2017).

**Biomedical applications**

Applications related to proteases, chaperones and transcription factors from extremely thermophilic archaean are of increasing interest due to their potential biomedical implications. Sometimes features in extremely thermophilic archaean have wide-ranging evolutionary relationships with less thermophilic organisms and cells. For example, *P. furiosus* protease I (PfpI)
is a cysteine protease that has been grouped into the DJ-1/ThiJ/PfpI superfamily (Blumentals et al., 1990, Halio et al., 1996, Halio et al., 1997, Bandyopadhyay & Cookson, 2004, Larson & McPherson, 2017) of relevance in human biology and disease. Mutations in, or down-regulation of, the eukaryotic analog of PfpI, DJ-1, causes familial Parkinson’s disease (Wilson et al., 2003, Olzmann et al., 2004), tumors (Le Naour et al., 2001), and infertility (An et al., 2011).

By preventing misfolding, thermostable chaperones have the potential to serve as therapeutics in protein aggregation diseases (Glover & Clark, 2015, Peng et al., 2017). Hsp60, prefoldin, and sHsp chaperones from P. furiosus have been expressed in E. coli to confer short-term resistance to translation inhibiting aminoglycoside antibiotics (Peng et al., 2017). Prefoldin and Hsp60 improved growth of E. coli in the presence of streptomycin due to a reduction in intracellular aggregation of misfolded proteins. Furthermore, the γ-prefoldin from Methanococcus jannaschii forms highly stable filamentous structures with potential uses in tissue engineering and drug delivery (Glover et al., 2016). Archaeal transcription factors can serve as templates for antimicrobial peptides, such has been demonstrated for the VLL-28 peptide derived from the Stf76 transcription factor from Sulfolobus islandicus (Notomista et al., 2015). This peptide was recently shown to selectively induce apoptosis in murine fibroblasts and human tumor lines, expanding the antimicrobial properties to a potential cancer therapeutic (Gaglione et al., 2017).

**In vitro biotransformations**

As mentioned previously, enzymes from extreme thermophiles have also been used to construct components of *in vitro* systems that convert mixtures of simple (xylose, glucose) or complex sugars (starch, corn stover) to useful products with high productivities and yields. Using thermostable glycoside hydrolases with glycolytic enzymes has led to cell-free starch conversion for novel applications including production of sugar alcohols, such as myo-inositol (Fujisawa et al., 2017, You et al., 2017) and direct production of energy for enzymatic fuel cells (Zhu et al., 2014, Cheng et al., 2015). For all of these applications, the process is essentially the same. α-1,4 linkages of amylodextrin are broken by an α-glucan phosphorylase to produce glucose-1-phosphate. A phosphoglucomutase converts glucose-1-phosphate to glucose-6-phosphate. For energy production, glucose-6-phosphate is oxidized by NAD⁺ and G6P dehydrogenase. NADH is then oxidized by diaphorase which transfers electrons to the anode of the fuel cell (Zhu et al., 2014, Cheng et al., 2015). The proof-of-concept system yielded a consistent power density of 0.32
mW cm$^{-2}$ over a 60 hour period (Zhu et al., 2014). Isoamylase from *Sulfolobus tokodaii* was added to debranch amylopectin, resulting in a doubling of the power output from the starch fuel cell (Cheng et al., 2015). For cell-free myo-inositol production, glucose-6-phosphate is then converted to myo-inositol-3-phosphate by inositol synthase, then dephosphorylated to myo-inositol. Fujisawa and co-workers used the maltodextrin phosphorylase and phosphoglucomutase from *T. kodakarensis* and the inositol synthase from *A. fulgidis* (Fujisawa et al., 2017). A similar pathway was used with the addition of both a 4-α-glucanotransferase from *T. litoralis* to remove free glucose and an isoamylase from *S. tokodaii*, thereby eliminating the need to continuously add NAD$^+$ and facilitating scale up (You et al., 2017).

While many archaeal enzymes can be directly used as components of *in vitro* pathways, engineered variants can provide more flexibility for *in vitro* pathway design. For example, the AdhD from *Pyrococcus furiosus* preferentially uses NAD(H) over NADP(H) in the oxidation of secondary alcohols and reduction of ketones (Machielsen et al., 2006). AdhD is active on numerous secondary alcohols and ketones, and is able to reduce ketones enantioselectively (Machielsen et al., 2006). Modification of the cofactor binding pocket allowed for nicotinamide mononucleotide (NMN) to serve as the cofactor in an enzymatic fuel cell, resulting in a 40% increase in current density (Campbell et al., 2012, Solanki et al., 2016). Because of their thermostability, ability to selectively synthesize chiral alcohols, and capacity to use less expensive electron acceptors, engineered AdhD variants have applications in the synthesis of pharmaceuticals and speciality chemicals (Machielsen et al., 2006, Solanki et al., 2016).

**Genetics**

Significant strides have been made to develop and improve molecular genetic techniques for extremely thermophilic archaea in the past decade. At present, genetic manipulations have been performed in ten such archaea (see Table 2-5), including crenarchael *Metallosphaera* and *Sulfolobus* species and euryarchael *Thermococcus* and *Pyrococcus* species. While extremely thermophilic methanogenic archaea (e.g. *Methanothermus fervidus*, *Methanocaldococcus jannaschii*) are also of interest for biotechnological applications, no genetic systems have been developed, to date, in these species. Instead, archaeal methanogen engineering efforts have focused on mesophilic *Methanococcus maripaludis* (Moore & Leigh, 2005, VanDyke et al., 2009) and *Methanosarcina* species (Pritchett et al., 2004, Buan et al., 2011, Mondorf et al., 2012, Shea et al.,
2016), which are outside the scope of this review. Of the extremely thermophilic archaea with functional genetic systems, those of *P. furiosus*, *T. kodakarensis*, *S. solfataricus* and *S. acidocaldarius* stand out as being highly tractable model systems. The general molecular genetic strategies and techniques used in these systems are often able to be extended to related species. These systems also demonstrate methods for overcoming the challenges related to manipulating archaea in general. Thus, the proven methodologies and metabolic engineering successes in these model species will facilitate the development and improvement of other molecular genetic systems in a wider array of extremely thermophilic archaea. Additionally, new technologies and methodologies (e.g., DNA synthesis, transposon mutagenesis, CRISPR-Cas) are entering the molecular genetics toolbox for these organisms, enabling further advancements. Vector designs and workflow for genetic engineering, including many of the genetic tools (shuttle vectors, promoters, selectable markers, etc.) have been reviewed previously (Leigh et al., 2011, Atomi et al., 2012, Farkas et al., 2013, Loder et al., 2017, Peng et al., 2017). Here, we will focus, in general, on the challenges for developing robust genetic tools and systems in extremely thermophilic archaea. Recent discoveries and technologies enabled by genetics in these microorganisms will also be discussed.

**Genetic manipulation workflows at high temperatures**

The workflow of genetic manipulations in extremely thermophilic archaea vary slightly depending on the manipulation being performed (chromosomal deletion or insertion), whether the final modification is marked with a selectable marker, and whether the host acceptor strain can be transformed with linear or circular DNA. For most applications, genetic manipulations to the chromosome are performed in a two-step process. First, an acceptor strain is transformed with a circular DNA construct bearing a selectable marker and regions of homology to the chromosome. An intermediate strain, which has taken up this DNA by homologous recombination, is generated and identified with a selectable marker out of the population. The second step involves counter-selection for loss of the marker, resulting in either generation of the chromosomal modification or reversion to the parent genotype. Additional screening is required to identify the genetically modified strain from the revertant strain population. A final marker-less strain allows for the same marker to be used iteratively to make multiple chromosomal modifications to a strain. For transformations involving linear DNA, where there is a need to recycle the selectable marker, a
short pair of homologous DNA regions are placed on either side of the marker. After insertion of the DNA at the target location, the second step counter-selects for marker loss or “pop-out” via recombination at these short regions. Extra-chromosomal expression of a gene can also be achieved in some genetic systems via use of an autonomously replicating plasmid.

Systems of genetic selection

Selective agents coupled with corresponding selectable genetic markers for positive and negative selection strategies are one of the keystone elements to a successful genetic system. The selection and/or counter-selection ability of a marker, the selection stringency, and flexibility for selection in different media types determine the utility of the selectable marker for engineering a species. Antibiotics used widely in bacterial genetic systems are largely ineffective in archaea, so other antibiotics have been developed for archaeal genetic systems. In extremely thermophilic archaea, the choice of selection strategy is further limited by the thermal stability of both the antibiotic agent and its corresponding resistance protein, therefore only a few antibiotic strategies have been used successfully in extreme thermophiles so far. The statin-based antibiotics simvastatin, and to a lesser extent mevinolin, have been most widely used as positive selective agents, and these target the pathway involved in isoprenoid biosynthesis for archaeal phospholipid production. Overexpressing the gene for 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase confers resistance to these statin antibiotics in several systems: *P. furiosus* (Waege et al., 2010), *T. kodakarensis* (Matsumi et al., 2007), *S. islandicus* (Zheng et al., 2012). Another antibiotic strategy implemented only in *T. kodakarensis* involves negative selection on 6-methylpurine (6-MP) for the loss of a host marker gene (TK0664) involved in purine biosynthesis (Santangelo et al., 2010). Most antibiotic resistance selection methods offer the advantage of being compatible with use in rich media preparations, simplifying media formulation and potentially improving strain growth.

Because of the limited number of thermally stable antibiotics and their corresponding markers, nutritional selection strategies with auxotrophic mutant acceptor strains and defined media preparations have also been readily employed in extremely thermophilic archaea. Auxotrophic strains can be generated via selection of a random mutant or by targeted deletion employing antibiotic selection. The most widely used nutritional markers are the *pyrF* or *pyrE* genes in the uracil biosynthesis pathway, which when deleted produce uracil auxotrophic genetic
background strains. These markers are particularly useful as they can be used in both positive and negative selection strategies. Complementation of the auxotrophic strain with the marker restores uracil prototrophy, which can be selected for in defined medium lacking uracil. For strains which require further manipulation with the marker, loss of $pyrF$ or $pyrE$ can be counter-selected using 5-fluoroorotic acid (5-FOA), an analog of a uracil biosynthetic pathway intermediate. When both $pyrE$ and $pyrF$ are present, 5-FOA is converted to a compound toxic to the cells. Utilization of 5-FOA in the media allows for selection of those cells that have lost the marker via homologous recombination, and thus are no longer capable of converting to a toxic by-product. These uracil auxotrophy/5-FOA systems have been employed to modify almost all extremely thermophilic archaea to date (see Table 2-5).

Other nutritional selection auxotrophy/complementation selection strategies have also been employed. The majority of these have been used in $T. kodakarensis$, including: tryptophan (Sato et al., 2005, Santangelo et al., 2008), arginine/citrulline (Santangelo & Reeve, 2011), and agmatine (Santangelo et al., 2010). Agmatine-based selection has also been used in $S. solfataricus$ (Zhang et al., 2013), and positive selection with agamatine has the added advantage of being compatible with rich medium preparations. Histidine-based auxotrophy and complementation has been used in $T. barophilus$ (Thiel et al., 2014). Of the nutritional selection strategies, the counter-selectable $pyrF$ and $pyrE$ markers have a distinct advantage for employing schemes which necessitate recycling of markers for iterative genetic manipulations. Otherwise, application of two different markers, one for positive selection and one for negative selection, can be employed to achieve the same result.

**Growth requirement challenges during genetic manipulation**

The growth requirements of extremely thermophilic archaea present a challenge for developing a selective system, but also represent a hurdle for achieving clonal isolation of selected variants using solid medium. Agar substitutes such as Phytagel™ and Gelrite™ are used to formulate solid medium that remains solid at high temperatures, and glass petri dishes must be used instead of plastic. Incubation of plates under anaerobic high-temperature conditions can be achieved using modified metal “paint canisters”. Some anaerobic species, like $P. furiosus$, tolerate manipulations at room temperature on the bench instead of in an anaerobic Coy chamber. An additional challenge to manipulating these archaea is that colony formation requires incubation
periods of several days, regardless of their relatively rapid growth in liquid culture. \textit{P. furiosus} can take two to four days to grow distinct colonies on plates, while some \textit{Sulfolobus} species require up to seven days for colony growth. The time investment in selecting for, and clonally purifying, modified strains on plates is one of the major limitations to rapidly engineering these organisms. This increases the necessity for good transformation efficiency, stringent selection mechanisms, and well-understood foundational genetic background strains so genetic manipulations can be accomplished in a deliberate manner as opposed to one that is based on trial-and-error.

\textbf{DNA transformation barriers}

Overcoming barriers to DNA transformation are also of importance for developing a robust and efficient genetic system. Classical methodologies for transformation of microorganisms including CaCl\textsubscript{2} incubation, heat shock, and electroporation have all been used with varying success in extremely thermophilic archaea. Recently, a naturally competent \textit{P. furiosus} strain COM1 was identified (Lipscomb \textit{et al.}, 2011, Farkas \textit{et al.}, 2012), which is transformable with linear DNA and has transformation efficiencies approaching those of naturally competent bacteria. \textit{T. kodakarensis} KOD1 is also naturally competent for uptake of either linear or circular DNA, but with considerably lower transformation efficiencies than \textit{P. furiosus} COM1 (Sato \textit{et al.}, 2003, Hileman & Santangelo, 2012).

Additional barriers to transformation are DNA restriction modification systems present in the host organisms. \textit{S. acidocaldarius} restriction enzyme SuaI is known to cut unmethylated CCGG motifs (Prangishvili \textit{et al.}, 1985), while \textit{Sulfolobus islandicus} E322S restriction enzyme Suil cuts GCwGC motifs (Sollner \textit{et al.}, 2006). Plasmid DNA being transformed into \textit{S. acidocaldarius} can be methylated in \textit{E. coli} prior to transformation to prevent degradation of the transformed DNA by SuaI (Kurosawa & Grogan, 2005). Deletion of SuaI can also be performed to eliminate the need for methylation prior to transformation of \textit{S. acidocaldarius} (Suzuki & Kurosawa, 2016). SMRT sequencing technologies, which can detect certain types of DNA methylation, are a valuable tool for determining the native DNA methylation patterns that may need to be overcome for the successful transformation of various species.

Of additional concern in developing genetic systems are species with genetic insertion elements which, under certain conditions, may become mobile, potentially generating random phenotypic effects. Multiple identical insertion elements within a genome can also complicate
DNA sequence assembly, and in turn, complicate re-sequencing of modified strains. About 10% of the *S. solfataricus* P2 genome is comprised of insertion elements, and further, both *S. solfataricus* and *S. islandicus* have active mobile genetic elements (She *et al.*, 2001, Brugger *et al.*, 2004, Guo *et al.*, 2011). Thus, in the genus *Sulfolobus*, these two species are less desirable metabolic engineering hosts than *S. acidocaldarius*, which does not suffer from the abundance and mobility of various insertion elements.

**Developments enabled by improving genetic systems**

In the face of formidable barriers previously discussed, genetic systems developed in extremely thermophilic archaea have enabled recent advancements in the understanding of the physiology of these organisms. For example, deletion of reverse gyrase in *P. furiosus* demonstrated the necessity of this DNA modification protein for growth at temperatures above 95°C (Lipscomb *et al.*, 2017). The genetics of *P. furiosus* have also enabled *in vivo* study of PfAgo, an Argonaute DNA-guided nuclease (Swarts *et al.*, 2015). This PfAgo protein was recently developed into a biotechnological platform for creating ‘artificial restriction enzymes’ for cleavage of DNA at a programmable site of interest, guided by short DNA pieces (Enghiad & Zhao, 2017). This technology could be transformative to classical molecular biology approaches using restriction enzymes, potentially allowing precise control of an enzyme to cut at a specific DNA sequence of interest, programmed by a small guide DNA. Another recent development using *P. furiosus* genetics was the demonstration of the ability to construct a random mutant library (Guschinskaya *et al.*, 2016). For the library construction, *P. furiosus* genomic DNA was randomly transposed *in vitro* and these marker-bearing fragments were then transformed into *P. furiosus* COM1. The ability to perform this mutant library construction is facilitated largely by the relatively high transformation efficiency of the naturally competent *P. furiosus* COM1 strain. A random mutant library in this archaeon would enable competitive fitness screening, which could help identify genes which positively or negatively affect strain fitness under a variety of conditions (e.g. growth substrates, growth modes, stresses). Genetics and transformation procedures have also enabled the study of the native CRISPR-Cas systems of *P. furiosus* (Elmore *et al.*, 2015, Swarts *et al.*, 2015, Shiimori *et al.*, 2017) and *Sulfolobus* species (Deng *et al.*, 2013, Peng *et al.*, 2015). This understanding of the CRISPR-Cas system of extremely thermophilic archaea could enable future genetic engineering efforts in these species. For example, using a native CRISPR-Cas system that
targets mRNA in *S. solfataricus*, post-translational gene silencing could be performed with gene knockdown greater than 90% in some cases (Zebec et al., 2016). Genetics of *T. kodakarensis* have recently been used to study factor-independent transcriptional termination (Walker et al., 2017), to engineer strains for degrading and metabolizing chitin (Aslam et al., 2017), and to study DNA replication (Gehring et al., 2017). Other studies in *T. kodakarensis* have elucidated the endopeptidases which contribute to the maturation of a [NiFe]-hydrogenase in this species (Kanai et al., 2017). Recently developed genetics in *Thermococcus barophilus* and *Pyrococcus yayanosii* have enabled the study of stress responses and adaptation to piezophilic conditions in these species (Cario et al., 2015, Cario et al., 2016, Li et al., 2016). And, genetics in *S. acidocaldarius* were recently used to better understand a Crenarchaeal system for exchange of DNA (termed Ced), which is involved in DNA import and is induced by UV-light treatment (van Wolferen et al., 2016). Better understanding of DNA uptake in the crenarchaea may enable improved genetic transformation techniques for these species. The mechanisms of DNA transfer between archaea, including the Ced system, were recently reviewed (Wagner et al., 2017). The genetics methodologies in extremely thermophilic archaea have brought many new insights to these fascinating organisms which would otherwise not have been possible.

It is indeed an exciting time for the study of these archaea, with so many new genetic technologies on the horizon. The discoveries and developments, made in recent years, have been enabled primarily by the increasing tractability of genetics in extremely thermophilic archaea. Work to add new techniques to the molecular genetic toolbox of the most tractable genetic systems, and to extend these techniques to other species of archaea is ongoing. Thus, the future of using molecular genetics in extremely thermophilic archaea for basic science discovery and metabolic engineering is extremely bright.

**Metabolic Engineering**

While recruitment of individual genes and pathways from archaea for metabolic engineering applications has been employed, there has been much less done with extreme thermophiles as the platform organisms to date (Zeldes et al., 2015). In the last decade, numerous small molecule pathways have been produced on the scale of ≤ 1 gram per liter scale, along examples of protein overexpression, both homologous and heterologous (see Table 2-6). Progress in genetic system development has increased the pace at which extremely thermophilic archaea
are being utilized as platform organisms, thereby opening up new avenues to metabolic engineering.

There are examples of proteins encoded in extremely thermophilic archaeal genomes that have been recruited for moderately thermophilic metabolic engineering purposes. For example, genes encoding enzymes from *T. kodakarensis* and *P. horikoshii* were employed to construct a non-ATP-forming chimeric Embden-Meyerhof (EM) pathway expressed in the thermophilic bacterium *T. thermophilus* (Ye et al., 2012). Also, a portion of the archaeal isoprenoid ether lipid biosynthetic pathway was reconstructed in *E. coli* by expressing four genes from *Archaeglobus fulgidus* to form digeranylgeranylglyceryl phosphate (DGGGP) (Lai et al., 2009). Moving forward, cases where the extreme thermophiles themselves will be used as the platform microorganism will be reported, based on the development and application of systems biology-based tools. For example, efforts to develop genome-scale models for archaea have been described that seek to relate growth characteristics to metabolic pathways that have the potential to steer metabolic engineering efforts for substrate utilization and product formation (Ulas et al., 2012, Thor et al., 2017).

**Core metabolism**

Most extremely thermophilic archaea have a modified Embden-Meyerhof central glycolysis pathway which bypasses the NADH producing step of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and the ATP-producing step at phosphoglycerate kinase (PGK) with a single glyceraldehyde-3-phosphate ferredoxin oxidoreductase. Thus, the pathway from glucose to pyruvate is a net zero ATP process when considering only substrate phosphorylation. While utilizing such a pathway *in vivo* is not advantageous for ATP supply of cellular needs, this net zero ATP cycling can be strategic when constructing this pathway *in vitro*. By constructing an *in vitro* energy and reducing equivalent neutral pathway with enzymes from the thermophilic bacterium *T. thermophilus* and the hyperthermophilic archaea *T. kodakarensis* and *P. horikoshii*, a stoichiometric amount of lactate was produced from glucose (Ye et al., 2012).

Metabolism can be significantly influenced by environmental conditions and stress, especially temperature. Growth of *P. furiosus* below 77°C resulted in 0.5 g L⁻¹ acetoin while temperatures above 95°C, near optimal for the organism, did not result in detection of acetoin (Nguyen et al., 2016). Additionally, a lactate dehydrogenase from a bacterium with optimal growth
near 78°C was inserted into the *P. furiosus* genome under the control of a cold-shock promoter (Basen *et al.*, 2012). Approximately 0.3 g L⁻¹ lactate was produced at 72°C while virtually no lactate was detected at 98°C, demonstrating yet another example of temperature related control of metabolism in thermophilic archaeb.

**Products from biotechnology applications**

A limited number of alcohols and other small molecules have been produced within extremely thermophilic archaean hosts in recent years. In *P. furiosus*, ethanol was produced at 0.5 g L⁻¹ via reduction of its acetate by-product with an unusual ferredoxin-dependent aldehyde oxidoreductase (AOR) producing acetaldehyde (Basen *et al.*, 2014). The insertion of an AdhA from *Thermoanaerobacter* species completed the ethanol production via reduction with NADPH. This pathway was also shown to reduce exogenously-added organic acids, such as propionate and isobutyrate, producing the corresponding alcohol, demonstrating substrate promiscuity of both the AOR and AdhA. A synthetic pathway was constructed in *P. furiosus* to produce 1-butanol with four genes recruited from *Caldanaerobacter subterr anus* and one from each *Spirochaeta thermophila* and *Thermoanaerobacter ethanolicus*. Modest titers of 1-butanol (0.06 g L⁻¹) were achieved, yet this is the most complex metabolic pathway engineered into an extreme thermophile to date (Keller *et al.*, 2015). Genes for three enzymes, a multi-subunit carboxylase and two NADPH-dependent reductases, from the 3-HP/4-HB carbon fixation cycle found in *Metallosphaera sedula* were engineered into *P. furiosus* to produce 3-hydroxypropionate at 0.05 g L⁻¹ (Keller *et al.*, 2013). Further development of the pathway led to the addition of a biotin protein ligase, responsible for biotinylation of the carboxylase. A nearly ten-fold increase in titer was achieved with addition to the pathway producing nearly 0.5 g L⁻¹ (Lian *et al.*, 2016). When utilizing an enzymatic pathway with activity well below the optimal temperature of the host organism, an operational strategy is necessary to ensure that the host reaches high cell density within a reasonable time frame (see Figure 2-4). This was employed for a 3-hydroxypropionate pathway with demonstrated optimal activity at 72°C, nearly 30°C below the optimal temperature of the host, *P. furiosus*. A bioreactor scale temperature strategy allowed the organisms to multiply in exponential phase to nearly the maximal density (5 x 10⁸ cells mL⁻¹), before the temperature was lowered over a short period of time to 72°C. Prior to the temperature shift, 3-HP was not detected and acetate was the primary fermentation product. Following the temperature shift, 3-HP
production commenced after a lag phase lasting approximately six hours (Hawkins et al., 2015). Transcriptomic analysis revealed that the foreign genes were expressed at very high levels during both growth and production phases, as expected given the strong, constitutive promoter used. Central metabolic genes were generally down-regulated in the parent strain during the lower temperature production phase. But, surprisingly, these genes were highly transcribed in the recombinant strain during both phases, perhaps reflecting the bioenergetic needs to produce 3-HP (Figure 2-4). While the temperature shift strategy presents an opportunity to exploit thermophily for high temperature metabolic engineering platforms, more work is needed to understand the physiological and metabolic consequences of this approach.

Viruses

Viruses associated with extremely thermophilic archaea co-exist and proliferate in the same environments as their hosts. Similar to their hosts, they demonstrate marked diversity and a reservoir of many uncharacterized enzymes and proteins (Prangishvili et al., 2017). Archaeal-specific viruses are currently classified into seventeen families, many containing thermophilic representatives. Many of these viruses have unique and previously unobserved morphologies, ranging from the champagne bottle-like Ampullaviridae (Haring et al., 2005) to the spindle-shaped Fuselloviridae and Bicaudaviridae (Krupovic et al., 2014) to the ‘tear drop’ shaped Guttaviridae (Mochizuki et al., 2011). On the other hand, some archaeal viruses closely resemble others seen in other domains of life. For instance, the globular viruses from Globuloviridae, the filamentous viruses from the families Clavaviridae, Rudiviridae, Lipothrixviridae, and Tristromaviridae closely resemble the filamentous viruses seen in bacteria and eukaryotes in structure. However, they are highly divergent with respect to their central components (e.g., utilizing dsDNA versus ssDNA or RNA) (Krupovic et al., 2018) and, in the case of the latter two families, a lipid envelope is present, derived from its hosts (Rensen et al., 2016, Kasson et al., 2017). Finally, there are numerous viruses that are dubbed ‘cosmopolitan’ in nature. This is due to their similarities to the viruses seen in other domains of life and to their ability to infect a multitude of hosts, such as the Turriviridae, which have an icosahedral capsid structure and are known to exclusively infect the Sulfolobales (Veesler et al., 2013).

Even though many archaeal morphologies have been discovered, fewer archaeal viruses have been isolated and classified than bacterial viruses (roughly 2% as many) (Ackermann &
and still this diversity of discovered morphologies continues to grow (Prangishvili et al., 2017). Beyond the detectable homologs seen among virus families in the archaea, few if any homology-based connections can be made to the better studied bacterial and eukaryotic viruses (Gliozzi et al., 1983, Iranzo et al., 2016). Additionally, there are some viruses that lack genes with any homologs (Prangishvili et al., 2006). The fact that many of these viruses are co-habitating in extremely thermophilic hosts suggests that enzymes, with biotechnological relevance, remain to be discovered.

The robustness of these viral capsids has already inspired some biotech applications for thermophilic viruses. Specifically, the *Sulfolobus* monocaudavirus 1 (SMV1) was recently investigated as a nano-trafficker capable of surviving the acidic environment of the stomach in mice, while avoiding interactions with gut microbes and endothelial cells, and potentially unwanted immune response (Uldahl et al., 2017). Further, the unique morphologies of some of these viruses may play a role in their selective uptake in certain types of cells (Uldahl et al., 2016). Another application suggested and explored is the ability of these extremely stable capsid proteins to self-assemble, forming novel viral nanoparticle materials (Steinmetz et al., 2008).

An underappreciated application of archaea could involve their CRISPR systems, which to-date are much less studied in comparison to those present in bacteria. In fact, archaea may be a better reservoir for new CRISPR applications and enzymes given that recent estimates suggest that 70% of archaeal genomes contain CRISPR systems, in comparison to 50% of bacterial genomes (Makarova et al., 2011). In contrast, a recent report suggests the numbers are dramatically lower in uncultured prokaryotes (Burstein et al., 2016). This higher prevalence appears to be a reasonable outcome of life at higher temperatures, where population sizes are lower, leading to lower viral mutation rates, and a stronger impetus for adaptive immunity; this contrasts well with mesophilic conditions where the arms-race is much faster paced and cells may not derive as much evolutionary benefit from squandering resources on a quickly circumvented route to immunity (Weinberger et al., 2012, Iranzo et al., 2013). Add to this the fact that many archaea appear to have a “primed immunity” fueled by low specificity for mismatches on some protospacers (Maniv et al., 2016, Mousaei et al., 2016), a not uncommon feature in type-I and type-III systems, the dominant systems in archaea (Makarova et al., 2011).

These systems are more complex in structure (requiring more proteins, some of which are still being characterized and understood); however in the niche world of extreme thermophiles,
some are being utilized for molecular biology pursuits, such as genome editing (Peng et al., 2015) and gene regulation (Zebec et al., 2014, Li et al., 2016) (specifically in the Sulfolobales). Nonetheless, Cas9 from *Streptococcus pyogenes* has remained a focal point of the development of Cas-mediated genome editing and gene silencing tools, which have revolutionized the field of molecular biology (Donohoue et al., 2017). Although, one advantage of the type-II system is its simplicity (well-characterized activity from a single effector protein module), the lack of a thermally stable system presents challenges in making robust systems that could be used in medical and industrial applications. Thus, some efforts have moved to identifying more thermostolerant Cas9 proteins but thus far the limit has been a protein with short-term stability at a maximum of 70°C from *Geobacillus spp* (Harrington et al., 2017). Therefore, the future of CRISPR biotechnological advancement could focus on more thermophilic organisms that appear to be a reservoir of novel CRISPR systems, whose enzymes are waiting to be better characterized for broader application.

**Lipids**

Archaeal lipids differ in structure from their bacterial and eukaryotic counterparts, in that they contain ether linkages along isoprenoid moieties; the stereochemistry of these linkages is opposite that of their bacterial/eukaryotic counterparts. The structural features of archaeal lipids impart thermal stability and attributes that inspire biotechnological applications. These lipids can be utilized to encapsulate molecules and materials as liposomes or to create monolayer films. Archaeosomes, spherical vesicles comprising archaeal lipids, have been utilized as delivery systems for drugs, genes, antigens, and proteins. Archaeal lipids have also been employed to create thin films for sensors and filtration membranes. In archaea, the lipids that comprise their membrane components are a function of the need to survive in such extreme habitats (Konings et al., 2002). More details about these biotechnological opportunities are provided below.

Lipids from extremely thermophilic archaea contain C<sub>20</sub>-C<sub>40</sub> isoprenoid chains with linkages to either glycerol and/or nonitol; these backbones may have either polar or non-polar headgroups (Hanford & Peeples, 2002). This unique chemistry imparts structural advantages of relevance to biotechnology, including enhanced stability. Such durability and robustness is the result of many attributes including tetraether structure that spans the membrane becomes rigid at high temperature, the transmembrane cyclic structures reduce fluidity of the membrane, the alkyl
covalent bonds increases the strength of the membrane at high temperature, and the significant amount of glycosyl polar head groups that stabilize the membrane through hydrogen bonding (Ulrih et al., 2009). While diethers, such as archaeols, occur in almost all archaea, tetraethers, such as caldarchaeola and calditoglycerocaldarchaeols, occur in methanogenic, thermophilic, and psychrophilic archaea (Sprott, 1992). Other structures are observed more rarely (Langworthy, 1977, De Rosa et al., 1983, Grather, 1995). For example, tetraether backbones form monolayer membranes with the different polar head groups on each side (Gliozzi et al., 1983, Jacquemet et al., 2009). Many different lipid classes are observed in archaea, including phospholipids, phosphoglycolipids, glycolipids, sulpholipids, and aminolipids (Jacquemet et al., 2009). The architecture of archaeal monolayer membranes creates opportunities for covalent linkage in the middle layer, which protects against membrane lysis. Other noteworthy features of archaeal lipids include β-D-galactofuranosyl units that rapidly hydrolyze and cyclopentane rings that increase with environmental temperature (Tomoaia-Cotisel et al., 1992, Gambacorta et al., 1995, Itoh et al., 2001).

Archaeal lipid structures improve membrane function in a variety of ways that could be of biotechnological benefit. Ether groups increase stability and resistance to hydrolysis over ester groups imparting the toleration of a broader pH range. Branching methyl groups reduce crystallization and membrane permeability. Saturated chains provide protection against oxidative degradation. The stereochemistry of the glycerol backbone provides resistance against attacks by phospholipases (Kates, 1992). In extreme thermoacidophiles and methanogens, certain components within the bipolar monolayer membrane enhance stability. For example, flexible cyclopentane units allow the membrane rigidity to remain constant with changing temperature and glycosylated lipids stabilize the membrane through glycosyl headgroup bonding (Benvegnu et al., 2008). Large sugar head groups close to the convex surface promote asymmetric orientation to aid in membrane organization, cyclic diethers aid in survival under high pressures, and the structure of the monolayer also reduces permeation of protons and small molecules (Maccioni et al., 1995, Patwardhan & Thompson, 2000, Gliozzi et al., 2001). Archaeal lipids are formed in the absence of cholesterol, and as such can be modified to provide more stability against enzymatic hydrolysis (Szoka & Papahadjopoulos, 1980, Patel & Sprott, 1999, Rethore et al., 2007, Benvegnu et al., 2013). Archaeal lipids can be prepared and stored in the presence of oxygen, sterilized by autoclaving, and freeze-dried without degradation (Brown et al., 2009, Uhl et al., 2016). All of
these features combine to make extremely thermophilic archaeal lipids biotechnologically interesting.

Archaeosomes

Archaeosomes are liposomes made of archaeal lipids and are spherical vesicles constructed at least partially of lipid bilayers. Generally defined as vesicles formed by only archaeal lipids, archaeosomes can also be utilized in combination with conventional lipids and thus termed ‘mixed vesicles’. As with liposomes, archaeosomes and mixed vesicles can be formed as either unilamellar vesicles (ULV) or multilamellar vesicles (MLV) (Jacobsen et al., 2017). Archaeosomes are typically formed by hydrating dry lipids, stirring the mixture for an extended time, and then extruding the mixture through a membrane designed for liposome creation (Parmentier et al., 2011). To test their stability in vitro, markers such as $^{14}$C-sucrose, 5(6)-carboxyfluorescein (CF) and fluorescein isothiocyanate (FTIC)-dextran can be utilized to track their successful migration to the point of interest (Patel et al., 2000, Benvegnu et al., 2005, Parmentier et al., 2011).

The aforementioned structure of the membrane-spanning tetrathers cause archaeosomes to be more stable and less permeable than other liposomes (Szoka & Papahadjopoulos, 1980, Lichtenberg & Barenholz, 1988). Unlike other liposomes, which require a large proportion of cholesterol to form, minimal or no cholesterol is needed for archaeosomes (Lichtenberg & Barenholz, 1988). Archaeosomes are much more stable in the presence of high temperature, acidic or alkaline pH, phospholipases, bile salts, and serum media (Patel et al., 2000, Benvegnu et al., 2005). They can also be formed in the presence of oxygen at physiological temperatures or lower (Patel & Chen, 2010). In vitro and in vivo studies have demonstrated that archaeosomes are safe and non-toxic in mice (Patel & Sprott, 1999, Omri et al., 2003). This biocompatibility and the superior stability of archaeosomes makes them a promising candidate application in gene therapy and vaccine and drug delivery.

Drug delivery

Archaeal lipids show promise as candidates for improving upon the drug delivery methods developed for liposomes that have demonstrated stability issues. Oral delivery of therapeutic proteins, peptides, and small molecules must initially resist the harsh, acidic environment of the
digestive tract (Jacobsen et al., 2017). The use of lipids from extreme thermoacidophiles, organisms that must maintain their membrane integrity for their very survival, are excellent candidates for this task (Parmentier et al., 2011). For example, archaeosomes made from methanogen lipids were used to encase Paclitaxel, a small molecule drug utilized in chemotherapy (Alavi et al., 2014). The archaeosome treatment was found to increase the therapeutic value of the drug over its standard administration. Archaeosomes, 400 nm in diameter, have been prepared from the extreme thermophile *Aeropyrum pernix* for encasement of molecules of various sizes (Zavec et al., 2014). For delivery of indicator molecules like calcein into epithelial cells, 40% effectiveness has been reported.

Utilization of pure archaean lipids may be prohibitively expensive, but their improved properties could be advantageous in lower fractions as stabilizers. A liposome mixture of egg phosphatidylcholine (EPC), cholesterol and the total lipid extract (TLE) from *Sulfolobus islandicus* was examined in the presence of simulated intestinal fluid at various taurocholate (a bile salt) levels. The mixture containing 18% TLE from *S. islandicus* showed minor reduction in loss of calcein, the fluorescent marker utilized for tracking (Jensen et al., 2015). In a study with rats, a statin was encapsulated in a mixture of (EPC) and specific lipid components derived from *S. acidocaldarius*, specifically dipalmitoyl phosphatidylcholine (DPPC) and glycerlcaldityl tetraether (GCTE). Radiolabeled tetraether lipids were incorporated into the archaesome mixture with radioactivity tracked after oral administration as a proxy for liposome degradation. The mixtures with increased GCTE to DPCC ratios outperformed the mixture of EPC and cholesterol. Addition of a synthetic analog of DPPC, produced by adding an amine containing branch to each end to impart a cationic nature, resulted in nearly five times lower plasma membrane concentrations of the radiolabeled statin (Parmentier et al., 2011).

**Gene delivery**

Many lipid-DNA complexes have been produced for lipofection, or the transfection of DNA into cells through the use of liposomes. However, the lipids employed, such as DOPE or cholesterol, were often unstable and clinically unsatisfactory (Barenholz, 2001) The stability and rigidity of archaeosomes could overcome this barrier. Synthetic archaean lipids are stable enough even for oral delivery, and they have also proven to be effective enhancing lipids for *in vitro* gene delivery when compared to DOPE and cholesterol (Brard et al., 2007). To demonstrate transfection
ability, a series of diether and tetraether archaeal-like lipids were synthesized and derivatized with polyethylene glycol and a folic acid group. Utilizing this mixture, a plasmid containing a gene for luciferase (fluorescent reporter protein) expression was transfected into HeLa cells (Laine et al., 2008). The lipids derivatized with folate demonstrated improved transfection due to their ability to bind to receptors on the target cells. Thus, synthesized and derivatized archaeal lipids can be functionalized for specific targets in addition to harnessing their inherent improved stability.

**Antigen delivery**

Antigens delivered within archaeosomes have demonstrated effective immune response and subsequent antibody production, underlining their potential as antigen delivery systems. Phagocytic cells take in archaeosomes more than three-fold faster than conventional ester-based liposomes (Tolson et al., 1996). Utilizing archaeosomes as a delivery mechanism targeting phagocytic antigen processing cells, a superior immune response was demonstrated in mice. The lipid extracts from mesophilic methanogenic archaea, *M. smithii*, *M. mazei*, and the thermophilic archaeon *Thermoplasma acidophilum* were utilized to construct separate BSA containing archaeosomes. Each displayed immune responses comparable to the effective but toxic immunopotentiator Freud’s adjuvant (Sprott et al., 1997). These archaeosomes have shown good immunostimulation, whether delivered subcutaneously, intramuscularly, or intraperitoneally. They are thus excellent candidates for the delivery of vaccines, and have also been successfully used as immunostimulatory carriers for tumor antigen cancer vaccines (Krishnan et al., 2003).

**Peptide and protein delivery**

Archaeal lipids can be used both as a matrix and a delivery system for proteins and peptides. Experiments have shown successful *in vitro* reconstitution of proteins from their native lipid bilayers into archaeosomes (In't Veld et al., 1992, Elferink et al., 1993). After their incorporation into archaeosomes, beef heart cytochrome-c-oxidase and leucine transport systems were active, showing transport functions and the creation of transmembrane potentials. Proton gradients have also been successfully generated by a bacteriorhodopsin in the main phospholipid from *T. acidophilum* and coupled to ATP synthesis in an archaeosome (Freisleben et al., 1995). These tetraether lipid-based formulations do form a matrix for proton pumps, but the rigidity of the membrane compromises their operation at high tetraether concentration. Isoprenylcysteine
carbon methyl transferase activity was also successfully demonstrated in synthetic bipolar archaeosomes (Febo-Ayala et al., 2006).

Archaeosomes could be useful in delivering peptide or proteins taken orally since they can withstand the stressful conditions of the stomach and intestines with minimal leakage. Multilamellar vesicles (MLV) liposomes, which retained more than 90% of their contents after 90 minutes at low pH, are generally more stable than unilamellar vesicles (ULV) liposomes (Patel et al., 2000). The nature of the head groups was also important, with amphiphilic phosphocholine groups contributing most to the stability of the liposome (Benvegnu et al., 2005). Small molecules exit the archaeosome more easily than large molecules, such as peptides; this is initially favorable, but could also indicate a high permeability for protons that could denature the enclosed proteins (Parmentier et al., 2011). Mixed vesicles containing archaeal lipids and a bioenhancer, such as cholylsarcosine, octadecanethiol, or TPGS 1000, are promising for the delivery of peptides, since they are stable under digestive conditions, non-toxic, and safe when delivered both orally and intravenously. Insulin delivery to mice with induced diabetes was effective with archaeosomes made from the lipids of *S. acidocaldarius*, producing a decline in glucose (Li et al., 2010). While not as effective as intraperitoneally-injected insulin, it was more effective than orally-delivered encapsulated insulin or insulin enclosed in standard liposomes, which showed no effect.

**Archaeolipid films**

Archaeal lipids can also be used to make thin films, which can be useful in the creation of sensing elements or membranes in the study of transmembrane systems. Lipids derived from the extreme thermophile, *S. solfataricus*, have been employed in conjunction with the cyclic polypeptide valinomycin, a compound highly selective for potassium over sodium ions, to create Langmuir-Blodgett films. These films have applications for potassium ion-selective membranes with applications in sensors for potassium ions (Sprott et al., 1997). The stable bipolar tetraether lipids derived from *S. acidocaldarius* are also useful in the creation of free-standing planar membranes on micropores in printed circuit board fluidics (Ren et al., 2014). Compared to other membranes, those constructed with archaeal lipids demonstrate higher resistance and better stability, which is useful in studying channel proteins and transmembrane events. This suggests biotechnological uses of tetraether lipids derived from extreme thermophiles in high-throughput
drug screening, artificial photosynthesis, microfluidics, and membrane-based lab-on-chip applications.

**Future Prospects**

Once a biological curiosity, the time is now ripe to consider extremely thermophilic archaea for biotechnological applications that go beyond their use as sources of intrinsically thermostable and thermoactive enzymes. In the post-genomics era, new tools have emerged for microbial biotechnology that have been used to examine the biochemical and physiological characteristics of extreme thermophiles, as well as their ecological roles. Molecular genetic systems and tools for several extreme thermophiles have matured to the point that they can be considered as metabolic engineering platforms. Now it is up to the imagination of biotechnologists to find ways to leverage the biological properties that define life at elevated temperatures to create new products and processes that exploit archaeal extreme thermophily.

**Acknowledgments**

This work was supported by grants to RMK and MWWA by the US Department of Energy Research ARPA-E Electrofuels Program (DE-AR0000081), the US National Science Foundation (CBET-1264052, CBET-1264053), to RMK by the US Air Force Office of Scientific Research (AFOSR) (FA9550-13-1-0236, FA9550-17-1-0268), and to MWWA by the US Department of Energy (DE-FG05-95ER20175). JMC, CTS and JRC acknowledge support from a US Department of Education GAANN Fellowship (P200A140020, P200A160061), JAC acknowledges support from a US NIH Biotechnology Traineeship (T32 GM008776-16).
Figure 2-1. *In vitro* synthetic pathway for hydrogen production

*Adapted from (Myung et al., 2014, Kim et al., 2017)*

**Abbreviations:** g1p, glucose-1-phosphate; g6p, glucose-6-phosphate; 6pgl, 6-phosphogluconolactone; ru5p, ribulose-5-phosphate; x5p, xylulose-5-phosphate; r5p, ribose-5-phosphate; s7p, sedoheptulose-7-phosphate; g3p, glyceraldehyde-3-phosphate; e4p, erthrose-4-phosphate; dhap, dihydroxyacetone phosphate; fdp, fructose-1,6-diphosphate; f6p, fructose-6-phosphate. Sucrose phosphorylase (SP)* and xylose isomerase (XI)* are mesophilic enzymes.
Figure 2-2. CO₂ fixation cycles including those found exclusively in extremely thermophilic archaea. The DC/4-HB (blue) and 3-HP/4-HB (green) are found exclusively in extremely thermophilic archaea. These cycles share many intermediates with the more ubiquitous reverse TCA (rTCA) cycle as well as the 3-HP bicycle from green non-sulfur bacteria.

Abbreviations: TCA, tricarboxylic acid; 4-HB, hydroxybutyrate; 3-HP, 3-hydroxypropionate; DC, dicarboxylate
Figure 2-3. Sulfur and iron oxidation pathways in extremely thermophilic archaeal species applicable to biomining. Biotic oxidation of ferrous iron (Fe$^{2+}$) drives the supply of ferric iron (Fe$^{3+}$) to abiotically dissolve the ore sulfides to elemental sulfur and polysulfides, rendering sulfur species available for biotic oxidation. Biotic and abiotic reactions are listed in Table 2.
Figure 2-4. Metabolic Engineering Temperature Shift Strategy Demonstrated in *P. furiosus*
(a) Heterologous pathway from *M. sedula* inserted in *P. furiosus*;
(b) Temperature shift strategy with growth phase near $T_{\text{opt}}$ of host organism and production phase at $T_{\text{opt}}$ of pathway (c) Transcriptomes revealed that central metabolism in recombinant strain was minimally affected by temperature shift compared to parent strain.

Abbreviations: Mal-CoA = Malonyl CoA Reductase; Mal SA = Malonyl Semialdehyde Reductase
<table>
<thead>
<tr>
<th>Archaeon</th>
<th>Genome/ Isolation Reference</th>
<th>$T_{\text{opt}}$</th>
<th>$pH_{\text{opt}}$</th>
<th>Anaerobe</th>
<th>Aerobe</th>
<th>Growth on α,β-glucans</th>
<th>Growth on peptides</th>
<th>CO$_2$ Fixation</th>
<th>CO Oxidation</th>
<th>Carboxydromelic Oxidation</th>
<th>Sulfur Oxidation</th>
<th>Iron Oxidation</th>
<th>H$_2$ Production</th>
<th>Genetic System</th>
<th>Biotech Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyrococcus furiosus</td>
<td>(Fiala &amp; Stetter, 1986)</td>
<td>100</td>
<td>6-7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>ME Platform</td>
<td>Platform (3-HP, ethanol, butanol, oxidize CO); source of proteases and glycoside hydrolases</td>
</tr>
<tr>
<td>Thermococcus kodakarensis</td>
<td>(Fukui et al., 2005)</td>
<td>85</td>
<td>6-7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>ME platform</td>
<td>Source of proteases and glycoside hydrolases</td>
</tr>
<tr>
<td>Thermococcus onnurineus NAI</td>
<td>(Lee et al., 2008)</td>
<td>80</td>
<td>8.5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Demonstrated use in producing H$_2$ from steel mill gas effluents</td>
<td></td>
</tr>
<tr>
<td>Methanococcus jannaschii</td>
<td>(Bult et al., 1996)</td>
<td>85</td>
<td>5-7</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Methane generation from H$_2$ and CO$_2$</td>
<td></td>
</tr>
<tr>
<td>Sulfolobus solfataricus</td>
<td>(She et al., 2001)</td>
<td>80</td>
<td>2.4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>*</td>
<td>*</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Source of proteases and glycoside hydrolases; Overexpression of biotech enzymes</td>
<td></td>
</tr>
<tr>
<td>Sulfolobus acidocaldarius</td>
<td>(Chen et al., 2005)</td>
<td>75</td>
<td>2-3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>*</td>
<td>*</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Potential ME platform</td>
<td></td>
</tr>
<tr>
<td>Sulfolobus metallicus</td>
<td>(Huber &amp; Stetter, 1991)</td>
<td>70</td>
<td>2-3</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Prolific iron and sulfur oxidizer utilized in bioleaching of ores</td>
<td></td>
</tr>
<tr>
<td>Metallosphaera sedula</td>
<td>(Auernik et al., 2008)</td>
<td>73</td>
<td>2-3</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Prolific iron oxidizer utilized in bioleaching of ores</td>
<td></td>
</tr>
<tr>
<td>Acidianus brierelyi</td>
<td>(Segerer et al., 1986)</td>
<td>70</td>
<td>2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>Prolific iron and sulfur oxidizer utilized in bioleaching of ores</td>
<td></td>
</tr>
</tbody>
</table>

*reported for some species
<table>
<thead>
<tr>
<th>Species</th>
<th>Oxidation State of Sulfur</th>
<th>Enzyme</th>
<th>Net Reaction</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CuS</td>
<td>-2</td>
<td>ABIOTIC</td>
<td>CuS + Fe^{3+} + H^{+} \rightarrow Cu^{2+} + \frac{1}{2} H_{2}S_{n} + Fe^{2+}</td>
<td></td>
</tr>
<tr>
<td>H_{2}S_n</td>
<td>0</td>
<td>ABIOTIC</td>
<td>\frac{1}{2} H_{2}S_{n} + Fe^{3+} + H^{+} \rightarrow S^{0} + Fe^{2+} + H^{+}</td>
<td></td>
</tr>
<tr>
<td>H_{2}S</td>
<td>-2</td>
<td>Sulfur Quinone Reductase (SQR)</td>
<td>H_{2}S \rightarrow S^{0} + 2H^{+} + 2e^-</td>
<td>Reduced Quinone to Terminal Oxidase</td>
</tr>
<tr>
<td>S^{0}</td>
<td>0</td>
<td>Sulfur Oxygenase Reductase (SOR)</td>
<td>2S^{0} + \frac{1}{2} O_{2} + 2H_{2}O \rightarrow H_{2}S + HSO_{3}^{-} + H^{+}</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABIOTIC</td>
<td>S^{0} + \frac{3}{4} H_{2}O \rightarrow \frac{1}{4} SO_{3}^{2-} + 4HS^{-} + 8H^{+}</td>
<td>Favored at pH &gt;7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABIOTIC</td>
<td>S^{0} + HSO_{3}^{-} \rightarrow S_{2}O_{3}^{2-} + 8H^{+}</td>
<td>Favored at pH &gt;5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adenylylsulfate Reductase (APSR)</td>
<td>HSO_{3}^{-} + AMP \rightarrow PAPS + 2e^-</td>
<td>ADP from AMP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adenylylsulfate:Phosphate Adenylyltransferase (APAT)</td>
<td>PAPS + PO_{4}^{3-} \rightarrow ADP + SO_{4}^{2-} (PAPS = Phosphoadenylyl sulfate)</td>
<td>*Adenylate Kinase (AK) transforms (2 ADP \rightarrow ATP + AMP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sulfite:Acceptor Oxidoreductase (SAOR)</td>
<td>HSO_{3}^{-} + H_{2}O \rightarrow SO_{4}^{2-} + 3H^{+} 2e^-</td>
<td>Reduced Quinone to Terminal Oxidase</td>
</tr>
<tr>
<td>HSO_{3}^{-}</td>
<td>+4</td>
<td>S_{2}O_{3}^{2-} (-2,+6)</td>
<td>2S_{2}O_{3}^{2-} \rightarrow S_{2}O_{4}^{2-} + 2e^-</td>
<td>Reduced Quinone to Terminal Oxidase</td>
</tr>
<tr>
<td>S_{4}O_{3}^{2-}</td>
<td>(+5,0,0,+5)</td>
<td>S_{4}O_{3}^{2-} (+5,0,0,+5)</td>
<td>S_{4}O_{3}^{2-} + H_{2}O \rightarrow S^{0} + SO_{2}^{2-} + S_{2}O_{3}^{2-} + 2H^{+}</td>
<td>Favored at pH &gt;4</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>Terminal Oxidase (TO)</td>
<td>2H^{+} + \frac{1}{2} O_{2} + 2 e^- \rightarrow H_{2}O</td>
<td>ATP via ATP Synthase</td>
</tr>
</tbody>
</table>
Table 2-3. Extremely Thermophilic Archaeal Enzymes used in Biotransformations

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Origin</th>
<th>Application</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Polymerases</td>
<td>Thermococcus barophilus Ch5</td>
<td>PCR</td>
<td>(Kwon et al., 2016)</td>
</tr>
<tr>
<td></td>
<td>Pyrococcus furiosus</td>
<td>PCR, Error-prone PCR</td>
<td>(Biles &amp; Connolly, 2004, Wang et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Sulfolobus solfataricus</td>
<td>Error-prone PCR</td>
<td>(McDonald et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Thermococcus kodakaraensis</td>
<td>PCR, one-step RT-PCR</td>
<td>(Elshawady et al., 2014, Okano et al., 2017)</td>
</tr>
<tr>
<td></td>
<td>Thermococcus JDF-3</td>
<td>DNA sequencing</td>
<td>(Arezi et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>Thermococcus 9°N-7</td>
<td>DNA sequencing, Sequence-by-synthesis</td>
<td>(Gardner &amp; Jack, 2002, Ja et al., 2006)</td>
</tr>
<tr>
<td>Ligases</td>
<td>Thermococcus sp. 1519</td>
<td>Gibson Assembly</td>
<td>(Smagin et al., 2008)</td>
</tr>
<tr>
<td></td>
<td>Staphylothermus marinus</td>
<td>Ligase chain reaction</td>
<td>(Seo et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Pyrococcus furiosus</td>
<td>Ligase chain reaction</td>
<td>(Tanabe et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>Methanothermobacter thermautotrophicus</td>
<td>5’-adenylation</td>
<td>(Sriskanda et al., 2000)</td>
</tr>
<tr>
<td></td>
<td>Thermococcus kodakaraensis</td>
<td>RNA sequencing</td>
<td>(Zhang &amp; Tripathi, 2017)</td>
</tr>
<tr>
<td>Inteins</td>
<td>Pyrococcus horikoshii</td>
<td>Protein purification</td>
<td>(Lennon et al., 2016)</td>
</tr>
<tr>
<td></td>
<td>Purococcus furiosus</td>
<td>Biosensor</td>
<td>(Iwai et al., 2001)</td>
</tr>
</tbody>
</table>
Table 2-4: Extremely thermophilic archaea with reported genetic systems

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Origin</th>
<th>Application</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta-Glucosidase</td>
<td><em>Pyrococcus</em>, <em>Sulfolobus</em>,</td>
<td>flavanone and ginsenoside production; lactose removal from milk; steryl glucoside removal from</td>
<td>(Li et al., 2013, Shin et al., 2013, Peiru et al., 2015)</td>
</tr>
<tr>
<td></td>
<td><em>Thermococcus spp.</em></td>
<td>biodiesel</td>
<td></td>
</tr>
<tr>
<td>Lactonase</td>
<td><em>Sulfolobus solfataricus</em></td>
<td>quorum sensing disruption, detoxification of organophosphates</td>
<td>(Del Giudice et al., 2016, Remy et al., 2016)</td>
</tr>
<tr>
<td>Superoxide reductase</td>
<td><em>Pyrococcus furiosus</em></td>
<td>reduced mortality of plants and insects to heat stress</td>
<td>(Im et al., 2009, Geng et al., 2016, Jiang et al., 2017)</td>
</tr>
<tr>
<td>Ferritin</td>
<td><em>Pyrococcus furiosus</em></td>
<td>PO₄, AsO₄ removal from water; enzyme immobilization</td>
<td>(Sevcenco et al., 2015, Tetter &amp; Hilvert, 2017)</td>
</tr>
<tr>
<td>Isoamylase</td>
<td><em>Sulfolobus tokodaii</em></td>
<td>enhanced productivity from amylopectin component of starch for chemical and energy production</td>
<td>(Cheng et al., 2015, You et al., 2017)</td>
</tr>
<tr>
<td>Glycolytic enzymes</td>
<td><em>Thermococcus kodakarenensis</em>;</td>
<td>cell-free metabolic engineering for energy or chemical production</td>
<td>(Fujiisawa et al., 2017, You et al., 2017)</td>
</tr>
<tr>
<td></td>
<td><em>Archaeoglobus fulgidus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcription factors</td>
<td><em>Sulfolobus islandicus</em></td>
<td>identification of anti-microbial and anti-cancer peptide; serve as potential templates for other</td>
<td>(Notomista et al., 2015, Gaglione et al., 2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>medical uses</td>
<td></td>
</tr>
<tr>
<td>Chaperones</td>
<td><em>Pyrococcus furiosus</em>; <em>Methanocaldococcus jannaschii</em></td>
<td>prevent protein misfolding in the presence of translation inhibitors; template for novel protein scaffolds</td>
<td>(Peng et al., 2012, Glover &amp; Clark, 2015, Glover et al., 2016)</td>
</tr>
<tr>
<td>Carbamate kinase</td>
<td><em>Thermococcus barophilus</em></td>
<td>N⁴ removal from wastewater</td>
<td>(Hennessy et al., 2017)</td>
</tr>
<tr>
<td>Fuel Cell</td>
<td><em>Pyrococcus furiosus</em></td>
<td>bioelectricity</td>
<td>(Sekar et al., 2017)</td>
</tr>
<tr>
<td>Xylanase</td>
<td><em>Thermococcus</em></td>
<td>biomass conversion</td>
<td>(Gavriloj et al., 2016)</td>
</tr>
<tr>
<td>Chitinase</td>
<td><em>Thermococcus sp.</em></td>
<td>chitin hydrolysis</td>
<td>(Horiuchi et al., 2016)</td>
</tr>
<tr>
<td>Alcohol dehydrogenase</td>
<td><em>Pyrococcus furiosus</em></td>
<td>cofactor specificity</td>
<td>(Solanki et al., 2016)</td>
</tr>
<tr>
<td>Amylase</td>
<td><em>Thermococcus sp.</em></td>
<td>commercial enzyme</td>
<td>(Richardson et al., 2002)</td>
</tr>
<tr>
<td>Organism</td>
<td>$T_{opt}$</td>
<td>Uracil / 5-FOA</td>
<td>Tryptophan / 6-MP</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Metallosphaera sedula</td>
<td>75°C</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Sulfolobus islandicus</td>
<td>78°C$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermococcus barophilus</td>
<td>85°C</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Thermococcus onnurineus</td>
<td>80°C</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Pyrococcus abyssi</td>
<td>100°C</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Pyrococcus yayanosii</td>
<td>98°C</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Pyrococcus furiosus $^*$</td>
<td>100°C</td>
<td>X X</td>
<td></td>
</tr>
</tbody>
</table>

*Naturally competent strains are available. $^a$ $T_{opt}$ values from (Grogan, 1989) and (Albers, 2014).
References: $^a$ Isolation/metabolism, $^b$ genetic methods
Table 2-6. Genetically Engineered Products and Protein Overexpression involving Extremely Thermophilic Archaea

<table>
<thead>
<tr>
<th>Organism</th>
<th>$T_{opt}$ (°C)</th>
<th>Recombinant Outcome</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sulfolobus solfataricus</em></td>
<td>80</td>
<td>cellulose degradation with overexpressed endoglucanase*</td>
<td>(Girfoglio et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sulfur Oxygenase Reductase (SOR)**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABC Class Fe-S Protein*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>membrane-associated ATPases*</td>
<td>(Albers et al., 2006, Girfoglio et al., 2012)</td>
</tr>
<tr>
<td><em>Sulfolobus acidocaldarius</em></td>
<td>75</td>
<td>manipulate pentose uptake for biomass degradation applications</td>
<td>(Wagner et al., 2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>heat stable green fluorescent protein expression**</td>
<td>(Hence et al., 2012)</td>
</tr>
<tr>
<td><em>Sulfolobus islandicus</em></td>
<td>78</td>
<td>protein expression*</td>
<td>(Peng et al., 2012)</td>
</tr>
<tr>
<td><em>Thermococcus kodakarensis</em></td>
<td>85</td>
<td>increased chitinase production*</td>
<td>(Takemasa et al., 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>proteases*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased $\text{H}_2$ production via hydrogenase and reductase deletions</td>
<td>(Santangelo et al., 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased $\text{H}_2$ production via overexpression of hydrogenase*</td>
<td>(Kanai et al., 2015)</td>
</tr>
<tr>
<td><em>Thermococcus onnurineus</em></td>
<td>80</td>
<td>increased $\text{H}_2$ production from CO via overexpression of CO dehydrogenase*</td>
<td>(Kim et al., 2013)</td>
</tr>
<tr>
<td><em>Pyrococcus furiosus</em></td>
<td>100</td>
<td>lactate (0.3 g/L)</td>
<td>(Basen et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3-hydroxypropionate (0.4 g/L)</td>
<td>(Keller et al., 2013, Lian et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ethanol (1.1 g/L)</td>
<td>(Basen et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>butanol (0.07 g/L)</td>
<td>(Keller et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>acetoin (0.5 g/L)</td>
<td>(Nguyen et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{H}_2$ production from CO</td>
<td>(Schut et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased cytoplasmic hydrogenase expression</td>
<td>(Chandrayan et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>overexpression of tagged RNA polymerase*</td>
<td>(Waage et al., 2010)</td>
</tr>
</tbody>
</table>

*Homologous Protein Overexpression

** Heterologous Protein Overexpression
References


Paques Vol. 2017 pp. PAQUES - Biological and Wastewater Treatment.


Shin KC, Nam HK & Oh DK (2013) Hydrolysis of flavanone glycosides by beta-glucosidase from
Pyrococcus furiosus and its application to the production of flavanone aglycones from

Shin KC, Choi HY, Seo MJ & Oh DK (2015) Compound K production from red ginseng extract
by beta-glycosidase from Sulfolobus solfataricus supplemented with alpha-L-

Siebers B, Zaparty M, Raddatz G, et al. (2011) The complete genome sequence of Thermoproteus

upgrading systems. Biogas Science and Technology Vol. 151 (Guebitz GM, Bauer A,
Publishing.

Smagin VA, Mardanov AV, Bonch-Osmolovskaia EA & Ravin NV (2008) [Isolation and
characteristics of new thermostable DNA ligase from archaea of the genus Thermococcus].
Prikl Biokhim Mikrobiol 44: 523-528.

thermostable alcohol dehydrogenase (AdhD) from Pyrococcus furiosus. Biotechnol J 11:
1483-1497.

Solbak AI, Richardson TH, McCann RT, et al. (2005) Discovery of pectin-degrading enzymes and
9431-9438.

from Sulfolobus islandicus. Extremophiles 10: 629-634.

Sondergaard D, Pedersen CN & Greening C (2016) HydDB: A web tool for hydrogenase

555-566.

Sprott GD, Tolson DL & Patel GB (1997) Archaeosomes as novel antigen delivery systems. FEMS


CHAPTER 3: Lignocellulose Solubilization and Conversion by Extremely Thermophilic *Caldicellulosiruptor bescii* Improves by Maintaining Metabolic Activity

Christopher T. Straub¹,#, Piyum A. Khatibi¹,#, Jonathan K. Otten¹,
Michael W.W. Adams² and Robert M. Kelly¹

¹Department of Chemical and Biomolecular Engineering
North Carolina State University, Raleigh, NC 27695-7905

²Department of Biochemistry and Molecular Biology
University of Georgia, Athens GA 30602

#contributed equally to this work

Abstract

The extreme thermophile *Caldicellulosiruptor bescii* solubilizes and metabolizes the carbohydrate content of lignocellulose, a process that ultimately ceases because of biomass recalcitrance, accumulation of fermentation products, inhibition by lignin moieties, and reduction of metabolic activity. Deconstruction of low loadings of lignocellulose (5 g/l), either natural or transgenic, whether unpretreated or subjected to hydrothermal processing, by *C. bescii* typically results in less than 40% carbohydrate solubilization. Mild alkali pretreatment (up to 0.09 g NaOH/g biomass) improved switchgrass carbohydrate solubilization by *C. bescii* to over 70% compared less than 30% for no pretreatment, with two-thirds of the carbohydrate content in the treated switchgrass converted to acetate and lactate. *C. bescii* grown on high loadings of switchgrass (50 g/l) retained in a pH-controlled bioreactor slowly purged (τ = 80 h) with growth media without a carbon source improved carbohydrate solubilization to over 40% compared to batch culture at 29%. But more significant was the doubling of solubilized carbohydrate conversion to fermentation products, which increased from 40% in batch to over 80% in the purged system, an improvement attributed to maintaining the bioreactor culture in a metabolically active state. This strategy should be considered for optimizing solubilization and conversion of lignocellulose by *C. bescii*. 
Introduction

*Caldicellulosiruptor bescii* can co-metabolize five-carbon (C5) and six-carbon (C6) sugars derived from plant biomass (Blumer-Schhuete et al. 2014), and rapidly and extensively convert lignocellulose derived complex carbohydrates, such as arabinan (C5), mannan (C6), galactan (C6), xylan (C5) and cellulose (C6), into fermentation products (Zurawski et al. 2015; Zurawski et al. 2017). *C. bescii* has been engineered to produce ethanol from switchgrass (Chung et al. 2014). Furthermore, improvements in the genetics toolbox for this bacterium (Lipscomb et al. 2016; Williams-Rhaesa et al. 2017) bode well for developing strains with enhanced carbohydrate degradation capacity (Conway et al. 2016) and increased levels of metabolically engineered products (Williams-Rhaesa et al. 2018). The complex structure of plant biomass, primarily the heterogenous lignin matrix, exacerbates the microbial degradation of the carbohydrate content present in cellulose and hemicellulose, even at low biomass loadings, that otherwise are readily deconstructed in their purified forms. For *C. bescii*, in the absence of pretreatment, carbohydrate solubilization at low loadings (5 g/l) is typically below 40% (see Table 3-1), but can be as high as 70% for transgenic switchgrass lines, if alternating and multiple hydrothermal and microbial treatments are used (Zurawski et al. 2017). At high loadings (50 g/L) in batch fermentations, *C. bescii* solubilized 60% of microcrystalline cellulose, but only 20% of unpretreated switchgrass (Basen et al. 2014); microbial activity was limited by the formation of fermentation products (primarily acetate), an unidentified inhibitor generated from the plant biomass, and transition of the culture from exponential to stationary phase. Herein, new data on switchgrass solubilization following chemical (mild alkali) pretreatment is reported, in addition to a bioprocessing strategy to slowly and continuously purge inhibitory fermentation products, cells, and soluble lignin residuals to maintain metabolic activity of the bioreactor *C. bescii* culture for fermentation of lignocellulosic carbohydrates.

Since lignin is the primary barrier to the microbial degradation of plant biomass, pretreatments aimed at lignin reduction are used to decrease recalcitrance (Kumar et al. 2009). Alkali treatments can selectively dissolve lignin, leaving the cellulose component nearly completely intact with minimal degradation of hemicellulose, as measured by the xylan and arabinan content (Chen et al. 2013; Gupta and Lee 2010). In most microbial based lignocellulose bioconversions, the biomass product following alkali treatment requires further processing with cellulases and hemicellulases to generate fermentable sugars. However, *C. bescii* produces an array
of multi-domain glycoside hydrolases which are highly active on complex polysaccharides, such as those found in cellulose and hemicellulose, rendering addition of exogenously added enzymes unnecessary (Conway et al. 2018). The use of mild alkali prior to C. bescii degradation and conversion of switchgrass was, therefore, examined in order to determine if this could increase carbohydrate release and conversion. To reduce recalcitrance, switchgrass was pretreated with 0.03, 0.06, and 0.09 g NaOH /g biomass at 95°C for 1 h, and then washed with water to remove residuals. For the two abiotic controls of switchgrass processed without sodium hydroxide (with and without heat treatment), carbohydrate content did not change. However, for the 0.03, 0.06, and 0.09 g NaOH / g biomass alkali pretreatments, carbohydrate content of the post-treatment material increased from 62% to 67%, 69%, and 76% respectively, as the non-carbohydrate fraction was removed by the alkali more so than carbohydrate content. In addition to carbohydrate enrichment, sodium hydroxide pretreatments can remove acetyl groups from hemicellulose, which is supported in this case by lower acetate generation in the abiotic controls.

Following alkali pretreatment of the biomass, C. bescii had higher initial growth rates and higher maximum cell densities in line with the severity of pretreatment (Figure 3-1A). More importantly, alkali pretreatment enhanced carbohydrate solubilization and conversion (Figure 3-1B). For the 0.09 g NaOH / g biomass pretreatment, more than 70% of the carbohydrate in the pretreated switchgrass was solubilized during fermentation, compared to less than 30% for the unpretreated material, while 27 mM organic acid (acetate and lactate) was generated compared to 10 mM for the unpretreated material (Figure 3-1C). The more severe alkali pretreatments resulted in carbohydrate loss (C5 - 31%, C6 – 10%), but removed 55% of the inert material. Without alkali pretreatment, removal of C5, C6, and inert components was minimal during the washing steps (5-8%), but more than 60% of the carbohydrate still remained in the biomass, even after fermentation by C. bescii. Thus, alkali pretreatment resulted in more carbohydrate solubilized and converted to fermentation products by C. bescii despite the upfront loss of C5 and C6 content (Figure 3-1D-F).

Previously, solubilization of unpretreated switchgrass by C. bescii on the 10 L batch scale at 50 g/L loading ceased at 20%, as mentioned above (Basen et al., 2014). To address this issue here, the unpretreated feedstock was retained by using an outlet filter (200 mesh) in a pH-controlled bioreactor, allowing the liquid contents to be slowly and continuously turned over and thereby purging fermentation products, inhibitors and cells from the vessel (Supplementary
The objective was to maintain the bioreactor culture in a metabolically active state with minimal loss of feedstock. This was in contrast to batch studies, from previous reports and from this work, where *C. bescii* grew rapidly to reach its maximal cell density after approximately 72 - 96 h, after which biomass was still solubilized by enzymes in the secretome but the hydrolyzed carbohydrate was not converted to fermentation products. For the continuously liquid purged bioreactor, following *in situ* washing of switchgrass (four reactor volumes) at 70°C, fermentation was initiated by inoculating at $10^7$ cells/mL with the volumetric turnover rate ($\tau$) of 80 h. Initial cell growth was rapid, increasing to approximately $2 \times 10^8$ cells/mL after 48 h, reflecting utilization of the most readily accessible carbohydrate by *C. bescii* (Figure 3-2). For comparison, the cell densities in the batch mode reactor, under the same conditions but in the absence of purging, stabilized at $3 \times 10^8$ cells/mL within 4-5 days, while in the purged bioreactor cell densities peaked at similar densities between 7-10 days. Acetate volumetric productivity (Supplementary - Figure 3-S2) was utilized as an indicator for metabolic activity. In batch mode, acetate productivity was at the same level as the abiotic control after approximately 7 days (less than 10 mg/L/h). In contrast, acetate productivity in the continuously liquid purged bioreactor peaked at 50-60 mg/L/h, and even at 28 days was still significantly above the abiotic control. By continuously purging the bioreactor of liquid yet retaining the biomass substrate, metabolic activity was retained, even as the biomass became more recalcitrant as the most accessible carbohydrates were consumed. The composition of the residual switchgrass, following either biotic or abiotic processing, was similar to the starting material (approximately 60-70% carbohydrate) (Supplementary - Figure 3-S3). However, more than twice as much of the solubilized carbohydrate was converted to fermentation products in the liquid purged compared to batch mode. That is, assuming that solubilized carbohydrate was stoichiometrically converted to acetate by *C. bescii*, 40% of theoretical conversion (solubilized carbohydrate only) was observed for batch operation, while over 80% of theoretical was noted when the bioreactor was continuously liquid purged (Table 3-2). Furthermore, purging improved overall carbohydrate solubilization from 29% for batch operation to 41%. Thus, by slowly turning over bioreactor contents, *C. bescii* remained in a metabolically active state, and as a consequence, significantly improved switchgrass solubilization and conversion.

While it has been demonstrated that *C. bescii* can access, solubilize and convert significant amounts of the carbohydrate content of unpretreated plant biomass, even at high loadings, the
efficacy of this bacterium producing bio-based chemicals at industrial scales requires further improvements. We show here that mild alkali pretreatment of the biomass, but without the need for added enzymes, is a possible route. More promising is the processing strategy that involves purging of inhibitory components from the bioreactor to maintain metabolic activity. The combination of this approach with transgenic rather than natural feedstocks, whose composition and lignin structure better align with the physiology and metabolism of the *C. bescii*, is the next step and such studies along these lines are currently underway.
Methods and Materials

Harvested switchgrass (variety - Cave-in-Rock) was obtained from the National Renewable Energy Laboratory, Golden, CO. The switchgrass was ground on a Wiley Mill with a 20-mesh screen and then sieved to collect the 40/80 mesh fraction. The unwashed switchgrass was utilized for the treatments specified and for the bioreactor trials, both batch and continuously liquid purged.

Growth of Microorganisms

*C. bescii* (DSMZ6725) was grown in modified DSMZ 671 medium (0.25 g/L sodium sulfide being replaced by 1 g/L cysteine HCl). All trials, both biotic and abiotic, were conducted at 70°C. Serum bottle experiments were performed in a sealed serum bottle (150 mL total volume) with 50 mL media. To transition *C. bescii* onto complex substrates, freezer stocks were recovered on 5 g/L cellobiose and passaged onto 5 g/L biomass with 0.5 g/L cellobiose and further passaged to 5 g/L biomass as the sole carbon source. This culture was then utilized as the inoculum for both the serum bottle and bioreactor experiments.

Mild alkali treatment

The unwashed milled switchgrass was utilized for the pre-treatment. The unwashed switchgrass (5 g) was added to a serum bottle with 40 mL of deionized water. The switchgrass was allowed to soak for 1 h at 4°C. After 1 h of soaking in DI water, a 5.0 M sodium hydroxide (NaOH) solution was added to bring the sodium hydroxide loading to the equivalent of 0.03, 0.06, and 0.09 g NaOH / g biomass. For 5 g switchgrass at 0.00, 0.03, 0.06, and 0.09 g NaOH / g biomass this is 0.0 mL, 0.75 mL, 1.5 mL, and 2.25 mL of 5.0 M NaOH. Water was added to bring the total volume of the mixture to 50 mL for a 100 g/L treatment loading. The serum bottles were sealed and transferred to a 95°C bath shaking at 150 RPM for 1 h. A control with no sodium hydroxide but no heat treatment (No HT Ctrl) was similarly prepared and incubated by shaking at 150 RPM and 25°C for 1 h. After 1 h, the bottles were removed and placed in room temperature water to cool for 15 minutes. The resulting biomass was removed from the serum bottles and transferred to 50 mL conical tubes and centrifuged at 4,700 x g for 10 minutes in a swinging bucket centrifuge to pellet the biomass. The supernatant was carefully withdrawn from the pelleted biomass and discarded. Deionized water was added to the sample to bring the volume to 50 mL, after which the
pellet was re-suspended in the conical by shaking vigorously. This sample was again centrifuged; this washing step was repeated two additional times. The final pelleted biomass was transferred to tared aluminum weigh boats for drying at 50°C for 48 hours. The resulting treated material was analyzed for composition and also utilized in cultures (at 5 g/L) as the only carbon source.

**Bioreactor Configuration**

An Applikon 3 L bioreactor was utilized for the high loadings experiment (Supplementary Figure 3-S1). The mixing shaft contained two impellers, a four blade Ruston impeller on the bottom of the shaft (about 3 cm from the dished bottom of the bioreactor) and a vortexing marine impeller positioned at liquid level. The continuous liquid feed to the bioreactor was controlled with a variable speed MasterFlex L/S 600 RPM peristaltic pump from a 5 L bottle of filter sterilized media (salts and vitamins only, no carbon source) which was also continuously purged with inert gas (CO₂/N₂). The liquid effluent was continuously collected via a custom dip leg, constructed from a 1” piece of stainless-steel tubing. The end internal to the bioreactor was machined to fit with a cap which held a circular (approx. 2 cm diameter) 200 mesh screen (0.0074 mm) in place. The opening on the dip leg was positioned at approximately 5 cm below the liquid level. The bioreactor working volume l was controlled at 2 L volume by a liquid contact level sensor which controlled a peristaltic pump connected to the discharge line, removing effluent from the bioreactor (through the screened dip leg). A porous polyethylene filter (200 mesh) (McMaster Carr – Part #8680T51) was placed on the effluent line outside of the bioreactor (before the pump) to trap any biomass material that may traveled beyond the single layer screen in the bioreactor. This material was removed, dried, and accounted for in calculations as unsolubilized biomass. It was not reintroduced to the bioreactor.

**Bioreactor Operation**

Two liters of 671D media were added to the bioreactor with 100 g (dry weight basis) of milled and sieved (40/80 mesh) unwashed switchgrass. Temperature was brought to 70°C with agitation at 400 RPM. After 30 min, 671D media (no carbon source) was fed to the reactor at 1,000 mL per hour (reactor turnover time of 2 h), with the level control keeping the reactor volume at 2 L. This washing continued for 8 h (or 4 reactor volumes) to wash the switchgrass until the effluent was no longer turbid or brown. The reactor was then purged at 33 mL/min of inert gas (CO₂/N₂
20/80) for 1 h to establish anaerobic conditions. The temperature, agitation, and gas purge rate were maintained throughout all bioreactor experiments. The reactor was inoculated to a cell density of approximately $10^7$ cells/mL from cultures grown in serum bottles with the same media and switchgrass. For batch reactor experiments, no liquid flushing or level control was utilized. For continuous liquid flushing experiments, immediately after inoculation, the flow rate was turned to 25 mL/hr (a reactor turnover time of 80 h). The level control was also initiated to maintain the reactor at 2 L. Samples were taken from the effluent line for cell counts and acetate concentration.

**Switchgrass Carbohydrate Content Analysis**

Switchgrass composition was determined via a modified version of NREL procedure - Determination of Structural Carbohydrates and Lignin in Biomass (NREL/TP-510-42618). Briefly, 600 μL of 72 wt% sulfuric acid was added to 40 mg of dry biomass in a 18 mm x 150 mm Balch tube. The tubes were immersed in a 30°C water bath and gently stirred with a glass rod for 1 h. The samples were diluted to 4 wt% sulfuric acid by addition of 16.8 mL deionized water. The tubes were sealed and autoclaved at 121°C for 1 h. A 1 mL sample of the recovered liquid was centrifuged to precipitate solids and was immediately analyzed for sugars via HPLC (see below). As per the NREL procedure, sugar recovery standards were utilized to account for the slow degradation of sugars in sulfuric acid.

**Acetate and Sugar Analysis via HPLC**

Acetate and sugar concentrations were determined by HPLC utilizing a 5 mM sulfuric acid (0.6 mL/min) as the mobile phase and a Rezex-ROA column (300 mm by 7.8 mm; Phenomenex). Acetate and sugars were detected via a Waters Refractive Index Detector (Model 2414).

**Acknowledgements**

This work was supported in part by the US National Science Foundation (CBET-1264052 to RMK, CBET-1264053 to MWWA), the US Department of Agriculture (2018-67021-27716), and the Bioenergy Science Center a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. CT Straub acknowledges support from a US DoEd GAANN Fellowship (P200A160061).
Figure 3-1. Alkali Pretreated Switchgrass Fermentations (5 g/L) with *C. bescii*  (A) Cell density over the course of 7-day fermentation. (B) Total biomass solubilization of pretreated and washed biomass (5 g/L) following 7-day fermentation (dark blue) and abiotic controls (gray). (C) Organic acid production (acetate and lactate) (red) with abiotic controls (gray). (B – F) X-axis: mass of sodium hydroxide per mass of biomass utilized during pretreatment step (g NaOH / g biomass). Treatment at 95°C for 1 h. No heat-treatment control (No HT Ctrl) was maintained at 25°C for 1 h. Switchgrass mass balance accounting with C5 content (D - blue), C6 content (E - green), and inert content (F - tan). Fraction of each component solubilized and discarded via washing following alkali pretreatment (bottom bar of stack), fraction of each component solubilized during biotic fermentation (middle bar of stack), and fraction of original component which remained in residual material after pretreatment and biotic fermentation (top bar of each stack).
Figure 3-2. Bioreactor-scale Lignocellulose Fermentation by C. bescii (A) Cell density over the course of 28-day fermentation on 50 g/L switchgrass. (B) Acetate levels (mM) over course of 28-day fermentation (abiotic batch mode - dotted grey line) (abiotic purged mode – dotted light blue line).
Figure 3-S1. Configuration of Continuous Purged Bioreactor for *C. bescii* Lignocellulose Fermentation  Batch trials (both biotic and abiotic) performed in same bioreactor configuration and conditions, but without addition of fresh media and purging of bioreactor liquid.
Figure 3-S2. Acetate Productivity for Gas Sparged Bioreactor Fermentation, with and without Continuous Liquid Purging
Figure 3-S3. Composition of Feedstock (switchgrass) and Residual Material from Gas Sparged Bioreactor Fermentation, with and without Continuous Liquid Purging
Table 3-1. Summary of *Caldicellulosiruptor* Species Solubilization/Conversion of Lignocellulosic Carbohydrate

<table>
<thead>
<tr>
<th>Biomass</th>
<th>Calderum</th>
<th>Pretreatment</th>
<th>Load (g/L)</th>
<th>Size (mm)</th>
<th>Reactor (L)</th>
<th>Time (days)</th>
<th>T (°C)</th>
<th>Total Carbohydrate (%</th>
<th>Solub. (g/L)</th>
<th>Conversion (mM)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switchgrass</td>
<td><em>C. sac</em></td>
<td>Washed &amp; Unwashed</td>
<td>30</td>
<td>-0.085</td>
<td>0.025</td>
<td>6</td>
<td>65</td>
<td>W 7.1 UW 3.7</td>
<td>NR</td>
<td>W 2 NR</td>
<td>(Talluri et al. 2013)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed - 78°C/18 h</td>
<td>5</td>
<td>0.177 – 0.841 (20/80 Mesh)</td>
<td>0.05</td>
<td>2</td>
<td>78</td>
<td>33 NR</td>
<td>4.7 0.2</td>
<td>27 NR 9.7 1.7</td>
<td>(Basen et al. 2014)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed at 25°C</td>
<td>5</td>
<td>0.177 – 0.841 (20/80 Mesh)</td>
<td>0.05</td>
<td>7</td>
<td>70</td>
<td>40 NR</td>
<td>7.4 0.2</td>
<td>40 NR 6.8 0.2</td>
<td>(Zurawski et al. 2015)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Autoclaved 121°C – 60 min</td>
<td>2.9*</td>
<td>14.3d</td>
<td>0.05</td>
<td>5</td>
<td>75</td>
<td>19 24 30 8.0 1.0</td>
<td></td>
<td></td>
<td>(Paye et al. 2016)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td>5</td>
<td>0.177 – 0.841 (20/80 Mesh)</td>
<td>0.05</td>
<td>12</td>
<td>70</td>
<td>30 22 26 8.2 0</td>
<td></td>
<td></td>
<td>(Zurawski et al. 2017)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38 36 45 14.1 1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40 37 36 10.2 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45 45 52 15.6 3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td>5</td>
<td>0.177 – 0.841 (20/80 Mesh)</td>
<td>0.05</td>
<td>12</td>
<td>70</td>
<td>20 0 16 5.5 0</td>
<td></td>
<td></td>
<td>(Zurawski et al. 2017)</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25 20 34 10.0 0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td>5</td>
<td>0.177 – 0.841 (20/80 Mesh)</td>
<td>0.05</td>
<td>28</td>
<td>722</td>
<td>28 75 13.5 3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35 30 40 12.6 0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42 35 36 9.5 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>43 46 57 13.5 3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 180°C/25 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>43 46 57 13.5 3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td>5</td>
<td>0.177 – 0.420 (40/80 Mesh)</td>
<td>0.05</td>
<td>7</td>
<td>70</td>
<td>36 31 24 10.0 0.2</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31 34 22 10.8 0.1</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>HT 70°C for 60 min/Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41 50 43 14.3 0.6</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>0.03 g NaOH/g Biomass 95°C - 60 min/Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53 63 55 17.0 4.5</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>0.06 g NaOH/g Biomass 95°C - 60 min/Washed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61 71 73 18.0 8.8</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td>Switchgrass</td>
<td><em>C. bes</em></td>
<td>Washed at 70°C in situ</td>
<td>5</td>
<td>0.177 – 0.420 (40/80 Mesh)</td>
<td>0.05</td>
<td>7</td>
<td>70</td>
<td>35 28 28 63 NR</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
<tr>
<td></td>
<td><em>C. bes</em></td>
<td>Purged at 1 reactor vol/2 h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45 39 40 141 NR</td>
<td></td>
<td></td>
<td>*This Work</td>
</tr>
</tbody>
</table>
Table 3-1. (continued)

<table>
<thead>
<tr>
<th>All biomasses were generated by milling harvested biomass and some were then sieved to desired particle size.</th>
<th>W = Washed; UW = Unwashed; NR = Not Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organism Abbreviations:</strong> Caldicelluliruptor bescii (Cbes); Caldicelluliruptor saccharolyticus (Csac); Caldicelluliruptor kronotskyensis (Ckro)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>a Gas sparged, pH controlled</td>
<td></td>
</tr>
<tr>
<td>b Harvested mid-season</td>
<td></td>
</tr>
<tr>
<td>c 1 g/l glucan basis substrate loading</td>
<td></td>
</tr>
<tr>
<td>d 5 g/l glucan basis substrate loading</td>
<td></td>
</tr>
<tr>
<td>e Fermentation products calculated from Zurawski et al. 2017.</td>
<td></td>
</tr>
<tr>
<td>f See methods section for NaOH treatment protocols</td>
<td></td>
</tr>
<tr>
<td>g Gas (N₂/CO₂) sparged, pH controlled</td>
<td></td>
</tr>
<tr>
<td>h Gas (N₂/CO₂) sparged, pH controlled, liquid purged</td>
<td></td>
</tr>
<tr>
<td>i Acetate concentration normalized to batch mode (total acetate production over course of run / reactor volume)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3-2. *C. bescii* solubilization and conversion of switchgrass in bioreactor

<table>
<thead>
<tr>
<th></th>
<th>Starting Mass (g/L)</th>
<th>Solub. Mass (g/L)</th>
<th>Solub. Mass (%)</th>
<th><em>a</em>Initial Carbohydrate (g/L)</th>
<th><em>a</em>Solubilized Carbohydrate (g/L)</th>
<th>Theoretical Acetate from Fermentation (g/L)</th>
<th>Total Measured Acetate (g/L)</th>
<th>Calculated Abiotic Acetate Contribution (g/L)</th>
<th>Calculated Biotic Acetate Production (g/L)</th>
<th>Theoretical Acetate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch</td>
<td>50.0</td>
<td>17.6</td>
<td>35.2</td>
<td>31.5</td>
<td>9.1</td>
<td>5.9</td>
<td>3.7</td>
<td>1.3</td>
<td>2.4</td>
<td>40</td>
</tr>
<tr>
<td>$\tau = 80$ h</td>
<td>50.0</td>
<td>22.3</td>
<td>44.5</td>
<td>31.5</td>
<td>12.5</td>
<td>8.2</td>
<td>8.3</td>
<td>1.4</td>
<td>6.9</td>
<td>85</td>
</tr>
<tr>
<td>(#1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau = 80$ h</td>
<td>50.0</td>
<td>23.2</td>
<td>46.3</td>
<td>31.5</td>
<td>13.2</td>
<td>8.6</td>
<td>8.4</td>
<td>1.4</td>
<td>7.0</td>
<td>81</td>
</tr>
<tr>
<td>(#2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch</td>
<td>50.0</td>
<td>7.5</td>
<td>14.9</td>
<td>31.5</td>
<td>0.4</td>
<td>0.2</td>
<td>1.3</td>
<td>1.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(Abiotic)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau = 80$ h</td>
<td>50.0</td>
<td>8.9</td>
<td>17.7</td>
<td>31.5</td>
<td>1.8</td>
<td>1.1</td>
<td>1.4</td>
<td>1.34</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(Abiotic)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Carbohydrate content includes glucan, xylan, and arabinan
References
Paye JM, Guseva A, Hammer SK, Gjersing E, Davis MF, Davison BH, Olstad J, Donohoe BS, Nguyen TY, Wyman CE and others. 2016. Biological lignocellulose solubilization:


CHAPTER 4: Recalcitrance and Convertibility of Transgenic Populus trichocarpa by Metabolically Engineered Caldicellulosiruptor bescii Tracks with Lignin Content

Christopher T. Straub\textsuperscript{1}, Ryan G. Bing\textsuperscript{1}, Jack P. Wang\textsuperscript{2}, Vincent L. Chiang\textsuperscript{2}, Michael W.W. Adams\textsuperscript{3} and Robert M. Kelly\textsuperscript{1}

\textsuperscript{1}Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695

\textsuperscript{2}Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695

\textsuperscript{3}Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

To be submitted to: Bioresource Technology
Abstract

Microbial conversion of lignocellulosic biomass is significantly hindered by the recalcitrance of the feedstock. In addition to thermal, chemical, and enzymatic treatments, recalcitrance can be addressed by modification of the plant secondary cell wall through genetic engineering, breeding, strategic harvest windows, and manipulation of nutrients and growth conditions. The efficacy of such strategies can be advantageously assessed through a direct assay that is based on lignocellulose-utilizing microorganisms metabolically engineered to produce an industrially relevant chemical. Here, the recalcitrance and convertibility of transgenic black cottonwood (*Populus trichocarpa*) lines were assessed through incubation with an engineered ethanol-producing *Caldicellulosiruptor bescii* strain that simultaneously solubilized the biomass and converted both hexose and pentose carbohydrate fractions to ethanol and acetate. A strong inverse correlation between conversion efficiency and lignin content was observed, although low lignin content did not necessarily determine fitness of the transgenic line. A correlation with the syringyl to guaiacyl (S/G) ratio of lignin was less evident. The S/G ratio has previously presented a positive correlation to reduced recalcitrance in lignin-modified feedstocks. Microbial conversion correlated with previous results with enzymatic saccharification for most of the transgenic lines tested. Of the monolignol biosynthetic genes targeted, coumarate 3-hydroxylase 3 (C3H3) down-regulation was the most effective in achieving high levels of *C. bescii* solubilization and conversion. One such line, with a growth phenotype similar to the wild-type control, generated more than three times the fermentation products of the control, suggesting that excellent digestibility can be achieved without compromising fitness.
Introduction

Lignocellulosic biomass contains the three most abundant polymers on earth - cellulose, hemicelluloses, and lignin – all generated via solar-powered fixation of carbon dioxide. This abundant resource is available from crop wastes, dedicated cultivated feedstocks, and sustainable harvest of forest lands to name a few sources (Lange, 2007). In addition, more than half of land-based biomass is carbohydrate in the form of biopolymers, primarily cellulose, a homogenous linear polymer of β-1,4-linked glucose. The other major carbohydrate component, hemicelluloses, is a heterogenous polymer composed of primarily xylose along with smaller amounts of arabinose, mannose, rhamnose, galactose, glucose and glucuronic acid (Kenney et al., 2013). While there are many microorganisms capable of converting these simple sugars to chemicals, such as ethanol, those that can ferment lignocellulose directly are far more limited. The key barrier to biomass conversion is the recalcitrance of renewable feedstocks (Himmel et al., 2007), which is a strong function of lignin content (Novaes et al., 2010). Transgenic lignocelluloses have been generated through a variety of molecular strategies (Tatsis & O'Connor, 2016), although the efficacy of microbial conversion of these biomasses to fermentation products is highly variable (Phitsuwan et al., 2013; Poovaiah et al., 2014) and significantly dependent on pretreatments. Striking a balance between low lignin content and fitness under field conditions is a key challenge for transgenic biomass.

In order to determine the suitability of transgenic feedstocks for bio-based chemicals, an assay based on microorganisms capable of Consolidated BioProcessing (CBP) (Lynd et al., 2005) (Olson et al., 2012) should be more insightful than evaluations based on simultaneous saccharification and fermentation (SSF). *Caldicellulosiruptor bescii* is an extremely thermophilic bacterium, capable of natively deconstructing and metabolizing cellulose and hemicelluloses comprising plant biomass (Blumer-Schuette et al., 2010). Furthermore, *C. bescii* has been metabolically engineered to produce non-native fermentation products, such as ethanol (Chung et al., 2014). In addition to prospects for using *C. bescii* as a platform organism for CBP, there is also the prospect of using this bacterium to screen for the efficacy of transgenic manipulations of lignocellulosic biomass to reduce recalcitrance. In other words, *C. bescii* could be used in a direct bio-assay that reports not only how recalcitrance levels impact biomass solubilization, but also how carbohydrate availability (C5 and C6 sugars) relates to fermentation products. Here, low
lignin transgenic poplar lines were screened with *C. bescii* to assess both recalcitrance and convertibility to bio-based ethanol.

**Results and Discussion**

**Monolignol Targeted Poplar Lines**

Monolignol biosynthesis can be viewed in terms of three steps: 1) creation of the key building block, 4-coumaric acid, from the amino acid phenylalanine; 2) modification of aromatic ring side groups to either alcohols or methoxy moieties at positions 3 and 5; and 3) conversion of the three-carbon branch from an organic acid to an alcohol (Figure 4-1). Various strategies to modify the monolignol pathway have been used to generate transgenic poplar lines with reduced recalcitrance (see Table 4-1). For example, to create transgenic *Populus alba x grandidentata*, RNAi was used to reduce the coumarate 3-hydroxylase (C3H) transcript to about 5% of the wild type level such that H monolignols were increased from less than 1% to 31%, thereby reducing lignin content to 10.5% (Ralph et al., 2012). In another study, fourteen lines of *P. tremula x P. alba* were generated to target the 4-coumarate-CoA ligase (4CL) gene with RNAi and field grown for two years (Voelker et al., 2010), generating lines with transcript levels as low as 20% of the control. While many of the samples had significant differences from the control, including reduced biomass productivity, reduced lignin, and increases in extractives, no significant change in enzymatic saccharification (180°C for 40 minutes followed by 72 h enzymatic treatment) was observed (Voelker et al., 2010). Another effort targeted cinnamoyl-CoA reductase (CCR) in *P. tremula x alba*; the resulting lines were tested in both the greenhouse (at 6 months) and in the field (at 10 months and at 20 months in another). In the 10-month field test, one line (FS40) had a 40% improvement (12.8% vs 9.2% dry wt) in saccharification efficiency of unpretreated material, while the 20-month field test line (FAS13) demonstrated a 140% improvement in unpretreated saccharification efficiency (18.9% vs 7.9% dry wt) (Van Acker et al., 2014). Despite the improvements, total saccharification yield for these lines was less than 20%. There have been attempts to reduce recalcitrance via other methodologies that do not involve the lignin synthesis pathway. Examples include the overexpression xylem development regulatory genes (Jung et al., 2013), down-regulation of pectin synthesis (Biswal et al., 2018), and overexpression of cell wall degrading enzymes, such as xyloglucanases (Park et al., 2004), glycosyl hydrolases (Xiao et al.,
2016) and xylanases (Shen et al., 2012). Natural variants with desirable saccharification properties have also been considered (Meng et al., 2017; Studer et al., 2011; Yoo et al., 2018).

Systems biology-based approaches may be best in dealing with the intrinsic complexity of monolignol biosynthesis as it relates to reducing recalcitrance in lignocellulosic biomass. Along these lines, (Wang et al., 2018) sought the most promising avenues for modifying P. trichocarpa lignin structure and content with an eye towards favorable wood characteristics and plant fitness. The down-regulation of 21 genes involved in monolignol biosynthesis individually and by gene-pairs and gene families were considered as a basis for a mathematical model that predicted wood traits as a function of transgenic changes. A collection of transgenic lines with a broad set of phenotypes were generated, including biomass composition, lignin monomer and linkage properties, that were assessed through enzymatic saccharification assays; these are summarized in Tables 4-2 and 4-3 along with the transcript relative to the wild type of the target gene(s). Here, these samples were subjected to solubilization and conversion by a metabolically engineered strain of C. bescii to determine those genetic and transcriptomic alterations in the transgenic poplar lines that mapped to optimal CBP results and point to further favorable outcomes for bio-based chemical production.

**C. bescii fermentation of transgenic Lines of P. trichocarpa**

Based on previous work (Wang et al., 2018), 17 transgenic samples of P. trichocarpa were selected for treatment with a metabolically engineered strain of C. bescii in which the adhE gene (bi-functional alcohol dehydrogenase) from Clostridium thermocellum was inserted to enable the generation of ethanol, in addition to its natural fermentation products, acetate, H₂ and CO₂ (Williams-Rhaesa et al., 2018). Poplar stems (bark removed) were milled and sieved to 40/80 mesh, water washed and dried, and incubated with C. bescii for 7 days at 65°C. Prior to this study, C. bescii had been examined on two lines of transgenic switchgrass with reduced lignin content, resulting in small improvements in biomass solubilization and fermentation (Zurawski et al., 2017). However, the broader sample set available here provided an opportunity to examine C. bescii efficacy as a function of recalcitrance factors, especially lignin content.

Lignin is the primary component responsible for biomass recalcitrance (Li et al., 2016). Many studies have attempted to not only reduce its content as a percentage of the biomass, but also to modify its structure by down-regulation of genes in the biosynthesis of the lignin monomers
or the lignin synthesis enzymes (Novaes et al., 2010). As is shown in Tables 4-2 and 4-3, the lines generated by (Wang et al., 2018) varied significantly in terms of biomass solubilization (15% to 79%) and fermentation products (6.8 to 29.6 mM), with some lines performing more poorly than the wild type control (i.e., 20.1% lignin and 7.3 mM). Overall, conversion of poplar carbohydrates to fermentation products (acetate, ethanol) by C. bescii directly correlated with biomass solubilization ($R^2 = 0.81$) (Figure 4-2A), and inversely with lignin content ($R^2 = 0.79$) (Figure 4-2B). However, there were some unexpected results. One outlier of interest was the a4-3 (PAL5) transgenic poplar line. With a lignin content of 14.5%, an expected improvement in solubilization (56%) compared to wild type (20.1%) was observed. However, the total fermentation products (9.1 mM) were comparable to wild type poplar (7.3 mM), even though previously reported saccharification levels were substantially above wild type (Wang et al., 2018). The reasons for this were unclear. Yet, upon examining the lignin properties previously reported (Table 4-2), this line has the highest proportion of spirodienone ($\beta$-1) interunit linkages (2.9% vs 2.3% for wild-type), while the lines targeting C3H3 that performed substantially better than the wild-type had 0.0% for lines i20-5, i69-4, and i69-13 and 0.4% for i20-10. Although more information is needed to draw a definitive conclusion here, one possibility is that the biomass released a compound that was inhibitory to C. bescii, suggesting that solubilization was mostly abiotic. Another possibility is that the carbohydrates remained bound to lignin moieties and, while solubilized, were not available in a form that C. bescii could utilize.

Saccharification results (glucose and xylose release) from (Wang et al., 2018) correlated with C. bescii conversion to fermentation products (acetate, ethanol) (Figure 4-2C & 4-2D); note that saccharification assays were performed with wood samples that had been treated with acetone to remove extractives, while the wood utilized here for the C. bescii treatment was milled without any other form of chemical, thermal, or prior enzymatic treatment.

While the overall lignin content was negatively correlated to fermentation performance, the type of lignin present can additionally affect the recalcitrance of the lignocellulosic feedstock. A higher ratio of syringyl to guaiacol monomers (S/G ratio) present in the lignin has been shown to improve the saccharification capability of P. trichocarpa and subsequent ethanol yield from fermentation of the saccharified biomass with yeast (Yoo et al., 2018). For the lines tested here with available data on S/G ratio, the higher ratios weakly correlated with increased fermentation performance ($R^2 = 0.41$) (Figure 4-3). The importance of the S/G ratio has been reported in prior
work and is summarized in Table 4-1 (Yoo et al., 2018). However, a much weaker link to recalcitrance was noted in the C. bescii assay.

Growth Productivity of Lignocellulosic Feedstock

A viable lignocellulosic feedstock must not only be more readily digestible, either by a naturally cellulolytic and hemicellulolytic organism, such as C. bescii, or by more traditional enzymatic saccharification; it must also have good growth performance and productivity. Lower lignin is generally correlated with growth defects (Novaes et al., 2010). This study found a similar correlation with those lines demonstrating the highest fermentation performance, i20-5 and i35-7, having stem volumes of 50% and 26% of the wild type, respectively. Yet, there are outliers in which the lower lignin content did not result in a growth defect. Line i20-10 (lignin content 13.3%), developed with the same construct as i20-5, targeting C3H3 but with slightly less down-regulation, had stem volume of 95% of the wild type, thus demonstrating low lignin composition and excellent fermentation performance without a penalty to biomass productivity.

To account for these factors, a fermentation-growth factor was created in which the fermentation product quantity (in mM) was multiplied by the stem volume (normalized to wild type) and the overall factor normalized to wild type set at 1.0 (Figure 4-4). Those lines generated targeting the C3H3 gene stand out as lines with the desirable properties for further improvement of biomass feedstocks. In fact, line i20-10 performs three times as well as the wild-type after accounting for both growth factors and fermentation performance. Thus, this suggests that C3H3 is a highly promising target for low recalcitrance biomass. However, there may only be a narrow transcript window in which the growth phenotype is maintained while also generating a less recalcitrant feedstock, considering that line i20-5 had a C3H3 transcript level of 13% while i20-10 was at 17%. Line i20-2 produced by (Wang et al., 2018) had a C3H3 transcript level at approximately 50% of wild type but still demonstrated lignin content, wood composition, and saccharification results in line with wild type. Thus, control of transcript levels may be required to generate lignocellulosic feedstocks with the desired properties. Most efforts summarized in Table 4-1 were performed by RNA interference using Agrobacterium, which does not allow fine down-regulation control due to random genome integration of transgene. Thus, more surgical genetic tools are needed to exert precise control of transcript level, localization, and impact on specific
cell types. CRISPR-Cas9 based genome editing may be the solution to more strategic control of monolignol biosynthesis and hence the desired reduction in biomass recalcitrance.

Conclusions

Recent development of novel genetic tools and procedures are now permitting the ability to create precise and significant phenotypic modifications in transgenic plants. A direct screen by a bacterium capable of both deconstructing the carbohydrate content and fermenting the resulting carbohydrate is an effective and insightful alternative to assessment by enzymatic saccharification and fermentation. Additionally, this bio-assay provides insight into how genetic alterations to the transgenic plant affect biomass solubilization and conversion. While the fermentation capabilities of *C. bescii* require further improvement for consideration of this organism for commercial use, the utilization of such a screen provides another informative tool for characterizing biomasses proposed as a lignocellulosic feedstock for bioprocessing.

Methods and Materials

Biomass Preparation

All wild type and transgenic greenhouse grown *Populus trichocarpa* samples were created and prepared as described in (Wang et al., 2018). The untreated stems were stripped of bark and air dried for approximately 72 h. The dried stem segments were milled utilizing a Wiley Mill and sieved to 40/80 mesh. The 40/80 mesh material was water-washed by adding 1.5 g of material to a 50 mL conical centrifuge tube and filling with deionized water. The centrifuge tube was centrifuged and the supernatant discarded. This was repeated twice more and the pelleted material was transferred to an aluminum weigh boat for drying at 50°C.

Biomass Solubilization

Following the 7-day incubation of *C. bescii* with the biomass, the sealed serum bottles were removed from the shaking incubator and allowed to cool to room temperature. The entire 50 mL contents were transferred to a 50 mL conical centrifuge tube and centrifuged, as described above. A portion of the supernatant was sterile-filtered and saved for fermentation product analysis. The remainder of the supernatant was discarded. Each serum bottle was rinsed with deionized water to remove any remaining biomass and this was added to the biomass pellet in the centrifuge tube.
The centrifuge tube was filled with water up to 45 mL, shaken to loosen the pellet, and centrifuged again to pellet the biomass. The supernatant was again removed and another wash performed. After the final wash and removal of the supernatant, the pellet was transferred to a tared aluminum weigh boat and dried at 50°C. Following drying, the gross weight was recorded and utilized to calculate biomass solubilization based upon the initial weight.

**Biomass Properties**

All biomass properties, such as those reported in Tables 4-2 and 4-3 and other measured solubilization and fermentation products, were analyzed and reported previously by (Wang et al., 2018).

**Microbial Growth on Biomass**

*C. bescii* was cultured at 50 mL in sealed serums bottles on 5 g/L DSMZ671 defined media with the biomass as the only substrate as described previously (Straub et al., 2019). Cultures were incubated at 65°C for 7 days (with shaking at 150 RPM) after which fermentation products were analyzed and biomass solubilization was measured, as described above.

**Analysis of Fermentation Products**

The sterile-filtered supernatant obtained from the culture was utilized for fermentation product analysis. Acetate was quantified utilizing high performance liquid chromatography (HPLC) with a Water Model 2489 UV/Vis detector. Ethanol was quantified via gas chromatography utilizing a Shimadzu GC-2014 (Phenomenex ZB-WAXplus column; Part No. 7HK-G013-22). Nitrogen was utilized as the carrier gas and detection via FID.

**Acknowledgements**

This work was supported by the U.S. Department of Energy BER Award DE-SC0019391, and the US Department of Agriculture (NIFA 2018-67021-27716). CT Straub acknowledges support from a US DoEd GAANN Fellowship (P200A160061).
Figure 4-1. Monolignol synthesis pathway present in angiosperms  Monolignol biosynthesis from phenylalanine with enzymes responsible for conversion of phenylalanine to 4-coumaric acid as base building block for all monomers, enzymes responsible for creation of alcohol and methoxy side groups on aromatic ring, and enzymes responsible for conversion of organic acid to alcohol on three carbon branch at 1° position on aromatic ring.
Figure 4-2. *C. bescii* fermentation production from poplar lines  (A) Fermentation production as a function of biomass solubilization after 7-day treatment with *C. bescii*; (B) Lignin composition of poplar lines; (C) Glucose and xylose release from saccharification assay for unpretreated (No Prt) and pretreated (Prt) (five minutes in water at 180°C) following 72 h enzymatic digestion.
Figure 4-3. S/G Ratio Effect on C. bescii Fermentation Products  Syringyl (S) over Guaiacyl (G) monolignol ratio measured in wood samples and its effect on fermentation products. [Line i20-5 data point is (9.9, 29.6) but not charted for scale.]
Figure 4-4. Fermentation-Growth Factor of Trangenics Compared to Wild Type for *C. bescii* Fermentation. The product of estimated stem volume and fermentation product titer was to assess growth phenotype conversion by *C. bescii*. (ND = No Data for stem volume.)
Table 4-1. Summary of Efforts to Generate Transgenic Poplar Lines

<table>
<thead>
<tr>
<th>Populus Species</th>
<th>Target Gene(s)</th>
<th>Genetics</th>
<th>Number Lines Reported</th>
<th>Transgenic Lignin %</th>
<th>S/G Ratio</th>
<th>Carbohydrate Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Populus tremuloides Michx.</em></td>
<td>4-coumarate:CoA ligase (4CL)</td>
<td>Antisense Downregulation</td>
<td>8</td>
<td>11.84 – 20.60 (21.6)</td>
<td>-</td>
<td>-</td>
<td>(Hu et al., 1999)</td>
</tr>
<tr>
<td><em>Populus alba x grandidentata</em></td>
<td>coumarate 3-hydroxylase (C3H3)</td>
<td>Antisense Downregulation</td>
<td>9</td>
<td>10.5 – 21.03 (23.8)</td>
<td>1.8 – 3.1 (1.8)</td>
<td>%Glucose = 46.0 – 51.1 (45.7) %Xylose = 17.3 – 21.1 (17.9)</td>
<td>(Coleman et al., 2008)</td>
</tr>
<tr>
<td><em>Populus tremuloides</em></td>
<td>4-coumarate:CoA ligase (4CL)</td>
<td>Antisense Downregulation</td>
<td>5</td>
<td>13.1 – 16.0 (22.2)</td>
<td>2.0 – 2.3 (2.2)</td>
<td>Cellulose:Lignin = 2.7 – 3.6 (1.9)</td>
<td>(Li et al., 2003)</td>
</tr>
<tr>
<td></td>
<td>coniferaldehyde 5-hydroxylase (CAld5H)</td>
<td>Sense Upregulation</td>
<td>5</td>
<td>19.7 – 22.4 (22.2)</td>
<td>3.0 – 5.5 (2.2)</td>
<td>Cellulose:Lignin - 1.8 – 2.2 (1.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*Both 4CL down-regulation &amp; CAld5H up-regulation</td>
<td></td>
<td>4</td>
<td>10.7 – 13.7 (22.2)</td>
<td>2.0 (2.2)</td>
<td>Cellulose:Lignin - 3.6 – 5.0 (1.9)</td>
<td></td>
</tr>
<tr>
<td><em>Populus tremula x alba</em></td>
<td>4-coumarate:CoA ligase (4CL)</td>
<td>Antisense Downregulation</td>
<td>14</td>
<td>-</td>
<td>1.4 – 2.2 (1.8)</td>
<td>-</td>
<td>(Voelker et al., 2010)</td>
</tr>
<tr>
<td><em>Populus tremula x alba</em></td>
<td>cinnamoyl-CoA reductase (CCR)</td>
<td>Antisense Downregulation</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>19.1 – 19.6 (20.7)</td>
<td>1.9 (1.9)</td>
<td>Sacccharification Yield&lt;sup&gt;c&lt;/sup&gt; % 32.0 – 33.8 (27.9)</td>
<td>(Van Acker et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>16.6 – 19.8 (21.7)</td>
<td>1.9 – 2.1 (2.0)</td>
<td>Sacccharification Yield&lt;sup&gt;c&lt;/sup&gt; % 23.6 – 33.7 (18.9)</td>
<td></td>
</tr>
<tr>
<td><em>Populus tomentosa</em></td>
<td>4-coumarate:CoA ligase (4CL)</td>
<td>Antisense Downregulation</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>22.2 – 22.7 (24.1)</td>
<td>-</td>
<td>Sacccharification Yield&lt;sup&gt;d&lt;/sup&gt; % 3.6 – 5.3 (4.3)</td>
<td>(Wang et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>caffeoyl-CoA 3-O-methyltransferase (CCoAOMT)</td>
<td>Antisense Downregulation</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>21.6 – 21.8 (24.1)</td>
<td>-</td>
<td>Sacccharification Yield&lt;sup&gt;d&lt;/sup&gt; % 16.8 – 17.2 (4.3)</td>
<td></td>
</tr>
<tr>
<td><em>Populus nigra x maximowiczii</em></td>
<td>4-coumarate:CoA ligase (4CL)</td>
<td>Antisense Downregulation</td>
<td>3&lt;sup&gt;e&lt;/sup&gt;</td>
<td>22.5 – 23.7 (23.6)</td>
<td>*S/V Ratio 0.6 – 1.9 (1.8)</td>
<td>-</td>
<td>(Xiang et al., 2017)</td>
</tr>
</tbody>
</table>

*Red denotes value reported for control / wild-type.

a Field Trial #1
b Field Trial #2
c Saccharification yield reported with 6.25 mM NaOH pre-treatment. See Van Acker et al for methods and raw data.
d Saccharification yield reported without pre-treatment. See Wang et al for methods and data.
e Xiang et al reports data for years 2 and 3 from lines grown in a mountain region and coastal region. Data reported here for year 3 trees grown in mountain region.
### Table 4-2. Physical and Chemical Properties of Poplar Samples Tested with *C. bescii* Treatment Results

| Line  | Target Genes | Transcript (% of Wild Type) | Lignin | Total Carb | Glucose | Xylose | Other Carb | Glucose (UnPrT) | Glucose (PrT) | Xylose (UnPrT) | Xylose (PrT) | Height (%WT) | Diameter (%WT) | Stem Volume (%WT) | Che Solub (% | Che Ferm (mM) |
|-------|--------------|-----------------------------|--------|-----------|---------|--------|-----------|--------------|--------------|---------------|--------------|--------------|----------------|-----------------|-----------------|----------|-------------|
| i20-5 | C3H3         | 13                          | 9.9    | 86.6      | 59.5    | 22.7   | 4.4       | 0.425        | 0.343        | 0.137         | 0.174        | 62.2         | 94.6           | 49.9            | 79.3      | 29.6        |
| i20-10| C3H3         | 17                          | 13.3   | 80.7      | 55.0    | 20.7   | 5.0       | 0.420        | 0.348        | 0.140         | 0.174        | 90.8         | 109.0          | 94.7            | 52.0      | 22.5        |
| i69-4 | C3H3         | 14                          | C4H1   | 9         | 96.5    | 66.9   | 23.9      | 5.7          | 0.341        | 0.431         | 0.082         | 0.161        | 66.7          | 101.8          | 60.8            | 64.4      | 28.1        |
|       | C3H2         | 27                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i69-13| C3H3         | 20                          | C4H1   | 14        | 91.1    | 62.3   | 23.3      | 5.5          | 0.268        | 0.399         | 0.063         | 0.157        | 71.5          | 103.6          | 67.4            | 51.1      | 23.3        |
|       | C4H2         | 36                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| a10-8 | C4H1         | 23                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | C4H2         | 36                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| a12-10a|    |                | C4H1   | 23        | 66.4    | 48.2   | 16.2      | 2.0          | 0.117        | 0.188         | 0.011         | 0.079        | 105.7         | 101.1          | 55.8            | 15.9      | 6.8         |
| i15-3a|    |                | C4H1   | 16        | 70.1    | 52.2   | 14.1      | 3.4          | 0.242        | 0.334         | 0.024         | 0.104        | ND            | ND             | ND              | 21.5      | 11.7        |
| i33-5 | CAD1         | 5                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i33-5 | CAD1         | 94                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i35-7 | CAD1         | 6                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | CAD2         | 81                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i24-1 | CCoAOMT1     | 6                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | CCoAOMT2     | 7                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i21-6a|    |                | CCoAOMT3| 24        | 68.9    | 50.1   | 15.5      | 3.3          | 0.122        | 0.295         | 0.014         | 0.104        | 94.2          | 95.7           | 99.7            | 26.6      | 9.9         |
| i30-1 | AldOMT2      | 5                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i6-9  | PAL1         | 4                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | PAL2         | 4                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i8-1  | PAL1         | 31                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | PAL2         | 62                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | PAL3         | 38                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | PAL4         | 120                         |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
|       | PAL5         | 21                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i4-3  | PAL5         | 21                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| i19-4a|    |                | HCT1   | 81        |        |        |           |              |              |               |              |              |                |                 |                 |           |
|       | HCT6         | 44                          |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |
| WT    | -             | -                           |        |           |         |        |           |              |              |               |              |              |                |                 |                 |           |

*a*Sample contained mix of three technical repeat plants in order to obtain sufficient samples from repository for testing. (See supplemental Table S2 for data on each technical repeat.)
<table>
<thead>
<tr>
<th>Line</th>
<th>Target Genes</th>
<th>Target Gene Transcript (% WT)</th>
<th>PB</th>
<th>H</th>
<th>G</th>
<th>S</th>
<th>S/G Ratio</th>
<th>Aldehydes</th>
<th>Alcohols</th>
<th>Aα</th>
<th>Bβ</th>
<th>Ca</th>
<th>SDα</th>
<th>XIγ</th>
<th>Cbe Solub (%)</th>
<th>Cbe Ferm (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i20-5</td>
<td>C3H3</td>
<td>13</td>
<td>6.3</td>
<td>44.8</td>
<td>5.1</td>
<td>50.2</td>
<td>9.93</td>
<td>0.0</td>
<td>100</td>
<td>80.7</td>
<td>6.6</td>
<td>12.7</td>
<td>0.0</td>
<td>7.0</td>
<td>79.3</td>
<td>29.6</td>
</tr>
<tr>
<td>i20-10</td>
<td>C3H3</td>
<td>17</td>
<td>5.4</td>
<td>17.4</td>
<td>17.5</td>
<td>65.1</td>
<td>3.73</td>
<td>0.0</td>
<td>100</td>
<td>86.4</td>
<td>3.4</td>
<td>9.8</td>
<td>0.4</td>
<td>6.1</td>
<td>52.0</td>
<td>22.5</td>
</tr>
<tr>
<td>i69-4</td>
<td>C3H3</td>
<td>14</td>
<td>5.4</td>
<td>32.8</td>
<td>17.8</td>
<td>49.4</td>
<td>2.77</td>
<td>0.0</td>
<td>100</td>
<td>85.8</td>
<td>5.1</td>
<td>9.0</td>
<td>0.0</td>
<td>5.9</td>
<td>64.4</td>
<td>28.1</td>
</tr>
<tr>
<td>i69-13</td>
<td>C3H3</td>
<td>20</td>
<td>4.2</td>
<td>24.9</td>
<td>18.5</td>
<td>56.6</td>
<td>3.06</td>
<td>0.0</td>
<td>100</td>
<td>87.0</td>
<td>3.7</td>
<td>9.3</td>
<td>0.0</td>
<td>6.5</td>
<td>51.1</td>
<td>23.3</td>
</tr>
<tr>
<td>a10-8</td>
<td>C4H1</td>
<td>14</td>
<td>3.7</td>
<td>1.6</td>
<td>33.0</td>
<td>65.4</td>
<td>1.98</td>
<td>0.0</td>
<td>100</td>
<td>87.2</td>
<td>3.9</td>
<td>7.1</td>
<td>1.7</td>
<td>7.0</td>
<td>15.9</td>
<td>6.8</td>
</tr>
<tr>
<td>a12-10a</td>
<td>4CL3</td>
<td>16</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>21.5</td>
<td>11.7</td>
</tr>
<tr>
<td>i15-3a</td>
<td>4CL5</td>
<td>10</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>28.5</td>
<td>11.8</td>
</tr>
<tr>
<td>i33-5</td>
<td>CAD1</td>
<td>5</td>
<td>8.3</td>
<td>0.3</td>
<td>28.9</td>
<td>70.8</td>
<td>2.45</td>
<td>29.2</td>
<td>70.8</td>
<td>95.9</td>
<td>1.2</td>
<td>1.4</td>
<td>1.5</td>
<td>4.8</td>
<td>15.0</td>
<td>8.5</td>
</tr>
<tr>
<td>i33-5</td>
<td>CAD1</td>
<td>94</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>24.2</td>
<td>9.1</td>
</tr>
<tr>
<td>i35-7</td>
<td>CAD1</td>
<td>6</td>
<td>14.8</td>
<td>2.7</td>
<td>24.2</td>
<td>73.1</td>
<td>3.02</td>
<td>30.4</td>
<td>69.6</td>
<td>95.8</td>
<td>1.5</td>
<td>1.5</td>
<td>1.2</td>
<td>4.6</td>
<td>78.2</td>
<td>27.4</td>
</tr>
<tr>
<td>i24-1</td>
<td>CCoAOMT1</td>
<td>6</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>21.5</td>
<td>8.9</td>
</tr>
<tr>
<td>i21-6a</td>
<td>CCoAOMT2</td>
<td>7</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>26.6</td>
<td>9.9</td>
</tr>
<tr>
<td>i30-1</td>
<td>AIdOMT2</td>
<td>5</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>46.7</td>
<td>18.8</td>
</tr>
<tr>
<td>i6-9</td>
<td>PAL1</td>
<td>4</td>
<td>3.8</td>
<td>0.3</td>
<td>28.8</td>
<td>70.9</td>
<td>2.46</td>
<td>8.9</td>
<td>91.1</td>
<td>89.2</td>
<td>2.0</td>
<td>6.7</td>
<td>2.1</td>
<td>6.9</td>
<td>30.8</td>
<td>12.7</td>
</tr>
<tr>
<td>i8-1</td>
<td>PAL1</td>
<td>31</td>
<td>3.0</td>
<td>0.2</td>
<td>27.8</td>
<td>72.0</td>
<td>2.59</td>
<td>8.4</td>
<td>91.6</td>
<td>88.3</td>
<td>2.1</td>
<td>7.1</td>
<td>2.5</td>
<td>6.1</td>
<td>28.5</td>
<td>11.3</td>
</tr>
<tr>
<td>a4-3</td>
<td>PAL5</td>
<td>21</td>
<td>2.9</td>
<td>1.0</td>
<td>24.2</td>
<td>74.8</td>
<td>3.08</td>
<td>0.0</td>
<td>100</td>
<td>90.0</td>
<td>1.6</td>
<td>5.6</td>
<td>2.9</td>
<td>5.4</td>
<td>55.7</td>
<td>9.1</td>
</tr>
<tr>
<td>i19-4a</td>
<td>HCT1</td>
<td>81</td>
<td>5.4</td>
<td>21.6</td>
<td>22.5</td>
<td>56.0</td>
<td>2.49</td>
<td>0.0</td>
<td>100</td>
<td>86.3</td>
<td>4.5</td>
<td>8.8</td>
<td>0.4</td>
<td>5.5</td>
<td>41.2</td>
<td>15.8</td>
</tr>
<tr>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>3.5</td>
<td>0.3</td>
<td>31.9</td>
<td>67.8</td>
<td>2.16</td>
<td>4.0</td>
<td>96.0</td>
<td>86.5</td>
<td>3.7</td>
<td>7.5</td>
<td>2.3</td>
<td>6.3</td>
<td>20.1</td>
<td>7.3</td>
</tr>
</tbody>
</table>

PB: p-hydroxybenzoic acid; H: H-subunits; G: G-subunits; S: S-subunits
Aα: β-aryl ether (β-O-4); Bβ: phenylcoumaran (β-5); Ca: resinol (β-β); SDα: spirodienone (β-1); XIγ: end-groups.
*Sample contained mix of three technical repeat plants in order to obtain sufficient samples from repository for testing.
<table>
<thead>
<tr>
<th>Target Gene(s)</th>
<th>Samples from Wang et al</th>
</tr>
</thead>
<tbody>
<tr>
<td>i20-5</td>
<td>C3H3</td>
</tr>
<tr>
<td>i20-10</td>
<td>C3H3</td>
</tr>
<tr>
<td>i69-4</td>
<td>C4H1</td>
</tr>
<tr>
<td>i69-13</td>
<td>C4H2</td>
</tr>
<tr>
<td>i10-8</td>
<td>C4H1</td>
</tr>
<tr>
<td>a12-10</td>
<td>4CL3</td>
</tr>
<tr>
<td>i15-3'</td>
<td>4CL3, 4CL5</td>
</tr>
<tr>
<td>i33-5</td>
<td>CAD1</td>
</tr>
<tr>
<td>i35-7</td>
<td>CAD1, CAD2</td>
</tr>
<tr>
<td>i24-1</td>
<td>CCoAOMT1, CCoAOMT2</td>
</tr>
<tr>
<td>i21-6</td>
<td>CCoAOMT3</td>
</tr>
<tr>
<td>i30-1</td>
<td>AldOMT2</td>
</tr>
<tr>
<td>i6-9</td>
<td>PAL1, PAL2</td>
</tr>
<tr>
<td>i8-1</td>
<td>PAL1, PAL2, PAL3, PAL4, PAL5</td>
</tr>
<tr>
<td>a4-3</td>
<td>PAL5</td>
</tr>
<tr>
<td>i19-4</td>
<td>HCT1, HCT6</td>
</tr>
</tbody>
</table>

Table 4-S1. Samples from Wang et al Utilized in this Study
Table 4-S2. Physical and Chemical Properties of Pooled Poplar Samples Tested with *C. bescii* Treatment Results

<table>
<thead>
<tr>
<th>Line</th>
<th>Target Genes</th>
<th>Transcript (% of Wild Type)</th>
<th>Lignin</th>
<th>Total Carb</th>
<th>Glucose</th>
<th>Xylose</th>
<th>Other Carb</th>
<th>Glucose (UnPrt)</th>
<th>Glucose (Prt)</th>
<th>Xylose (UnPrt)</th>
<th>Xylose (Prt)</th>
<th>Height (% WT)</th>
<th>Diameter (% WT)</th>
<th>Stem Volume (% WT)</th>
<th>Che Solub (%)</th>
<th>Che Ferm (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a12-10-1</td>
<td>4CL3</td>
<td>23</td>
<td>17.9</td>
<td>70.0</td>
<td>51.2</td>
<td>15.3</td>
<td>3.5</td>
<td>0.144</td>
<td>0.323</td>
<td>0.023</td>
<td>0.108</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>21.5</td>
<td>11.7</td>
</tr>
<tr>
<td>a12-10-2</td>
<td>4CL3</td>
<td>13</td>
<td>17.0</td>
<td>69.0</td>
<td>52.2</td>
<td>13.3</td>
<td>3.5</td>
<td>0.299</td>
<td>0.329</td>
<td>0.021</td>
<td>0.101</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>28.5</td>
<td>11.8</td>
</tr>
<tr>
<td>a12-10-3</td>
<td>4CL3</td>
<td>11</td>
<td>17.4</td>
<td>70.1</td>
<td>53.1</td>
<td>13.5</td>
<td>3.5</td>
<td>0.283</td>
<td>0.351</td>
<td>0.027</td>
<td>0.103</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>26.6</td>
<td>9.9</td>
</tr>
<tr>
<td>a12-10</td>
<td>4CL3</td>
<td>16</td>
<td>17.4</td>
<td>69.7</td>
<td>52.2</td>
<td>14.1</td>
<td>3.4</td>
<td>0.242</td>
<td>0.334</td>
<td>0.024</td>
<td>0.104</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i15-3-1</td>
<td>4CL3</td>
<td>3</td>
<td>19.2</td>
<td>70.8</td>
<td>52.6</td>
<td>14.9</td>
<td>3.3</td>
<td>0.201</td>
<td>0.312</td>
<td>0.014</td>
<td>0.095</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i15-3-2</td>
<td>4CL3</td>
<td>1</td>
<td>19.0</td>
<td>72.7</td>
<td>54.2</td>
<td>14.8</td>
<td>3.7</td>
<td>0.214</td>
<td>0.295</td>
<td>0.017</td>
<td>0.102</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i15-3-3</td>
<td>4CL3</td>
<td>5</td>
<td>19.1</td>
<td>71.5</td>
<td>54.5</td>
<td>13.8</td>
<td>3.2</td>
<td>0.261</td>
<td>0.309</td>
<td>0.013</td>
<td>0.088</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i15-3</td>
<td>4CL3</td>
<td>3</td>
<td>19.1</td>
<td>71.6</td>
<td>53.8</td>
<td>14.5</td>
<td>3.3</td>
<td>0.225</td>
<td>0.305</td>
<td>0.015</td>
<td>0.095</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i15-3</td>
<td>4CL5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i21-6-1</td>
<td>CCoAOMT3</td>
<td>13</td>
<td>18.8</td>
<td>68.8</td>
<td>49.7</td>
<td>15.6</td>
<td>3.5</td>
<td>0.116</td>
<td>0.267</td>
<td>0.012</td>
<td>0.100</td>
<td>94.4</td>
<td>93.5</td>
<td>95.8</td>
<td>26.6</td>
<td>9.9</td>
</tr>
<tr>
<td>i21-6-2</td>
<td>CCoAOMT3</td>
<td>31</td>
<td>18.9</td>
<td>68.1</td>
<td>49.2</td>
<td>15.3</td>
<td>3.6</td>
<td>0.112</td>
<td>0.335</td>
<td>0.016</td>
<td>0.110</td>
<td>104.7</td>
<td>99.1</td>
<td>117.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i21-6-3</td>
<td>CCoAOMT3</td>
<td>29</td>
<td>18.5</td>
<td>69.8</td>
<td>51.3</td>
<td>15.5</td>
<td>3.0</td>
<td>0.138</td>
<td>0.284</td>
<td>0.013</td>
<td>0.102</td>
<td>83.5</td>
<td>94.4</td>
<td>85.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i21-6</td>
<td>CCoAOMT3</td>
<td>24</td>
<td>18.7</td>
<td>68.9</td>
<td>50.1</td>
<td>15.5</td>
<td>3.3</td>
<td>0.122</td>
<td>0.295</td>
<td>0.014</td>
<td>0.104</td>
<td>94.2</td>
<td>95.7</td>
<td>99.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i19-4-1</td>
<td>HCT1</td>
<td>34</td>
<td>17.2</td>
<td>79.6</td>
<td>55.9</td>
<td>19.2</td>
<td>4.5</td>
<td>0.182</td>
<td>0.377</td>
<td>0.051</td>
<td>0.148</td>
<td>57.1</td>
<td>75.0</td>
<td>37.1</td>
<td>41.2</td>
<td>15.8</td>
</tr>
<tr>
<td>i19-4-2</td>
<td>HCT1</td>
<td>82</td>
<td>18.0</td>
<td>75.1</td>
<td>51.8</td>
<td>18.7</td>
<td>4.6</td>
<td>0.177</td>
<td>0.302</td>
<td>0.041</td>
<td>0.099</td>
<td>52.0</td>
<td>65.1</td>
<td>26.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i19-4-3</td>
<td>HCT1</td>
<td>127</td>
<td>17.6</td>
<td>75.5</td>
<td>52.8</td>
<td>17.9</td>
<td>4.8</td>
<td>0.170</td>
<td>0.290</td>
<td>0.040</td>
<td>0.103</td>
<td>34.6</td>
<td>42.0</td>
<td>8.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i19-4</td>
<td>HCT1</td>
<td>81</td>
<td>17.6</td>
<td>76.7</td>
<td>53.5</td>
<td>18.6</td>
<td>4.6</td>
<td>0.176</td>
<td>0.323</td>
<td>0.044</td>
<td>0.117</td>
<td>47.9</td>
<td>60.7</td>
<td>23.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i19-4</td>
<td>HCT6</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


CHAPTER 5: Quantitative Fermentation of Unpretreated Transgenic Poplar by 
*Caldicellulosiruptor bescii*

Christopher T. Straub¹, Piyum A. Khatibi¹.,& Jack P. Wang², Jonathan M. Conway¹, 
Amanda M. Williams-Rhaesa³, Ilona M. Peszlen⁴, 
Vincent L. Chiang², Michael W.W. Adams³, and Robert M. Kelly¹,*

¹Department of Chemical and Biomolecular Engineering, 
North Carolina State University, Raleigh, NC 27695

²Department of Forestry and Environmental Resources, 
North Carolina State University, Raleigh, NC 27695

³Department of Biochemistry and Molecular Biology, 
University of Georgia, Athens, GA 30602

⁴Department of Forest Biomaterials, 
North Carolina State University, Raleigh, NC 27695

&Present address: Novozymes Biologicals, Inc., Durham, NC 27709

Accepted (In Press): *Nature Communications* (Jul 2019)
Abstract

Microbial fermentation of lignocellulosic biomass to produce industrial chemicals is exacerbated by the recalcitrant network of lignin, cellulose and hemicelluloses comprising the plant secondary cell wall. However, transgenic poplar (*Populus trichocarpa*) lines could be solubilized without any pretreatment by the extreme thermophile *Caldicellulosiruptor bescii* that had been metabolically engineered to shift its fermentation products away from inhibitory organic acids to ethanol. Carbohydrate solubilization and conversion of unpretreated milled biomass was nearly 90% for two transgenic lines, compared to only 25% for wild-type poplar. Unexpectedly, approximately 70% of transgenic poplar solubilization and conversion were achieved even with unpretreated, intact poplar stems. The nearly quantitative microbial conversion of the carbohydrate content of unpretreated transgenic lignocellulosic biomass bodes well for full utilization of renewable biomass feedstocks.
Introduction

Industrial biotechnology aspires to produce industrial chemicals from renewable feedstocks to address the finite nature of fossil fuels and associated global warming concerns. Metabolically engineered model microorganisms (typically *Escherichia coli* and yeast) can produce a range of industrial chemicals from simple sugars. However, they are unable to directly metabolize the carbohydrate content of lignocellulosic biomass, such that physical, chemical and enzymatic pretreatment steps with their associated costs are a bioprocessing necessity. The bacterial genus *Caldicellulosiruptor* is globally distributed in terrestrial hot springs, with some species growing at temperatures up to 90°C by deconstructing and fermenting the two major complex carbohydrates found in plant biomass: cellulose and hemicelluloses. One species, *Caldicellulosiruptor bescii*, can completely solubilize and metabolize crystalline cellulose and hemicellulose in their purified forms. However, lignin, a complex heterogeneous aromatic polymer and the other major component of the secondary plant cell wall, introduces a significant barrier to plant biomass solubilization by this bacterium. Feedstocks with modified lignin content could be the solution to the recalcitrance of lignocellulosic biomass to microbial attack. However, such approaches must align with the physiology of the deconstructing microorganism. Even with switchgrasses transgenically-modified to achieve lower lignin levels, *C. bescii* solubilized at best 36% of the carbohydrate content from unpretreated samples, compared to 24% for the parent wild-type; the result for the transgenic was no better than for the solubilization of a natural variant switchgrass by this bacterium. Thus, lignin reduction, in and of itself, may not be sufficient to achieve significant microbial conversion of plant biomass. Herein, we sought to identify biomass feedstocks with attributes that aligned with the lignocellulosic features of *C. bescii* so that near complete carbohydrate solubilization and conversion without pretreatment could be achieved.

Results

Two low lignin, but otherwise unpretreated, transgenic poplar lines were compared with the corresponding parent wild-type (22% lignin, syringyl (S)/guaiacyl (G) = 2.1, 4% aldehyde content in lignin) to assess the extent of direct microbial conversion of the carbohydrate content to fermentation products by *C. bescii*. Line #54 was generated by down-regulating the coumarate 3-hydroxylase 3 (*PtrC3H3*) gene (to 12.5% of the wildtype transcript level), resulting in transgenic wood with a lignin content of 10% and a lignin S/G ratio of 9.9. Line #80 targeted the down-
regulation of two cinnamyl alcohol dehydrogenases (PtrCAD1, PtrCAD2, to 5.9% and 80.9% of the wildtype transcript level, respectively) that reduce cinnamaldehydes to their corresponding alcohols for lignin biosynthesis. This resulted in lowering the lignin content to 14% with a lignin-aldehyde content of 30%. The ultimate goal here was to determine the extent to which direct microbial conversion of the modified lignin poplar lines could be achieved without any physical, chemical and/or enzymatic pretreatment.

Based on a previous effort to metabolically-engineer C. bescii 9, but using a genetically stable strain of this bacterium 10,11, the lactate dehydrogenase gene (ldh) was deleted to eliminate lactate production, and the alcohol dehydrogenase (adhE) gene from Clostridium thermocellum was inserted to shift fermentation carbon flux to ethanol rather than acetate, reducing product inhibition (C. bescii strain Δldh::adhE) 12. The extent of poplar solubilization and conversion by engineered C. bescii was quantified in a series of batch fermentations. Cultures were grown on 5 g/L poplar (milled to 0.18 – 0.42 mm) as the sole carbon source. Total biomass solubilization of the wild-type poplar with engineered C. bescii after 7 days at 65°C (the optimal temperature of AdhE) was 20%, while lines #54 and #80 solubilized 79% and 78%, respectively; abiotic controls in these cases were 5%, 8%, and 11% for wild-type, #54 and #80, respectively (Figure 5-1A). Note that purified crystalline cellulose (Avicel) at the same loading was 87% solubilized compared to 2% for the abiotic control. Carbohydrate solubilization for the transgenic lines was nearly complete at 87% and 90% for #54 and #80, respectively, compared to 25% for the wild-type; note that carbohydrate release from the transgenic lines was nearly the same as for Avicel (90%), demonstrating the impact of modifying lignin structure and composition (Fig. 1B). The engineered C. bescii strain converted most of the carbohydrate content from the transgenic lines into fermentation products: 18.3 mM/11.3 mM and 16.5 mM/11.0 mM ethanol/acetate for lines #54 and line #80, respectively, compared to 2.4 mM/4.9 mM and 17.0 mM/12.4 mM for wild-type poplar and Avicel, respectively (Figure 5-1C).

For lignocellulosic biomass bioprocessing, pretreatments and subsequent economic viability concerns typically revolve around heat, chemicals, and enzymes. However, the energy and cost required to reduce the size of the lignocellulose feedstock can likewise have a dramatic impact on the economic considerations. A previous analysis of energy requirements for woody biomass feedstocks demonstrated that chipping mature trees to standard commercial chip size (10-50 mm in two dimensions, 5-15 mm in one dimension) consumes only 0.18 MJ/kg wood,
approximately 1% of the total theoretical thermal energy present in the wood. In contrast, chipping followed by disk milling to reduce to fibrous particles (<1 mm) consumes a total of 2.16 MJ/kg wood, or more than 12 times as much energy as chipping alone \(^{13}\). Therefore, any commercial bioprocess envisioned, based on extensive size reduction beyond chipping, would face significant economic hurdles.

Hence, to determine the impact of biomass particle size for the transgenic poplar lines examined here, fermentations utilizing whole segments of 6-month old trees, approximately 5 mm diameter, were examined. Single segments of the unwashed de-barked stems were cut to obtain a weight of 0.25 – 0.30 g and subjected to the same conditions as the milled poplar, i.e. 50 mL fermentations with 5 – 6 g/L loading. At equivalent mass loadings, the stem segments were solubilized at 12% for the wild-type poplar, but at 50% for line #54 and 52% for line #80; abiotic controls in these cases were 5%, 5%, and 11% for wild-type, #54 and #80, respectively. Thus, the intact transgenic poplar stems, with particle dimensions 100-1,000 times larger than the milled material, were solubilized to about two-thirds the extent of the comparable milled material. The stems also yielded fermentation products at 70% of that observed with the milled poplar. Surprisingly, measurements of the dried stem segments (diameters and length) following \(C.\ bescii\) fermentation revealed that the reduction in length and diameter was minimal. Following biotic treatment, stems were noticeably softer than the abiotic controls, although they retained their shape with no visually obvious degradation. Upon drying, extensive degradation was observed at the ends of the transgenic stems, while the wild-type stems and abiotic controls showed no such degradation (Figure 5-2). Bulk density of the wild-type stem decreased by 16% compared to over 60% for the transgenic lines, consistent with solubilization and conversion data. \(C.\ bescii\) likely accessed carbohydrate content via the cross-sectional vasculature of the plant rather than through radial penetration of the wood. Given that these plant cells range from 5-15 μm in diameter and \(C.\ bescii\) is a rod-shaped bacterium on order of 1-3 μm in length and 0.5 μm in diameter, this is a plausible hypothesis.

We also investigated the physical changes to the plant cell wall in stem material, brought on by modifying lignin in transgenic poplar, after incubation at 65°C (abiotic only), and after exposure of the poplar stem at 65°C to engineered \(C.\ bescii\). Figure 5-2A shows the cross section of the stem, demonstrating the dramatic difference following treatment with \(C.\ bescii\) at 65°C. Figure 5-2B and Figure 5-2C show progressively higher resolution images that reinforce the
degradation of the plant vascular structure as carbohydrate is solubilized and converted by *C. bescii*. Note in particular the thinning of the xylem fiber cell wall as a consequence of lignin modification and microbial attack. **Figure 5-2D** further illustrates the degradation of the xylem fiber plant cell wall from the longitudinal perspective of the cell wall.

**Discussion**

The results here with modified poplar lines can be put into perspective given other recent efforts to solubilize modified lignocellulosic feedstocks. *C. bescii* carbohydrate solubilization of unpretreated but milled transgenic switchgrass lines (14.7 g/L biomass loading) at 75°C was in all cases less than 30% and no statistically significant differences were observed when comparing parent and transgenic lines 14-16. *Clostridium thermocellum*, a moderately thermophilic anaerobe, was also tested on these feedstocks and performed best on the COMT Knock-Down (down-regulation of the caffeic acid 3-O-methyltransferase (COMT3)) switchgrass line with increased carbohydrate solubilization at 60°C from 45% to 61%. Unpretreated samples of a natural poplar variant with a mutation in a lignin synthesis pathway were also examined with *C. thermocellum* carbohydrate solubilization, but the increase from 20% to 31% for *C. thermocellum* solubilization is far less than the 25% to 90% carbohydrate solubilization observed here for *C. bescii* acting on lines #54 and #80. Although no carbohydrate conversion data were provided for the *C. thermocellum* poplar and switchgrass solubilization experiments 15,16, the inability of this bacterium to natively metabolize pentose sugars, which account for 20-50% of the carbohydrate content of lignocellulosic biomass, is clearly a barrier to complete carbohydrate utilization.

A key aspect of the work reported here is that high levels of both solubilization and conversion of plant carbohydrates were achieved without any biomass pretreatment. Not only could *C. bescii* solubilize cellulose and hemicellulose from the transgenic poplar, it also metabolized the resulting pentose and hexose sugars. While the results obtained here are for plant biomass loadings that are considerably below industrial requirements, *C. bescii* has previously been cultured at high switchgrass loadings (200 g/L) 17 and efforts are underway to examine high loadings at bioreactor scale for related transgenic poplar lines.
Methods and Materials

Bacterial Strains and Growth Conditions

A previously published recombinant *C. bescii* (MACB 1058) strain was utilized for the solubilization experiments. Briefly, a genetically stable parent strain (MACB 1034 - *Aldh*) was utilized with integration of the bifunctional alcohol dehydrogenase (AdhE) from *Clostridium thermocellum* DSM 1313 at the locus between genes _Athe_0949 and _Athe_0950.

Strains were grown in a shaking incubator oven at 65°C and 150 RPM in a modified version of DSMZ 671 defined medium under anaerobic conditions, with poplar or Avicel as the sole carbon source.

A freezer stock of the *C. bescii* strain was recovered in media containing 5 g/L cellobiose and then passaged (2% inoculum) onto media containing 0.25 g/L cellobiose and 5 g/L poplar substrate. After 48 hours, this was passaged (2% inoculum) onto a medium containing only poplar substrate as the carbon source.

Poplar Substrate Preparation

Various lines of genetically engineered *Populus trichocarpa* were generated as previously described. Of the 221 genetically engineered lines analyzed in the aforementioned work, two lines (referred to as #54 & #80) were suggested for further analysis based upon previously determined saccharification data and lignin content. For solubilization experiments, the poplar stems from 6-month-old greenhouse grown trees were stripped of bark and internodes 1 to 5 were removed. The stems were lightly scraped using a razor blade to remove the outer layers of differentiating xylem cells, and then allowed to air dry for 48 hours. Small samples of the stems were taken and prepared for further wood chemistry and properties examination with data and analysis reported elsewhere. The remainder of the stem samples were cut or ground to size, yet otherwise untreated, for microbial solubilization experiments.

The stem segments were obtained by cutting a section of the stem to a weight of 0.25 g to 0.30 g. The ends of the stems were lightly sanded with 400 grit sandpaper. The stems were dried for 24 hours in a 50°C oven and weighed before utilizing as a substrate in solubilization experiments.

The dry whole stems were milled on a Wiley Mill with a 40-mesh screen and sieved. The 40/80 mesh fraction was collected and washed with 50°C water by adding 1.5 g of biomass to a
50 mL conical centrifuge tube and filling to approximately 50 mL with deionized water. The conical centrifuge tube was shaken and then centrifuged for 10 minutes at 4696 x g in a Sorvall X1R swinging bucket centrifuge. The supernatant fraction was carefully removed with a pipette aid. 50°C deionized water was again added to bring the total volume to 50 mL. The conical centrifuge tube was again centrifuged and supernatant removed. This wash was repeated a second time. After the final removal of supernatant, the contents were transferred to an aluminum weigh boat with deionized water and allowed to dry at 50°C for 24 h. The washed and dried material was then utilized in solubilization experiments.

**Poplar Substrate Solubilization Experiments.**

Serum bottles (125 mL) containing 0.25 g (5 g/L) of poplar substrate (milled, sieved, washed and dried material or dried stem segments) were filled with 50 mL of DSMZ 671 defined media. Closed bottles were then brought to anaerobic conditions via vacuum purging cycles with 20/80 CO₂/N₂. Bottles were inoculated with 2% inoculum and placed in a shaking incubator oven (150 RPM) at 65°C for 7 days.

Following the microbial fermentation treatment, the stem segment was removed and transferred to a tared weigh boat. The remaining bottle contents, including any residual solids, were transferred to a 50 mL conical centrifuge tube. For milled poplar experiments, all contents were transferred to a 50 mL conical centrifuge tube. The conical centrifuge tubes were centrifuged at 4,696 x g in a Sorvall Legend X1R centrifuge for 10 min. Supernatant was removed with a pipette aid and sterile filtered for further analysis. Pelleted solids were washed with 50°C DI water another two times as described above. The biomass pellet was transferred to a tared aluminum weigh boat with deionized water. The material was dried at 50°C for 24 h.

The post-fermentation dry mass includes dry residual solids, dry bacterial weight, and dried stem segment (as applicable). Dry mass (including all solid substrate and dry cell weight) was utilized to calculate mass solubilization and utilized in calculation of carbohydrate balance.

**Calculation of Biomass Bulk Density**

Poplar stem mass was measured before solubilization experiments. An average of six diameter measurements were obtained and along with the stem length, these measurements were utilized to calculate a cylindrical volume as an estimate for bulk volume. Following drying of the
stem after biotic treatment, the same procedure was followed to establish a ‘before and after’ estimate of bulk density (mass over calculated cylindrical volume).

**Analysis of Fermentation Products**

Acetate concentrations in the fermentation medium were analyzed by HPLC (5 mM sulfuric acid mobile phase). A Rezex-ROA column (300 mm by 7.8 mm; Phenomenex) was utilized for separation with detection by a Waters Model 2414 Infrared detector and a Waters Model 2489 UV/Vis detector. Ethanol was analyzed via gas chromatography (Shimadzu GC-2014) using an FID detector and Phenomenex ZB-WAXplus column (Part No. 7HK-G013-22).

**Analysis of Poplar Carbohydrate Content**

Samples, both pre- and post-fermentation, along with abiotic controls, were analyzed for carbohydrate content via a modified version of NREL procedure - Determination of Structural Carbohydrates and Lignin in Biomass (NREL/TP-510-42618) as described elsewhere.

**Electron Microscopy**

From the treated stem segments, 1 mm thick disks with duplicates were hand cut using razor blades at wet condition. In addition, longitudinal cuts of approximately 3 mm were also prepared. The disks and the longitudinal sections were attached to stubs and air-dried first, before vacuum dried for examination with FEI Verios 460L field-emission scanning electron microscope (FESEM).
Figure 5-1. Solubilization and conversion of wild-type and low lignin transgenic poplar lines by ethanol-producing *C. bescii* over 7 days at 65°C. Greenhouse-grown poplar trees were harvested at 6 months, debarked, air dried, and prepared without further processing as approximately 50 cm long, 0.5 cm diameter stem segments. A portion of the stem material was milled and sieved to 40/80 mesh size, water washed, dried at 50°C and used at 5 g/L loading. Further stem segments were cut to a weight of 0.25 - 0.30 g and used directly in experiments. Solid bars represent milled biomass, hatched bars represent poplar stem segments. (A) Total biomass solubilization. (B) Total carbohydrate solubilization. (C) Carbohydrate conversion to *C. bescii* fermentation products. Error bars represent one standard error of three biological replicates.
Figure 5-2. Images of wild type and lignin-modified transgenic poplar stems (#54 and #80) with and without solubilization by *C. bescii*. Six-month-old stem segments were incubated for 7 days at 25°C (25°C control), at 65°C without *C. bescii* (abiotic control), or at 65°C with *C. bescii*. (A) End view of stems before and after treatment indicating the level of degradation. (B) Low magnification and (C) High magnification SEM images of stem cross-sections showing splits, fragmentations and the effects of lignin modification and *C. bescii* solubilization on the xylem cell wall fibers. (D) Comparison of longitudinal xylem fiber cell surfaces for different treatments.
References


9. Chung, D., Cha, M., Guss, A. M. & Westpheling, J. Direct conversion of plant biomass to


CHAPTER 6: A synthetic enzymatic pathway for extremely thermophilic acetone production based on the unexpectedly thermostable acetoacetate decarboxylase from *Clostridium acetobutylicum*

Benjamin M. Zeldes¹,#, Christopher T. Straub¹,#, Jonathan K. Otten¹,
Michael W.W. Adams², and Robert M. Kelly¹,*

¹Department of Chemical and Biomolecular Engineering
North Carolina State University, Raleigh, NC 27695

²Department of Biochemistry and Molecular Biology
University of Georgia, Athens, GA 30602

#B.M. Zeldes and C.T. Straub contributed equally to this work

Abstract

One potential advantage of an extremely thermophilic metabolic engineering host ($T_{\text{opt}} \geq 70^\circ\text{C}$) is facilitated recovery of volatile chemicals from the vapor phase of an active fermenting culture. This process would reduce purification costs and concomitantly alleviate toxicity to the cells by continuously removing solvent fermentation products such as acetone or ethanol, a process we are calling “bio-reactive distillation”. While extremely thermophilic heterologous metabolic pathways can be inspired by existing mesophilic versions, they require thermostable homologs of the constituent enzymes if they are to be utilized in extremely thermophilic bacteria or archaea. Production of acetone from acetyl-CoA and acetate in the mesophilic bacterium *Clostridium acetobutylicum* utilizes three enzymes: thiolase (Thl), acetoacetyl-CoA:acetate CoA transferase (CtfAB), and acetoacetate decarboxylase (Adc). Previously reported biocatalytic pathways for acetone production were demonstrated only as high as 55°C. Here, we demonstrate a synthetic enzymatic pathway for acetone production that functions up to at least 70°C *in vitro*, made possible by the unusual thermostability of Adc from the mesophile *C. acetobutylicum*, and heteromultimeric acetoacetyl-CoA:acetate CoA-transferase (CtfAB) complexes from *Thermosipho melanesiensis* and *Caldanaerobacter subterraneus*, composed of a highly thermostable $\alpha$-subunit and a thermally labile $\beta$-subunit. The three enzymes produce acetone *in vitro* at temperatures of at least 70°C, paving the way for bio-reactive distillation of acetone using a metabolically engineered extreme thermophile as production host.
Introduction

Acetone is widely used industrially as a solvent and polymer precursor. An acetone shortage in Great Britain during World War I led to one of the first instances of large-scale industrial fermentation, leveraging the ability of *Clostridium acetobutylicum* to convert starches and sugars into acetone, butanol, and ethanol (Sauer, 2016). Routes to produce acetone from petroleum eventually eclipsed fermentative processes, and recent research on *C. acetobutylicum* has predominantly focused on production of bio-butanol as a drop-in replacement for gasoline, in which context acetone is often seen as an undesired byproduct to be minimized (Jang et al., 2012). However, acetone remains a valuable product in its own right (Luo et al., 2016). As a commodity chemical, acetone prices in recent years have exceeded those for ethanol, although it remains less valuable than n-butanol. Acetone is a feedstock in the production of bisphenol A and methyl methacrylate-based polymers, used as a fuel additive, and of course as a solvent (direct solvent use represents ~30% of total demand) (Wu, Wang, Liu, & Huo, 2007). Acetone is also considerably less toxic to cells than n-butanol, which can cause severe toxicity even at low concentrations due to its fluidizing effects on the cell membrane (Peabody & Kao, 2016). Additionally, recombinant expression of enzymes for acetone production in native acetate producers has improved growth (Bermejo, Welker, & Papoutsakis, 1998; Shaw et al., 2015), since acetone appears to be less toxic than acetate, especially at low pH.

The high volatility of acetone (normal boiling point 56°C) makes it a potentially strategic metabolic engineering product for an extremely thermophilic host (T_\text{opt} \geq 70°C), where continuous recovery of product from the bioreactor, termed “bio-reactive distillation”, could be possible even at atmospheric pressure and relatively modest titers (**Figure 6-1**). In contrast, implementation of such a concept for mesophilic hosts would require near full vacuum conditions, which then requires energy intensive refrigeration at the distillation condenser (Cysewski & Wilke, 1977). Unfortunately, no microorganism growing at such elevated temperatures is known to produce acetone. In fact, production of solvents is rare among the fermentative extreme thermophiles, which tend to produce organic acids and hydrogen gas instead. The highest reported temperatures for native production of acetone, ethanol, and butanol are 43°C (Weimer, 1984), 72°C (Svetlitchnyi et al., 2013), and 58°C (Freier-Schroder, Wiegel, & Gottschalk, 1989), respectively, while metabolically engineered hosts expressing recombinant enzymes have allowed production
as high as 55°C (Shaw et al., 2015), 78°C (Basen et al., 2014), and 60°C (Keller et al., 2015), respectively, as summarized in Table 6-1.

The absence of known native acetone producing extremely thermophilic microorganisms suggests that creation of an acetone production pathway in an extremely thermophilic host requires establishing a synthetic production pathway with recruitment of enzymes from multiple thermophilic organisms. Such efforts have been reported previously, an example of which is n-butanol in *Pyrococcus furiosus* (Keller et al., 2015). The mesophilic three-enzyme pathway in *C. acetobutylicum* serves as a template for biological acetone production.

\[
\text{Thl:} \quad 2 \text{Acetyl-CoA} \rightarrow \text{Acetoacetyl-CoA + CoA} \\
\text{CtfAB:} \quad \text{Acetoacetyl-CoA + Acetate} \rightarrow \text{Acetoacetate + Acetyl-CoA} \\
\text{Adc:} \quad \text{Acetoacetate} \rightarrow \text{Acetone + CO}_2 \\
\text{Net Rxn:} \quad \text{Acetyl-CoA + Acetate} \rightarrow \text{Acetone + CO}_2 + \text{CoA}
\]

First, a thiolase (Thl) extends carbon chains by condensing two acetyl-CoA molecules into an acetoacetyl-CoA. Acetoacetyl-CoA:acetate CoA-transferase (Ctf) then transfers the CoA moiety from acetoacetyl-CoA to acetate, generating acetoacetate while transforming acetate to acetyl-CoA. Finally, acetoacetate decarboxylase (Adc) decarboxylates acetoacetate to acetone and CO2. The reaction catalyzed by Adc has also been found to occur non-enzymatically, but a mutant strain lacking the gene lags behind wild-type in acetone production (although final titers were found to be comparable) (Han, Gopalan, & Ezeji, 2011), suggesting Adc activity is highly advantageous in acetone production. Therefore, it was necessary to identify thermostable, thermoactive versions of these three enzymes.

The most straightforward method to identify thermostable enzymes is a search for homologs (based upon amino acid sequence) within genomes of thermophiles with optimum growth temperatures at or near the desired working temperature (in this case, at least 70°C). This was the approach previously adopted (Shaw et al., 2015), utilizing Thl from *Thermoanaerobacterium thermosaccharolyticum* (T_{opt} 60°C), Ctf from *Thermosipho melanesiensis* (T_{opt} 70°C), and Adc from *Bacillus amyloliquefaciens* (T_{opt} 50°C). While this collection of enzymes successfully produced acetone as a minor product at 55°C in T.
saccharolyticum, the Thl and Adc were from moderately thermophilic organisms, and thus unlikely to function in extreme thermophiles.

Thermophilic homologs can be found for many mesophilic enzymes, but in the case of highly specialized metabolisms, or reactions that are thermodynamically less favorable at higher temperatures, no thermophilic candidates may be available. However, there are cases where enzymes from mesophilic organisms exhibit unusual thermostability, such as the industrially relevant α-amylase from *Bacillus licheniformis* (Saito, 1973). Another example can be found in the acetone pathway; acetoacetate decarboxylase from *C. acetobutylicum* was reported to be active at 70°C when first characterized in partially purified forms (Davies, 1943). Subsequent work determined the enzyme retained 50% activity after 30 minutes at 80°C, and activity actually increased following an hour-long incubation at 70°C (Autor & Fridovich, 1970). *C. acetobutylicum* grows optimally at 35°C, thus it is unclear why the mesophile possesses an enzyme stable above 70°C, although it is worth noting that *C. acetobutylicum* spores are activated by brief heat shock at temperatures as high as 80°C (Al-Hinai et al, 2014).

The question addressed here is whether a synthetic biochemical pathway for acetone production can be designed and demonstrated to function at 70°C or higher. If available, this would pave the way for the development of metabolic engineering strains that form acetone at temperatures high enough to facilitate its recovery and purification.

**Methods and Materials**

**Identification of thermophilic gene candidates**

Thermophilic homologs to *C. acetobutylicum* acetone pathway enzymes Thl (AAK80816.1), CtfAB (NP_149326 & NP_149327), and Adc (NP_149328) were identified by BLASTp (NCBI) searches limited to microbial groups known to be made up primarily of thermophilic organisms. Query results were considered by coverage (the length of the protein that demonstrates a minimum level of homology) and amino acid identity (the number of amino acids that are identical at a given position and the number of amino acids that share the same R-group category of amino acid at a given position). Potential candidate enzymes were narrowed down to the most promising by focusing on organisms with optimum growth temperatures above 70°C.
Protein expression

Genes encoding candidate enzymes were cloned from genomic DNA of *C. acetobutylicum* ATCC 824 (CA_P0165 = *C. acetobutylicum* Adc), and *C. subterraneus* (TTE0549 = *C. subterraneus* Thl, TTE0720 = *C. subterraneus*-CtfA, TTE0720 = *C. subterraneus*-CtfB), or synthesized to match genes from *T. melanesiensis* (Tmel_1136 = *T. melanesiensis* CtfA (first base changed from T to A to change start codon from TTG to ATG), Tmel_1135 = *T. melanesiensis* CtfB), *Vulcanisaeta distributa* (VDIS_RS01295 = *V. distributa* Adc) and *Sulfolobus* sp. (WP_009989587 = *Sulfolobus* Adc) (Integrated DNA technologies, Skokie, IL), and inserted into PCR amplified plasmid backbones pET-46, pRSF, or pCDF (EMD Millipore, Billerica, MA) with flanking regions appropriate for Gibson assembly, which was performed using NEB Gibson Assembly® Master Mix (New England Biolabs, Ipswich, MA). All constructs included N-terminal 6-histidine tags. In addition, all three versions of the Adc enzyme candidates were constructed without tags as well. Cell lines used for cloning were chemically competent *E. coli* cells 5-α (New England Biolabs) for plasmid screening and amplification, and Rosetta 2 (EMD Millipore) for protein expression. Protein expression was carried out in shake flasks containing one liter of ZYM-5052 lactose autoinduction medium (Studier, 2005) and appropriate antibiotics at 37°C for 20-24 hours. Cells were harvested by centrifugation at 10,000 x g for 10 minutes.

Protein purification

Harvested *E. coli* cells were re-suspended in 5 mL/mg pellet weight of immobilized metal affinity chromatography (IMAC) buffer A (300 mM NaCl, 50 mM sodium phosphate, 1 mM MgCl₂, 20 mM imidazole, 10% glycerol, pH 8 for *C. subterraneus* Thl, increased to 20% glycerol with 100 mM sodium sulfate added for all Ctf subunit candidates), or lysis buffer (50 mM potassium phosphate, pH 5.9) for Adc candidates, and lysed in a French pressure cell. Lysed cells were heat-treated at 65°C for 10 minutes to denature *E. coli* proteins, and then centrifuged at 24,000 x g for 20 minutes to generate soluble heat-treated cell extract. No further purification was performed for the untagged Adc candidates. The heat-treated cell extract for the untagged Adc candidate proteins was buffer exchanged into storage buffer (50 mM Tris-HCl, 100 mM sodium chloride, 50% glycerol, pH 7.5) in Vivaspin 20 10,000 Da MWCO filters (Sartorius, Goettingen, Germany). All histidine-tagged proteins were purified by IMAC using 5 mL HisTrap HP columns. Binding was in IMAC buffer A (described above), followed by elution in a gradient up to 500mM
Imidazole. Fractions containing the elution peak were pooled, concentrated, and buffer exchanged as described above. Storage buffers consisted of: 50 mM Tris-HCl, 100 mM sodium chloride, 1 mM DTT, 50% glycerol, pH 7.5, for *C. subterraneus* Thl; 50 mM MOPS, 500 mM ammonium sulfate, 50% glycerol for all Ctf subunits. Purified enzymes were stored at -20°C.

**Enzyme assays**

Individual enzymes were assayed for activity on a Lambda 25 spectrophotometer with PTP-1 Peltier heaters (Perkin Elmer, Waltham, MA) using 100 μL Quartz Cuvettes (Starna Cells, Atascadero, CA). All assays were performed at 70°C unless otherwise noted.

Thiolase was assayed by coupling to the NADH consuming activity of 3-hydroxybutyryl-CoA dehydrogenase (Hbd), as described previously (Loder et al., 2015). The assay mixture consisted of 100 mM MOPS pH 7.9, 0.3 mM NADH, and appropriately diluted enzymes, with Hbd in considerable excess. The reaction was started by adding 0.5 mM acetyl-CoA substrate, which Thl converts to acetoacetyl-CoA and free CoA. The simultaneous consumption of acetoacetyl-CoA and NADH by Hbd was followed by monitoring the decline in NADH absorbance at 340 nm.

Acetoacetyl-CoA:acetate CoA-transferase was monitored by tracking consumption of the Mg-enolate form of acetoacetyl-CoA by decrease in absorbance at 310 nm (Cary, Petersen, Papoutsakis, & Bennett, 1990). The assay mixture consisted of 100 mM Tris, 150 mM potassium acetate, 20 mM MgCl₂ and 5% glycerol at pH 7.5. Acetoacetyl-CoA was added, and absorbance was monitored for 30 seconds to establish the non-enzymatic rate of acetoacetyl-CoA hydrolysis, then an appropriately diluted mix of the Ctf subunits was added (subunit beta was always in slight excess). For inhibition studies substrate concentrations ranged from 15 to 720 mM acetate and 0.03 to 0.4 mM acetoacetyl-CoA.

The acetoacetate decarboxylase assay was adapted from (Ho, Ménetret, Tsuruta, & Allen, 2009). Assay mix consisted of 50 mM potassium phosphate, 300 mM lithium acetoacetate, pH 5.9. The assay was started by adding appropriately diluted enzyme, and consumption of acetoacetate was monitored by the decline in absorbance at 290 nm.

Thermal inactivation studies involved diluting enzymes to appropriate concentration in assay buffer (without substrate), incubating at 70°C for a range of times, which were then assayed for residual activity. Thermal loss of enzyme activity is typically modeled as a simple exponential
decay, but in some cases a more complex two-step inactivation model is necessary. This form of thermal inactivation can be conceptualized as initial enzyme E being converted to an intermediate E\textsubscript{1} with fractional activity \( \beta \), before becoming fully deactivated enzyme E\textsubscript{2} (Epting, Vieille, Zeikus, & Kelly, 2005). The parameter \( \beta \) and inactivation rate constants \( k\textsubscript{1} \) and \( k\textsubscript{2} \) give the fraction of remaining enzyme activity (y) as a function of time:

\[
y(t) = \left( 1 + \frac{\beta k_1}{k_2 \cdot k_1} \right) e^{-k_1 t} - \left( \frac{\beta k_1}{k_2 \cdot k_1} \right) e^{-k_2 t}
\]

Values of \( \beta, k_1, \) and \( k_2 \) were calculated in Microsoft\textsuperscript{®} Excel using the Solver function to minimize the sum of squared differences between the model and experimental data.

The three enzymes were assayed together as an \textit{in vitro} pathway in a mixture containing 100 mM Tris, 10 mM MgCl\textsubscript{2}, 150 mM potassium acetate, 5 mM acetyl-CoA, pH 7.5, with enzymes added to activities of 5 U/mL for Thl and Adc, 15 U/mL for Ctf. Controls consisted of reaction mixture with each enzyme missing individually, and a no-enzyme control. The resulting mixtures were incubated at 70°C in a thermocycler, and acetone was detected by gas chromatography (GC).

\section*{Other methods}

Acetone was detected on a GC-2014 gas chromatograph (Shimadzu, Kyoto, Japan) equipped with a ZB-WAXplus 30 m long, 0.53-mm ID capillary column (Phenomenex, Torrance, CA) and flame ionization detector. The GC oven temperature was initially held at 35°C for 3 min, increased to 150°C at 20°C/min, and held for 6 min. The injector was held at 220°C and FID detector at 300°C. Nitrogen was used as the carrier gas at a column flow of 30 cm/s. Samples of 0.1 µL were injected with a 1:10 split ratio using an AOC-20i autosampler.

Protein concentrations were determined by the Bradford method with BSA standards. SDS-PAGE was done using BioRad TGX 4 to 12% gels with standard Tris-glycine buffer. Samples were heated at 95°C for 15 minutes to ensure denaturation of the thermostable enzymes, and gels were stained with GelCode blue. Blue native PAGE used the Novex NativePAGE kit, with a 4-16% Bis-Tris gel. Imaging relied on running the gel with Dark Blue Cathode Buffer, followed by overnight de-straining in 10% acetate, 50% methanol solution.
Results

Extremely Thermophilic Acetone Pathway Candidates

BLAST searches for homologs to the *C. acetobutylicum* acetone pathway enzymes indicated several promising candidates (Table 6-2). The thiolase from *Caldanaerobacter subterraneus* subsp. *tengcongensis* (formerly *Thermoanaeobacter tengcongensis*) is 68% identical at the amino acid level, and in fact was already characterized and utilized as part of a synthetic pathway for n-butanol production, which starts with the same condensation reaction of two acetyl-CoA molecules to form the four carbon acetoacetyl-CoA (Loder et al., 2015).

*C. subterraneus* also provided a candidate for the acetoacetyl-CoA:acetate CoA-transferase, with 65% identity for both subunits, followed closely by *Thermosipho melanesiensis* with 53% and 64% identity for the alpha and beta subunits, respectively. The *T. melanesiensis* enzyme hit was also not surprising, having been identified in the search for a more moderately thermophilic acetone pathway functional at 55°C (Shaw et al., 2015). However, while that report confirmed activity by acetone formation in *T. saccharolyticum* at 55°C and room temperature activity assay of unpurified cell extracts, no detailed biochemical characterization of the enzyme or its thermostability was reported.

The same study utilized an acetoacetate decarboxylase from *Bacillus amyloliquefaciens* with significant sequence identity (65%) to *C. acetobutylicum* Adc, but this organism’s optimum growth temperature of 50°C was lower than desired for an extremely thermophilic pathway. Unfortunately, the acetoacetate decarboxylase is a highly specialized enzyme for solvent production, and the only extremely thermophilic candidates identified with a coverage score over 80%, from *Vulcanisaeta distributa* and a metagenomics sequence annotated as ‘from a *Sulfolobus* species’, had low sequence homology (< 30% amino acid identity). One surprising finding in the first report characterizing the activity of the *C. acetobutylicum* Adc (Davies, 1943) indicated that the enzyme retained activity up to 70°C, suggesting that the native enzyme may be viable in an extremely thermophilic pathway despite its mesophilic origins.

Recombinant Production of Acetone Pathway Enzymes

**Thiolase:** The thiolase from *C. subterraneus* had been purified and characterized previously (Loder et al., 2015) and Figure 6-2 confirms that the IMAC purified preparation utilized in the *in vitro* experiments forms the expected homotetramer.
**Acetoacetyl-CoA:Acetate CoA-Transferase**: The CtfAB complex (αβ2) has been purified from *E. coli* (Sramek & Frerman, 1975b) and *C. acetobutylicum* (Wiesenborn, Rudolph, & Papoutsakis, 1989), and in both cases is described as a fastidious enzyme requiring a buffer with at least 20 wt% glycerol and 500 mM ammonium sulfate to remain stable. In both previous studies, the alpha and beta subunits were co-purified as a complex, as did a subsequent study expressing the *C. acetobutylicum* CtfAB enzyme in *E. coli* (Cary et al., 1990). The salting-out purification described by these earlier reports was utilized here for mixtures of the *T. melanesiensis* and *C. subterraneus* subunits expressed individually in *E. coli*. This resulted in partially purified cell extracts with Ctf activity at 70°C, which was the first confirmation that the *C. subterraneus* enzyme functions as a thermophilic Ctf, but purified enzymes were needed for subsequent assays.

Histidine-tagged fusion proteins are readily purified (as was the case for Thl), but the tag can interfere with enzyme activity if the tagged N or C terminus is near the enzyme active site or dimerization interfaces. Fortunately, examination of a crystal structure of the *E. coli* CtfAB complex (PDB ID: 5DBN) indicated that all subunits N and C termini are at the outer surface of the heterotetramer, well separated from the subunit interfaces and active sites. Given the structural similarities between family-I transferases, even those from vastly different lineages (Coros, Swenson, Wolodko, & Fraser, 2004), the extremely thermophilic candidates here seemed likely to share this general structure. Therefore, N-terminal histidine-tagged alpha and beta subunits of both CtfAB complexes were expressed in *E. coli* for purification. An initial lysis in standard IMAC buffer (lacking glycerol and sulfate) resulted in cell extracts with a strong band for both *T. melanesiensis* and *C. subterraneus* alpha subunits on SDS-PAGE, but bands for the beta subunits were only visible in the insoluble cell fraction. Subsequent lysis in IMAC A buffer reformulated for Ctf purification was able to recover both subunits in soluble form, but yields of the beta subunits were low, since most was still lost to the insoluble fraction (particularly for *C. subterraneus*). Purification using IMAC buffers containing glycerol and sulfate resulted in highly purified alpha subunits, as well as purified *T. melanesiensis* CtfB, but the *C. subterraneus* CtfB subunit was only partially purified (78%), as determined by SDS-PAGE densitometry (Figure 6-2).

In contrast to previous efforts, individually purified Ctf alpha and beta subunits from separate expressions were obtained for biochemical characterization; the *E. coli* alpha subunit had been previously purified alone for structural but not biochemical analysis (Korolev et al., 2002). The results shed light on the difficulty of recovering active Ctf in previous studies. It appears that
the cause of Ctf’s low stability is specifically attributable to the beta subunit, since alpha subunits from each thermophile purified easily even without ammonium sulfate and glycerol as stabilizers. It is also possible that the beta subunit could only be recovered alone in this case due to the greater stability inherent in thermophilic proteins. Blue Native PAGE analysis confirms that the *T. melanensiensis* CtfAB complex exists as the expected heterotetramer (α₂β₂), although bands for the heterodimer and heterooctamer are also visible (Figure 6-2). The Ctf subunits were also loaded individually at equal levels, but both beta subunits are only faintly visible as broad smears – likely reflecting denaturation or aggregation in the gel, and further evidence of their low stability. No oligomers are visible for the *C. subterraneus* CtfAB complex, although this could be a result of either the poor stability or low purity of the beta subunit.

**Acetoacetate Decarboxylase:** The acetoacetate decarboxylase candidate from *Sulfolobus* sp. was expressed in both His-tagged and untagged forms and its presence as a soluble protein was confirmed via SDS-PAGE gel. Here, neither form demonstrated any acetoacetate decarboxylase activity at temperatures ranging from 40°C to 80°C. The *V. distributa* enzyme was also expressed in *E. coli* in tagged and un-tagged forms, but neither could be solubilized under any conditions tested. Inclusion bodies containing the protein of interest were solubilized by urea extraction and SDS-PAGE gel confirmed the presence of the protein in the inclusion bodies but the protein could not be refolded to soluble form via dialysis after many trials. In contrast, the *C. acetobutylicum* Adc was functionally expressed in both his-tagged and untagged forms, but the tagged enzyme exhibited significantly reduced activity (data not shown). The active enzyme complex is a homododecamer (Ho et al., 2009), and it is plausible that the tag interferes with subunit assembly or interferes with substrate access to the active site. Therefore, untagged *C. acetobutylicum* Adc was expressed and the cell extract was heat-treated, resulting in a surprisingly pure protein (84% according to densitometry), which Native PAGE confirmed formed the expected 12-subunit complex (Figure 6-2).

**Biochemical Characterization of Extremely Thermophilic Acetone Pathway Enzymes**

**Thiolase:** As the *C. subterraneus* Thl had been previously characterized (Loder et al., 2015), only thermal stability studies were conducted on this enzyme before use in the three enzyme pathway assays. The functional enzyme complex is a homotetramer displaying remarkable thermostability.
High temperature incubation actually increases the enzyme activity initially, before starting a slow, linear decline, such that there is a small amount of residual activity even after 20 hours at 70°C (Figure 6-3d).

**Acetoacetyl-CoA:Acetate CoA-Transferase:** Purification of the individual alpha and beta subunits of Ctf from *T. melanesiensis* and *C. subterraneus* allowed assay of the individual subunits for activity: neither subunit from either species was capable of catalyzing the CoA transfer reaction alone.

While we observed no activity from either the alpha or beta subunits alone, activity was observed using hybrid assemblies (*C. subterraneus* CtfA with *T. melanesiensis* CtfB and vice-versa), but the hybrid complexes exhibited specific activities roughly 1% of *T. melanesiensis* CtfAB. Interestingly, the individual subunits also displayed dramatically different thermostabilities; incubation of the beta subunit of either species at 70°C for a few minutes was enough to eliminate activity, while the alpha subunit showed no reduction in activity after an hour (Figure 6-3b). The combined subunits of *T. melanesiensis* exhibited a stability intermediate between the two pure subunits, with a half-life at 70°C of 95 minutes. Thermal degradation of *T. melanesiensis* CtfAB complex followed a two-step inactivation, where activity was rapidly reduced to approximately 60% in the first 15 minutes of high temperature incubation, but then declined much more slowly (Figure 6-3a).

Inactivation parameters (equation in Methods) for *T. melanesiensis* CtfAB at 70°C were: \( \beta = 0.64, \ k_1 = 0.324 \text{ min}^{-1}, \) and \( \ k_2 = 2.68*10^{-3} \text{ min}^{-1}, \) which implies a half-life of 96 minutes. The low purity and poor stability of *C. subterraneus* CtfB complicated activity assays, but based on limited data the *C. subterraneus* CtfAB enzyme complex appears to have thermostability comparable to *T. melanesiensis*, with residual activity evident even after 12 h at 70°C.

Observed specific activities for the purified *T. melanesiensis* CtfAB were similar to those reported for other versions of the enzyme. With substrate concentrations of 25 mM acetate and 0.2 mM acetoacetyl-CoA we observed 80 µmol/min/mg protein (or U/mg), while under comparable conditions *E. coli* acetoacetyl-CoA:acetate CoA-transferase has specific activity of approximately 150 U/mg (Sramek & Frerman, 1975a). The value obtained for purified *C. acetobutylicum* Ctf under similar assay conditions was 29.1 U/mg, compared to 0.36 U/mg in raw cell extracts (Wiesenborn et al., 1989). A specific activity of 3.57 U/mg was reported in cell extracts of
recombinant T. saccharolyticum over-expressing T. melanesiensis Ctf, although that study assayed the thermophilic enzyme at room temperature utilizing a significantly different method (Shaw et al., 2015).

**Acetoacetate Decarboxylase:** The results here confirm previous evidence of heat activation (Autor & Fridovich, 1970; Neece & Fridovich, 1967) with incubation at 70°C leading to an increase in activity over the first hour (Figure 6-3c), resulting in two-step inactivation parameters $\beta = 2.28$, $k_1 = 0.077$ min$^{-1}$, and $k_2 = 0.5 \times 10^{-3}$ min$^{-1}$. Since the activity of the intermediate enzyme E1 is greater than the initial, $\beta$ has a value above 1. Reported $k_{\text{cat}}$ values for C. acetobutylicum Adc at room temperature range from 165 s$^{-1}$ (Ho et al., 2009) to 1560 s$^{-1}$ (Highbarger, Gerlt, & Kenyon, 1996) with acetoacetate. Given that these values differ by roughly a factor of 10, and calculation of $k_{\text{cat}}$ depends on molar concentration of enzyme, it seems likely that the larger value was calculated using the molecular weight of the dodecameric holozyme (330 kDa), while the smaller used the subunit weight of 27.5 kDa. If this is the case, converting the $k_{\text{cat}}$ values to specific activity gives $V_{\text{max}}$ values of 360 and 280 U/mg, respectively. Here, C. acetobutylicum Adc exhibited maximal activities in excess of 1,000 U/mg at 70°C, an increase which is probably attributable to the higher assay temperature, or to heat-activation during high-temperature incubations (not mentioned in either reference above).

**Acetone Production in vitro by the Extremely Thermophilic Enzyme Pathway**

A mixture of all three enzymes converted acetyl-CoA and acetic acid to acetone at 70°C; omitting any one enzyme eliminated acetone production (Figure 6-5). Sample chromatograms are shown alongside a 2.5 mM acetone standard in assay buffer, which seems to distort the acetone peak, since it displays a broad right shoulder in both the standard and in vitro reaction. However, when acetone standards were analyzed in water, no such shoulder was present. The relative size of the peaks suggests evaporation or incomplete conversion, as the reaction stoichiometry would predict that 5mM acetyl-CoA with excess acetate would result in 5 mM acetone. The peak at 4.1 min matches the acetoacetic acid standard, which unsurprisingly is most prominent in the no-Adc control, but some acetoacetate is also visible as an intermediate in the full reaction.
Discussion

Acetone is a promising candidate for production via bio-reactive distillation in an extreme thermophile using a pathway requiring three enzymes which function up to at least 70°C (Figure 6-6). The previously characterized \textit{C. subterraneus} thiolase meets this requirement, as does the surprisingly thermostable \textit{C. acetobutylicum} acetoacetate decarboxylase, which is fortunate since the putative thermophilic acetoacetate decarboxylase homologs from this study exhibited no relevant activity. The ease with which untagged \textit{C. acetobutylicum} acetoacetate decarboxylase was separated from contaminating \textit{E. coli} proteins and brought to greater than 80% purity by simple heat treatment serves as a reminder as to why recombinant thermophilic proteins can be strategic for biochemical studies. As such, other unusually thermostable proteins from mesophiles could be identified simply by heat-treating mesophilic cell extracts to remove the most labile proteins. The proteins could be separated by various means including liquid chromatography and the identity of the given thermostable proteins could be identified by mass spectrometry. Additionally, screening the heat-treated supernatant for residual activities may provide insights into the types of thermostable enzymes present.

\textbf{Thiolase:} The \textit{C. subterraneus} Thl was characterized in detail previously (Loder et al., 2015). The functional enzyme complex is a homotetramer displaying remarkable thermostability with high temperature incubation increasing initial activity. The \textit{C. subterraneus} Thl also functions well in the desired direction (formation of acetoacetyl-CoA), with \( V_{\text{max}} = 74 \text{ U/mg} \) and a strong affinity for the acetyl-CoA substrate (\( K_M = 271 \text{ µM} \)). The strong affinity for acetyl-CoA is desirable in order for the thiolase to operate in the thermodynamically unfavorable acetoacetyl-CoA forming direction. Many thermophilic thiolases are known to catalyze the reverse reaction, cleavage of acetoacetyl-CoA to two acetyl-CoAs, and are prevalent in the carbon fixation cycle of the extremely thermoacidophilic Sulfolobales (Berg, Kockelkorn, Buckel, & Fuchs, 2007). Thus all enzyme parameters must be considered when seeking a thermophilic analog of a known mesophilic enzyme.

\textbf{Acetoacetyl-CoA:Acetate CoA-Transferase:} With a previously characterized extremely thermophilic thiolase and the surprising stable acetoacetate decarboxylase from \textit{C. acetobutylicum}, the missing link for acetone production was an extremely thermophilic acetoacetyl-CoA:acetate
CoA-transferase, two of which have been reported here. Either *T. melanesiensis* or *C. subterraneus* could serve as the basis for a thermophilic acetone pathway, since both complexes display sufficient activity and thermostability. In addition to reporting the first purified thermostable acetoacetyl-CoA:acetate CoA-transferase subunits alpha and beta and their heterotetramer complex (α2β2), the results here shed light on the properties of acetoacetyl-CoA:acetate CoA-transferases. The function of the hybrid *T. melanesiensis*/*C. subterraneus* Ctf complexes is likely due to the strong structural and sequence similarities among acetoacetyl-CoA:acetate CoA-transferases, particularly around the active site, and warrants further investigation to determine if this cross-functionality is evident among homologs from other species as well. Meanwhile, the dramatic difference in thermostability between the Ctf alpha and beta subunits, and poor stability of the beta subunits in general, helps to explain why this class of enzymes has been so challenging to purify in the past. The dramatically improved stability of the full CtfAB complex compared to the beta subunit highlights the role that the alpha subunit plays in stabilizing its partner, and serves as example of how important quaternary structure interactions can be for protein stability (the surprisingly thermostable dodecameric *C. acetobutylicum* acetoacetate decarboxylase is yet another example).

While Ctf complexes from *E. coli* and *C. acetobutylicum* have been purified and characterized for kinetics and substrate preferences, subunits from Ctf enzymes have not been purified separately for analysis. There is one instance of the co-purified *E. coli* subunits being separated by subsequent urea denaturation, which claimed the pure beta subunit had approximately 2% of the activity of the intact complex (Frerman & Duncombe, 1979). However, since the two subunit complex was used initially, the possibility of a small amount of residual contamination of active complex in the individual subunit fractions cannot be ruled out. The same report indicated that the alpha subunit was involved in structural support or maturation, which agrees with the findings here that it dramatically increases the stability of the beta subunit. At the same time, this rules out the possibility of the alpha subunit playing a catalytic role because it does not contain the nucleophilic glutamate involved in catalysis. More recent reports indicate the alpha subunit plays an important role in binding the CoA group (Korolev et al., 2002), in addition to the structural support it provides the beta subunit. The presence of highly conserved residues neighboring the active site in each subunit also suggests that both contribute to catalysis (Figure 6-4).
Family-I CoA transferases are known to exhibit Ping-Pong enzyme kinetics (also called “double displacement” or “substituted-enzyme” kinetics). In this case, one substrate binds the enzyme, is modified, and then dissociates as the first product, leaving the enzyme in a modified intermediate state, followed by binding of the second substrate, which is modified and dissociates as the second product, recovering the original enzyme. In the case of acetoacetyl-CoA:acetate CoA-transferase, the net reaction can be broken up into the two component steps:

Net Rxn: \[ \text{Acetoacetyl-CoA} + \text{Acetate} \rightarrow \text{Acetoacetate} + \text{Acetyl-CoA} \]
Step #1: \[ \text{Acetoacetyl-CoA} + \text{Enzyme} \rightarrow \text{Enzyme-CoA} + \text{Acetoacetate} \]
Step #2: \[ \text{Acetate} + \text{Enzyme-CoA} \rightarrow \text{Enzyme} + \text{Acetyl-CoA} \]

The two substrates and sequential nature of the reaction leads to unusual kinetics, where substrate inhibition appears at relatively low concentrations, but can be overcome by increasing the concentration of the other substrate (Wiesenborn et al., 1989). This unusual substituted-enzyme substrate inhibition is easily apparent on single and double-reciprocal plots (Cornish-Bowden, 1995), which matched kinetic data from this study for \( T. \) melanesiensis, confirming that catalysis with this enzyme proceeds through the same Ping-Pong mechanism observed in mesophilic versions. The relative thermal stabilities of the enzymes in the synthetic pathway should be taken into account when designing cloning constructs for recombinant expression in extremely thermophilic hosts. While the discovery that \( T. \) melanesiensis acetoacetyl-CoA:acetate CoA-transferase has a half-life of over 1 hour at 70°C indicates that the proposed extremely thermophilic acetone pathway is viable, it remains the least thermostable of the three pathway enzymes, inactivating more rapidly than the \( C. \) subterraneus thiolase.

**Acetoacetate Decarboxylase:** \( C. \) acetobutylicum Adc has been thoroughly characterized including the precise nature of its active site and catalytic mechanism which have been of interest to biochemists (Highbarger et al., 1996; Ho et al., 2009), and historically its surprising thermostability was a topic of study (Autor & Fridovich, 1970; Neece & Fridovich, 1967). Unexpectedly, the mesophilic \( C. \) acetobutylicum Adc is the most stable enzyme in this studied extremely thermophilic acetone pathway. The dramatic increase in activity at the beginning of high
temperature incubation meant that even after 12 h, activity was still greater than prior to heat-treatment.

**Acetone Pathway:** Given that Adc also appears to be the most active enzyme, it could be expressed at the end of an operon, or separately under the control of a moderate transcription level promoter. Thl and Ctf have comparable stabilities and activities, so equivalent expression would be appropriate. Given the instability of Ctf beta subunit, and the fact that all Ctf genes appear to exist in a *ctfAB* operon (often with overlapping start/stop codons), co-expression of the two subunits seems to be essential, especially at high temperatures.

As shown in **Figure 6-6**, only the reaction catalyzed by Adc is strongly favored thermodynamically (*k*<sub>eq</sub> significantly above 1). The thiolase reaction dramatically favors the reverse direction, as indicated by the very low equilibrium constant, such that intracellular acetoacetyl-CoA concentrations will be two to three orders of magnitude lower than acetyl-CoA. One way to drive the CoA transferase reaction forward is to increase the concentration of acetate, which is a known factor in driving the switch to solventogenesis in *C. acetobutylicum* (Wiesenborn et al., 1989). Using a host that is an efficient natural acetate producer and tolerates relatively high concentrations of acetate may facilitate acetone production. Fortunately, there are a number of fermentative extreme thermophiles that produce acetate as a major fermentation product and for which genetic tools are available (Loder et al., 2017; Zeldes et al., 2015).

Additionally, this acetate consuming pathway may be useful in converting the acetate present in the lignocellulosic biomass in the form of acetyl groups. Lignocellulosic biomass contains an appreciable amount of acetate (2-5 dry wt%) in the form of acetylated compounds, primarily found in the hemi-cellulose component (Kong et al, 1992), and is thus a potential carbon source for conversion to biofuels and biochemicals. During biomass processing, whether chemical or biological, the acetyl groups are cleaved from carbohydrates via hydrolysis resulting in free acetate (Pawar et al, 2013). Acetate is typically not metabolized or consumed by microorganisms targeted for conversion of lignocellulosic materials to biofuels via carbohydrate fermentation, yet there is potential for a carbon efficiency improvement if this fraction can be converted to a useful product. The next steps are to demonstrate that the extremely thermophilic acetone pathway can indeed be utilized in an extremely thermophilic host and to achieve titers necessary for efficient bio-reactive distillation. Efforts targeting this objective are underway.
Acknowledgements

This work was supported with grants to RMK and MWWA from the US National Science Foundation (CBET 1264052, CBET1264053) and to RMK US Dept. of Agriculture (2018-67021-27716). BMZ acknowledges support from an NIH Biotechnology Traineeship (2T32GM008776). CTS acknowledges support from a US Department of Education (P200A140020) GAANN Fellowship. JKO acknowledges support from an REU Supplement to NSF CBET 1264052.
Figure 6-1. Vapor-liquid phase envelopes for acetone and ethanol at various pressures of “bio-reactive distillation” operation. Phase envelopes on T-xy diagram features bubble point (lower line of phase envelope) and dew point (upper line of phase envelope) lines for acetone (solid lines) and ethanol (dashed lines). Biologically reasonable fermentation titers for acetone or ethanol is shown in the shaded region. The fermentation vessel (reboiler in batch distillation terminology) must operate within the phase envelope for “bio-reactive distillation.” The condenser operates at the dew point (top line of phase envelope) corresponding to the relevant product purity. Distillation stages operate at gradually decreasing temperatures from the bottom (reboiler) to top (condenser). (Data from VLE-calc.com)
Figure 6-2. SDS-PAGE and Blue-native PAGE of acetone enzymes SDS-PAGE indicates that all enzymes have the expected Mr and are at least 80% pure according to densitometry (except for *C. subterraneus* CtfB). Blue-native PAGE also confirms that enzymes exhibit the expected multimolecular arrangement, with the exception of *C. subterraneus* Ctf (possibly a result of the low purity of the beta subunit). Neither Ctf beta subunit is visible on Native PAGE, despite all lanes loaded at equal mass. The smear in both Ctf beta lanes suggests these subunits may have denatured.
Figure 6-3. Thermal stability of thermophilic acetone production enzymes Residual enzyme activity (y-axis - log scale) following incubation at 70°C for various times (x-axis - minutes), relative to initial enzyme activity (normalized to 1). (a) *C. subterraneus* Thl exhibits brief heat activation followed by a long, approximately linear decay. (b) *T. melanesiensis* Ctf in complex exhibits the common two-step thermal inactivation. (c) *C. acetobutylicum* Adc shows significant activation following 1 hr incubation, after which activity slowly declines. (d) Individual subunits of *T. melanesiensis* Ctf incubated separately have very different thermostabilities: alpha shows no decline after 1 h (linear trendline), while beta loses 90% of activity within 5 min (exponential decay).
Figure 6-4. Alignment of conserved active site residues from Family I CoA transferases

Residues neighboring the active site glutamate (red arrow) are highly conserved in Family I CoA transferase alpha and beta subunits. Boxes indicate Prosite entries PS01273 and 01274, the alpha and beta subunit sequence motifs. Pig succinyl-CoA transferase consists of a single peptide, but shares sequence and structural similarities with the heteromeric bacterial enzymes. Sequence alignments made in Geneious 8.1 (Kearse et al., 2012)
Figure 6-5. *In vitro* function of the full acetone pathway  *In vitro* function of the three enzymes (Rxn) at 70°C is confirmed by production of acetone. Omitting any one enzyme eliminated acetone production, although a peak consistent with the acetoacetate intermediate is visible in the no - Adc control.
Figure 6-6. Three enzyme pathway to acetone production  The enzymes catalyzing the three steps of the acetone pathway (along with equilibrium constants $K_{eq}$ calculated at 1mM concentration, pH 7, $T = 25^\circ$C for the reactions in the direction shown). High acetate concentrations drive the reaction forward to favor products by improving the Ctf enzyme kinetics and thermodynamics. The pathway can be used with a native acetate producer, where it simultaneously detoxifies acetate while generating acetone. Equilibrium constants from eQuilibrator (Flamholz, Noor, Bar-Even, & Milo, 2012)
Table 6-1. Solvent production in native and in metabolically engineered hosts

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Native</th>
<th>T (°C)</th>
<th>Reference</th>
<th>Engineered</th>
<th>T(°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetone</td>
<td><em>Bacillus macerans</em></td>
<td>43</td>
<td>(Weimer, 1984)</td>
<td><em>Thermoanaerobacterium saccharolyticum</em></td>
<td>55</td>
<td>(Shaw et al., 2015)</td>
</tr>
<tr>
<td>Ethanol</td>
<td><em>Caldicellulosiruptor sp.</em></td>
<td>72</td>
<td>(Svetlitchnyi et al., 2013)</td>
<td><em>Pyrococcus furiosus</em></td>
<td>78</td>
<td>(Basen et al., 2014)</td>
</tr>
<tr>
<td>n-Butanol</td>
<td><em>Clostridium thermodonocarolyticum</em></td>
<td>58</td>
<td>(Freier-Schroder et al., 1989)</td>
<td><em>Pyrococcus furiosus</em></td>
<td>60</td>
<td>(Keller et al., 2015)</td>
</tr>
</tbody>
</table>
Table 6-2. Enzyme candidates for thermophilic acetone pathway based on *Clostridium acetobutylicum* pathway.

<table>
<thead>
<tr>
<th>Source Organism</th>
<th>$T_{\text{opt}}$ (°C)</th>
<th>Accession number</th>
<th>% amino acid identity</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thiolase (Thl) – <em>C. acetobutylicum</em>: AAK80816.1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. subterraneus subsp. tengcongensis</em></td>
<td>70</td>
<td>WP_011024972</td>
<td>68% (99% cov)</td>
<td>Characterized, 70°C (Loder et al., 2015)</td>
</tr>
<tr>
<td><strong>Acetate acetoacetyl-CoA transferase (CtfAB) – <em>C. acetobutylicum</em>: NP_149326 &amp; 27</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. subterraneus subsp. tengcongensis</em></td>
<td>70</td>
<td>WP_009610465</td>
<td>65% (97% cov)</td>
<td>Homology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WP_011025123</td>
<td>65% (100% cov)</td>
<td></td>
</tr>
<tr>
<td><em>Thermosipho melanesiensis</em></td>
<td>70</td>
<td>WP_012057350</td>
<td>53% (96% cov)</td>
<td>Activity, 55°C (Shaw et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WP_012057349</td>
<td>64% (94% cov)</td>
<td></td>
</tr>
<tr>
<td><strong>Acetoacetate decarboxylase (Adc) – <em>C. acetobutylicum</em>: NP_149328</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Clostridium acetobutylicum</em></td>
<td>35</td>
<td>NP_149328</td>
<td>100 (100% cov)</td>
<td>Activity, 70°C (Davies, 1943)</td>
</tr>
<tr>
<td><em>Vulcanisaeta distributa</em></td>
<td>85</td>
<td>WP_013335399</td>
<td>29% (92% cov)</td>
<td>Homology</td>
</tr>
<tr>
<td><em>Sulfolobus contig</em></td>
<td>75</td>
<td>WP_009989587</td>
<td>24% (98% cov)</td>
<td>Homology</td>
</tr>
</tbody>
</table>
References


CHAPTER 7: Bio-acetone production from microcrystalline cellulose by the metabolically engineered extreme thermophile *Caldicellulosiruptor bescii*

Christopher T. Straub\(^1\), Jonathan K. Otten\(^1\), Lisa M. Keller, Benjamin M. Zeldes, Michael W.W. Adams\(^2\), and Robert M. Kelly\(^1\)

\(^1\)Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695

\(^2\)Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

To be submitted to: *Biotechnology & Bioengineering*
Abstract

The production of volatile industrial chemicals from renewable feedstocks in metabolically engineered extreme thermophiles offers the possibility of facilitated recovery and separations concurrently with fermentation. An excellent target chemical for such a process is acetone ($T_b = 56°C$), ideally produced from renewable resources such as lignocellulosic biomass. *Caldicellulosiruptor bescii*, an extremely thermophilic fermentative bacterium naturally capable for deconstructing and fermenting lignocellulosic biomass was evaluated as a candidate for bio-acetone production. When *C. bescii* was grown in a medium with 400 mM (23.2 g/L) of added acetone, no significant growth inhibition was observed, demonstrating that this bacterium can withstand such titers of the target product. *C. bescii* was then metabolically engineered to produce acetone from microcrystalline cellulose via genes encoding a biosynthetic pathway comprised of thiolase (ThI) from *Caldanaerobacter subterraneus* subsp. *tengcongensis* (Tte_0549), acetoacetyl-CoA:acetate CoA transferase from *Thermosipho melanesiensis* (Ctf) (Tmel_1135, Tmel_1136), and an unusually thermostable acetoacetate decarboxylase (Adc) from the mesophile *Clostridium acetobutylicum* (CA_P0165), all functionally expressed in the extreme thermophile. Acetone production was observed up to 78°C, but optimally at 70°C, with exogenously added acetate (50 mM). However, the natural competitive pathway to form acetate in *C. bescii*, along with the low enzyme activity and unfavorable binding affinity of the CoA transferase, limited acetone titers to 4.5 mM. This evidence indicates the Ctf enzyme to be the limiting step which is not surprising given its physiological role may not be to perform this reaction.
Introduction

Establishment of a renewable carbon economy for fuels, chemicals, and materials will require many platform chemicals to be produced from carbon neutral sources, such as grasses and trees (Isikgor & Becer, 2015). While the petrochemical industry produces thousands of different chemicals, the renewable carbon economy could flourish with the production of a small subset of chemicals that can be subsequently converted via traditional industrially established processes (Vennestrom, Osmundsen, Christensen, & Taarning, 2011). For example, ethanol can be utilized as a fuel but also as a source of ethylene via catalytic dehydration (Morschbacker, 2009). Ethylene can then be utilized for production of polyethylene and ethylene oxide. Ethanol alone cannot provide the basis for an entire renewable chemical economy; a broader range of bio-based chemical building blocks are needed, and this could include acetone which at one time was industrially produced through fermentation (Vennestrom, Osmundsen, Christensen, & Taarning, 2011).

Acetone-butanol-ethanol (ABE) fermentation from sugars by the mesophilic fermentative anaerobe Clostridium acetobutylicum has been of interest for well over a century (Gabriel, 1928). These solvents were produced on an industrial scale in the United Kingdom during World War I in order to obtain acetone for gunpowder manufacture. Production was ceased by mid-20th century as petroleum routes to these chemicals became more economically favorable (Jones & Woods, 1986). However, recent interest in ABE fermentation has focused almost solely on 1-butanol with efforts to eliminate the acetone branch of the pathway (Jiang et al., 2009). Butanol has attracted attention as a direct substitute for gasoline, however, the difficulty of separating butanol ($T_b = 118^\circ$C) from water and its comparatively high toxicity to microorganisms have brought significant challenges to the economic prospects for bio-butanol (Ibrahim, Ramli, Kamal Bahrin, & Abd-Aziz, 2017).

To date, there have been few efforts to produce acetone as the sole target product from renewable carbohydrate sources in anaerobes, potentially due to inabilities to develop a redox balanced fermentative metabolism. Formation of two acetyl-CoAs from glucose through central glycolysis results in eight electrons, carried by NADH and reduced ferredoxin (Table 7-1). Thus, the producing microorganism must have a route to discard the electrons. However, very few organisms produce $H_2$ at the theoretical limit required for a balanced acetone pathway (4 $H_2$/glucose), a capability almost exclusively limited to thermophiles due to thermodynamic drivers at elevated temperatures (Thauer, Jungermann, & Decker, 1977; Verhaart, Bielen, van der Oost,
Stams, & Kengen, 2010). Thus, electron balance concerns may also explain why acetone as the sole fermentation product has not been pursued as a metabolic engineering effort in mesophilic anaerobes.

Given that most anaerobic, fermentative thermophiles produce acetate as their primary fermentation product, there could be thermophilic organisms which produce acetone as a means of removing acetate, the physiological function of the acetone pathway in \textit{C. acetobutylicum}. Likewise, the acetone would more readily evaporate from the elevated temperature environment, alleviating fermentation product accumulation, providing yet another potential ecological advantage to such an organism. However, the highest reported temperature optimum for acetone production as a native metabolite is 43°C by \textit{Bacillus macerans} (Weimer, 1984). To date, a thermophilic counterpart to \textit{C. acetobutylicum} has not been reported nor has any more thermophilic microbe been identified with the capacity to natively produce acetone in any significant amount. An effort to metabolically engineer the ethanologenic thermophile \textit{Thermoanaerobacterium saccharolyticum} (\(T_{\text{opt}} = 55^\circ\text{C}\)) sought to install the acetone synthesis pathway in order to remove acetate, as acetate is known to inhibit the organism, limiting the desired ethanol titers (Shaw et al., 2015). As is observed with \textit{C. acetobutylicum} and the \textit{E. coli} strains expressing the \textit{C. acetobutylicum} pathway enzymes, substantial amounts of acetate must be present in the media to drive the acetone production pathway. For \textit{T. saccharolyticum}, growth in a medium with 550 mM glucose equivalents and 180 mM of added acetate produced 775 mM ethanol and 14 mM acetone, with a net 50 mM reduction of acetate (Shaw et al., 2015).

Since acetone producing microorganisms are rare in nature, the alternative is to turn to metabolic engineering, with particular interest in hosts that can use renewable feedstocks. Certain members of the extremely thermophilic genus \textit{Caldicellulosiruptor} can utilize most of the carbohydrate forms present in lignocellulose and ferment the hexose and pentose fractions (Dam et al., 2011). In fact, a \textit{Caldicellulosiruptor bescii} strain engineered to produce ethanol can solubilize and metabolize nearly 90% of the carbohydrate fraction of transgenic poplar (Straub et al., 2019). The key metabolic steps for conversion of sugars to acetone in metabolically engineered \textit{C. bescii} are shown in Table 7-1. Hexose and pentose sugars derived from lignocellulosic biomass are channeled through central glycolysis and the non-oxidative pentose phosphate pathway branch, ultimately generating acetyl-CoA. The redox cofactors are regenerated by production of hydrogen gas via the bifurcating hydrogenase (Bfh). The native \textit{C. bescii} metabolism converts acetyl-CoA

\textbf{Table 7-1.}
to acetate, obtaining ATP via acetate kinase (Ack). While the acetone production pathway described here does require acetate as a substrate for the CoA transferase, acetate generation would ideally be limited to only the quantities required to supply the acetone pathway.

Since neither \textit{C. bescii}, nor any \textit{Caldicellulosiruptor} species, natively produce acetone, genes encoding key enzymes for this pathway must be recruited from other thermophilic microorganisms and engineered into \textit{C. bescii}. Here, we show that acetone production in this extreme thermophile is possible and also consider the bottlenecks that must be overcome to realize commercially significant levels of acetone production by \textit{C. bescii}.

\textbf{Results and Discussion}

In addition to the creation of a functioning bio-acetone pathway in \textit{C. bescii}, there are other considerations if this host is to produce acetone at commercially relevant levels. A primary concern is the potential toxicity of the desired metabolite as well as any deleterious effects it may have on the overall metabolism. Second, there is a potential for an unsuspected metabolic pathway to convert acetone to another metabolite. In a fermentative anaerobe, there are significant quantities of reducing equivalents available that could act on the ketone functional group of acetone to potentially convert acetone to isopropanol. Before developing strains of \textit{C. bescii} for bio-acetone, these considerations were examined.

\textit{C. bescii} Acetone Tolerance

Formation of heterologous products by metabolically engineered hosts creates potential toxicity issues that must be considered. As mentioned above, this problem has plagued efforts with bio-butanol (Peabody \& Kao, 2016). Solvents, such as ethanol, butanol, and acetone, can disrupt the hydrophobic cell membrane, although some organisms have mechanisms to resist such solvents through efflux pumps and modification of cell membrane lipid composition (Ezeji, Milne, Price, \& Blaschek, 2010). The toxicity of acetone for fermentative anaerobes is generally unknown, as most studies have focused on butanol toxicity and reported either only the butanol toxicity level or the combined butanol and acetone toxicity level. Wild-type \textit{C. acetobutylicum} can tolerate up to 13 g/L of 1-butanol, although strains have been developed that can withstand 20 g/L (Ezeji et al., 2010). In \textit{C. acetobutylicum}, acetone levels of 200 mM (11.7 g/L) have been achieved concurrently with 154 mM (11.4 g/L) butanol and 80 mM (3.7 g/L) ethanol (Peabody \& Kao,
A genetically engineered strain of *E. coli* produced 122 mM (7.1 g/L) acetone in a fed-batch reactor system and a strain of *E. coli* expressing the acetone production pathway from *C. acetobutylicum* produced 154 mM (9.2 g/L) in a purged batch reactor (Bermejo, Welker, & Papoutsakis, 1998; May et al., 2013). Thus, some bacteria can tolerate titers of acetone on the order of magnitude required for a commercially viable process. While no detailed economic analyses of acetone production via fermentation are available, numerous studies have been performed on cellulosic ethanol toxicity where titers of approximately 40 g/L are desired (Kazi et al., 2010). The theoretical yield of acetone from glucose or xylose is substantially lower than ethanol (32 wt% vs 51 wt%) which is partially compensated for by a higher selling price (Straathof & Bampouli, 2017). Furthermore, acetone does not form an azeotrope with water and, thus, does not require molecular sieve drying for its separation into purified forms, potentially reducing production costs.

In order to determine the acetone tolerance of *C. bescii*, the wild-type strain was passaged into a defined medium containing 5 g/L cellobiose as the sole carbon source, with each subsequent passage containing higher levels of acetone added to the medium. The addition of acetone to the sealed serum bottle was completed after the medium was brought to anaerobic conditions (by degassing and purging cycles) in order to prevent removal of acetone during this process. While it was necessary to serially passage *C. bescii* to adapt the organism to higher concentrations of acetone, growth was only moderately affected at 400 mM (23.2 g/L). This was apparently the upper limit, as the strain did not grow when passaged onto 425 mM or higher levels of acetone after multiple attempts and acclimation strategies. The acetone toxicity effect is abrupt and the solvent had little impact on growth rate until it became completely inhibitory. Given that acetone readily crosses cell membrane, this demonstrates that *C. bescii* could tolerate such titers if acetone is generated *in vivo* rather than exogenously added.

**Metabolic Acetone Conversion to Isopropanol in *C. bescii***

The reaction of acetone to isopropanol can be catalyzed by a secondary NAD(P)H dependent alcohol dehydrogenase (often referred to as AdhB). Many members of the genus *Thermoanaerobacter*, which like *Caldicellulosiruptor* species belong to the Firmicutes, contain an AdhB, including the well characterized example from *Thermoanaerobacter ethanolicus* (Burdette, Vieille, & Zeikus, 1996; Pei et al., 2010; Zhou et al., 2017). In addition, the *C. subterraneus* subsp. *tengcongensis* genome encodes a putative secondary alcohol dehydrogenase (Tte_0695); in fact,
secondary alcohol dehydrogenase activity has been demonstrated from cell extracts of this bacterium (Soboh, Linder, & Hedderich, 2004). There are five putative alcohol dehydrogenases encoded in the *C. bescii* genome (Athe_0537, Athe_0928, Athe_2083, Athe_2206, Athe_2244), the functions of which are currently unknown. At acetone concentrations as high as 400 mM, even low levels of secondary alcohol dehydrogenase activity would likely result in some conversion to isopropanol. However, during acetone toxicity testing, no isopropanol was detected (as measured via gas chromatography), even when cultured in media containing 400 mM exogenously added acetone for toxicity testing. Thus, there was no indication that secondary alcohol dehydrogenase activity in *C. bescii* would convert acetone to isopropanol. However, the possibility exists for construction of such a pathway by addition of one of the aforementioned AdhB enzyme candidates, if isopropanol production is desired.

**Processing Advantages of Acetone Production in Extreme Thermophiles**

One potential advantage to producing acetone in an extreme thermophile is the elevated temperature where direct removal of acetone can be considered without the need for gas stripping. Acetone-water phase equilibria data (Eduljee, Kumarkrishnaroo, & Rao, 1958) were utilized to determine a feasible *in-situ* separation concept which would allow the fermentation broth to remain at 20 g/L (344 mM), while continuing to produce acetone, resulting in an “effective titer” in excess of the toxicity limit.

A titer of 20 g/L acetone is only 0.64 mol% of the liquid phase. However, the difference in normal boiling points (56°C for acetone, 100°C for water) provides for substantial enrichment in the vapor phase (Figure 7-2). At 70°C, considering only acetone/water phase equilibria, a pressure of 0.35 bar is required to reach saturation and here the vapor phase fraction of acetone is enriched significantly. While such a concept is feasible at mesophilic temperatures, the system would require operation at 0.05 bar which would subsequently require cryogenic temperatures (<-50°C) at the condenser to sufficiently capture over 99% of the acetone. The condenser temperature can be increased by compressing the off-gas, but this comes at significant expense (Haelssig, Tremblay, & Thibault, 2008). In this proposed high temperature fermentation, when the fermentation substrate is exhausted, the pressure would remain at 0.35 bar and the temperature would only need to be raised to 75°C to extract all of the acetone. Thus, such a process can circumvent the 23.3 g/L toxicity level of *C. bescii* and provide *in situ* separation.
Biosynthetic Pathway for Acetone Production

A biosynthetic pathway for acetone production in vitro was recently described and serves as the basis for bio-acetone production in C. bescii (Zeldes, Straub, Otten, Adams, & Kelly, 2018). Here, factors that need to be considered for in vivo functioning of this pathway in C. bescii are considered.

Thiolase: There are many reported thermophilic thiolases that convert two acetyl-CoA to acetoacetyl-CoA and vice-versa. One such thiolase from the extreme thermophile Caldanaerobacter subterraneus subsp. tengcongensis was previously characterized, utilized in a synthetic butanol production pathway, and subsequently demonstrated in vivo in Pyrococcus furiosus at 70°C (Keller et al., 2015). With a half-life of over three hours at 70°C and a V\text{max} of 74 U/mg (Loder et al., 2015), this thiolase should provide sufficient activity for the desired pathway. However, the equilibrium favors acetyl-CoA over acetoacetyl-CoA by multiple orders of magnitude (Lan & Liao, 2012).

While the acetyl-CoA concentration in C. bescii has not been measured, metabolic data on the fermentative mesophile C. acetobutylicum reported acetyl-CoA concentrations during the acidogenic phase and solventogenic phase were 1.67 mM and 0.46 mM, respectively (Amador-Noguez, Brasg, Feng, Roquet, & Rabinowitz, 2011). Furthermore, for E. coli in exponential phase during aerobic growth on glucose, quite different from the conditions for C. bescii, acetyl-CoA concentrations of 0.61 mM were reported. Utilizing these two estimates, along with the previously reported thermodynamic equilibria for this condensation reaction, the intracellular concentration of acetoacetyl-CoA is likely on the order of magnitude of 1 μM. Thus, the two downstream enzymes must provide a strong thermodynamic driving force, with along with favorable kinetics, to allow generation of acetone.

CoA Transferase: There are a wide range of CoA-transferases encoded in microbial genomes with multiple physiological functions. Given the successful use of the Thermosipho melanesiensis CoA transferase in T. saccharolyticum (Shaw et al., 2015) and in an in vitro pathway for bio-acetone (Zeldes et al., 2018), this enzyme was the leading candidate here. Previously, the physiological function of this enzyme was unknown and its biochemical characteristics had not been extensively characterized. However, the metabolic features of the microorganism natively producing the
enzyme in conjunction with genomic context can provide important insights. Both subunits of the *Thermosipho melanesiensis* CoA transferase are natively co-expressed with a 4 bp overlap at the end of the *ctfA* gene and the beginning of the *ctfB* gene. Only 19 bp separated the end of *ctfB* and the beginning of a putative thiolase. A stem loop with a predicted melting temperature of over 70°C, followed by a poly-T stretch, indicated that the operon was terminated after the thiolase and, thus, it is highly likely that the *T. melanesiensis* thiolase and CoA-transferase are within a single operon. The presence of the thiolase in an operon with the CoA-transferase indicates that acetoacetyl-CoA (or acetoacetate) is likely one of the native physiological substrates. *T. melanesiensis*, a member of the order Thermotogales, is an anaerobic, heterotrophic marine bacterium that grows on yeast extract and carbohydrates, producing acetate and H₂ as fermentation products, and can reduce elemental sulfur (Antoine et al., 1997). Unfortunately, this context does not provide much insight into the second substrate for the enzyme nor its precise physiological and metabolic role.

The thermal stability of the *T. melanesiensis* Ctf leads to a half-life of the heterodimer of over one hour (Zeldes et al., 2018). Previously unpublished data from that study provided estimates of kinetic parameters for this enzyme: $V_{\text{max}}$ of approximately 0.5 U/mg, with binding coefficients ($K_m$) for acetoacetyl-CoA and acetate of approximately 0.02-0.2 mM and 250 mM, respectively. The $V_{\text{max}}$ is substantially lower than that for *C. acetobutylicum* (30 U/mg) (Wiesenborn, Rudolph, & Papoutsakis, 1989), yet the binding affinity shows the same trend with the binding of acetoacetyl-CoA (0.021 mM) and acetate (1200 mM). The strong affinity for acetoacetyl-CoA is necessary due to the aforementioned low intracellular concentration of this compound.

The thermophilic *in vitro* pathway considered another Ctf from *C. subterraneus* subsp. *tengcongensis* (Zeldes et al., 2018). Here, genomic context provided a clear picture of the physiological function of this CoA transferase. Given its location in an operon with genes annotated for lysine fermentation, it appears to be an acetoacetate:butyryl-CoA transferase involved in lysine fermentation, a pathway in which lysine is broken down into ammonia, acetate, and butyrate (Kreimeyer et al., 2007) (*Table 7-2*). The enzyme transfers a CoA from butyryl-CoA to acetoacetate, forming acetoacetyl-CoA and butyrate. The butyrate is a waste fermentation product and a thiolase is then able to convert the acetoacetyl-CoA into two acetyl-CoAs, which can then be converted to acetate to obtain ATP through acetate kinase (Ack). Multiple transformations utilizing these genes for the acetone pathway did not yield any positive
transformants after multiple attempts. Any transformants that did grow had mutations in the CtfAB locus. The only difference between the two vectors was the source of the CtfAB, *T. melanesiensis* versus *C. subterraneus*. The Thl and Adc remained the same as well as the promoter, operon order, and ribosomal binding sites. Why this vector did not yield any positive transformants could not be determined.

**Acetoacetate Decarboxylase:** Despite extensive searches and expression of two putative acetoacetate decarboxylases, a functional thermophilic enzyme that was homologous to the *C. acetobutylicum* acetoacetate decarboxylase could not be identified (Zeldes et al., 2018). Fortunately, extensive studies on the acetoacetate decarboxylase from *C. acetobutylicum* indicated that it is remarkably stable, being activated by incubation at 70°C for one hour and thereby doubling its activity and further stable for many hours at 77°C (Autor & Fridovich, 1970). Subsequent work on the catalytic mechanism determined the $V_{\text{max}}$ to be 256 μmol/min·mg with a $K_m$ of 8.2 mM at 25°C (Highbarger, Gerlt, & Kenyon, 1996). A later study on the catalytic active site determined the $V_{\text{max}}$ to be 27 μmol/min·mg with a $K_m$ of 4.1 mM at 25°C (Ho, Ménétre, Tsuruta, & Allen, 2009). Further, the acetoacetate decarboxylase is more than twice as active at 70°C as at 25°C (Davies, 1943), even before thermal activation. Thus, the thermal stability and activity of this enzyme indicate that it is suitable for a thermophilic *in vivo* bio-acetone pathway.

**Driving Acetone Production in Engineered *C. bescii*: Acetate Addition to Medium**

The acetone pathway requires acetate to be present in order for Ctf to transfer a CoA from acetoacetyl-CoA to acetate to form acetyl-CoA. Acetate is the primary natural fermentation product for wild type *C. bescii*, unless there is sufficient H$_2$ pressure to instead induce lactate formation (van Niel, Claassen, & Stams, 2003). Ideally, the organism would produce acetate fermentatively which would then be consumed in the biosynthetic acetone pathway. Following strain construction and initial testing on defined media, approximately 1 mM acetone was produced. Upon examination of the acetone pathway work in *T. saccharolyticum* (Shaw et al., 2015), it seemed that further acetate addition was required to increase acetone production by improving the output of Ctf. In addition, the Ctf from *C. acetobutylicum* has an exceptionally strong binding affinity ($K_m = 0.021$ mM) for acetoacetyl-CoA, while the binding affinity was five orders of magnitude less favorable for acetate ($K_m = 1200$ mM) and butyrate ($K_m = 660$ mM).
Thus, at least for *C. acetobutylicum*, the binding of the organic acid is a bottleneck when acetate concentrations are low. This is consistent with its physiological purpose, i.e. only produce acetone and butanol (solventogenesis phase) when acid concentrations are high such that the low binding affinity of the acids is overcome by the increased concentration. While the exact physiological purpose of the CoA transferase in *T. melanesiensis* is not known, the binding parameters from the *C. acetobutylicum* enzyme indicated that addition of acetate was required to drive the reaction. This is necessary to push the equilibrium in the desired direction due to thermodynamic considerations as well as partially offset the poor binding affinity for acetate. Also, acetate addition may slow the kinetics of the generation of additional acetate that results from the acetyl-CoA conversion to acetate via phosphotransacetylase (pta) and acetate kinase (ack).

Since acetate is the primary natural fermentation product of wild type *C. bescii*, acetate consumption by the biosynthetic acetone pathway can be strategic from a metabolic engineering perspective. However, the aforementioned potential for a low binding affinity for acetate and the consequential low thermodynamic equilibrium present a challenge in which driving the reaction requires addition of exogenous quantities of acetate. Growth on concentrations of acetate up to 125 mM (7.5 g/L acetate) demonstrated that there is the expected favorable initial increase in acetone production and selectivity. However, when grown at its optimal pH of 7.2, acetate is slightly inhibitory to wild type *C. bescii* at 75 mM acetate, and moderately inhibitory at 150 mM (Basen et al., 2014). Thus, additional acetate should drive acetone production in *C. bescii* at sub-inhibitory levels. Yet, this was not observed above 50 mM (*Figure 7-3a*), possibly because other metabolic factors come into play that are not clear at this point.

**Optimizing Bio-Acetone Production in C. bescii**

In *C. acetobutylicum*, a drop in pH in the growth environment triggers a switch from acetic acid and butyric acid production to acetone, butanol, and ethanol (Long, Jones, & Woods, 1984). For *C. bescii*, the heterologous CoA transferase likely binds undissociated acetic acid rather than the anion, acetate, so that lowering medium pH may impact intracellular pH and thereby the concentration of undissociated acetic acid. *C. bescii* grows well between a pH of 5.0 and 8.0 and thus a range of pH was tested at 70°C (*Figure 7-3b*). However, the impact on acetone production was minimal.
While two of the enzymes were sourced from thermophiles, thus far it has been shown that the acetoacetate decarboxylase is the most stable enzyme with a half-life at 70°C of over 12 hours and activity up to 82°C (Autor & Fridovich, 1970). The thiolase half-life of 3 hours at 70°C and the optimal growth temperature of its source, *C. subterraneus* subsp *tengcongensis* (75°C), led to the notion that this would not likely be the bottleneck. The CoA transferase was suspected to be the enzyme preventing the pathway from directing more acetyl-CoA to acetone rather than acetate. The optimal temperature for the pathway was found to be 70°C, which happens to be the optimal growth temperature of *T. melanensiensis* ($T_{opt} = 70°C$, $pH_{opt} = 6.5$). However, acetone was detected in cultures grown as high as 78°C (Figure 7-3c).

The acetone pathway was expressed on a replicating plasmid in one operon under the control of the constitutive promoter for the surface layer protein (*slp*) (Athe_2303). Transcript levels relative to *gapdh* (Athe_1406) were measured and all found to be within a fold change of the *gapdh* gene. (Figure 7-4) This is to be expected as the four genes (*adc, thl, ctfA, ctfB*) are expressed as one operon under the control of the *slp* promoter. Given the possibility that the Ctf enzyme is the limiting factor, efforts to further increase expression may improve the pathway yield and titer.

**Conclusion**

Evidence is provided here that *C. bescii* is suitably resistant to acetone for production of industrially relevant titers. Further, a process is possible in which acetone can be simultaneously removed from the fermentation without the need for gas stripping, vacuum pumps, and cryogenic heat exchangers. Acetone production was observed with greater than 50% selectivity for acetone over acetate, albeit at titers of only 4.5 mM. The bottleneck to achieving high titers of acetone appears to be the acetoacetyl-CoA:acetate CoA transferase. Improvements in the Ctf acetate binding affinity and the $V_{max}$ would provide significant enhancement to the acetone production rate. Additionally, it may ultimately be desirable to down-regulate or completely eliminate the genes encoding enzymes that generate acetate from acetyl-CoA (phosphotransacetylase and acetate kinase) and supply an exogenous source of acetate for the pathway. Significantly higher titers will be necessary to realize the engineering and separations advantages of producing acetone in a thermophile, yet this is a step forward in generating acetone at such temperatures in an organism natively capable of deconstructing and metabolizing renewable lignocellulosic biomass.
Methods and Materials

Strain Generation

The acetone producing strain was created by expressing four genes on a kanamycin resistant replicating vector in C. bescii. Genetics were performed, as previously described (Lipscomb, Conway, Blumer-Schuette, Kelly, & Adams, 2016). Briefly, fragments for the thiolase and acetoacetate decarboxylase genes were obtained by PCR from genomic DNA of C. subterreneus subsp. tengcongensis and C. acetobutylicum, respectively. The DNA fragments for the CoA transferase genes from T. melanesiensis were obtained as synthetic gene fragments from Integrated DNA Technologies (Coralville, IA). A kanamycin resistant plasmid was constructed via Gibson assembly and cloned into NEB 5-α chemically competent cells. Following plasmid production and purification, the plasmid was sequenced via Sanger sequencing (Genewiz, South Plainfield, NJ). Methylation, transformation and plate purification were performed as described previously (Lipscomb et al., 2016). Freezer stocks of the strains were prepared by freezing at -80°C in 15% glycerol. Strain characterization utilized cultures taken from freezer stocks and adapted to the media conditions of the experiment.

Fermentation Product Analysis

Cultures were allowed to cool to room temperature. A 5 mL sample was taken and centrifuged to pellet the cells. The remaining liquid utilized for further analysis with 1 mL transferred to a vial for gas chromatography analysis (Shimadzu GC-2014) to measure acetone and to determine if isopropanol was present. Acetone was quantified via gas chromatography utilizing a Shimadzu GC-2014. Nitrogen was utilized as the carrier gas and detection via FID. For HPLC, 1 mL samples were acidified with 1 μL 50% sulfuric acid. A Rezex-ROA column (300 mm x 7.8 mm; Phenomenex) was utilized for separation with a Water Model 2414 Refractive Index detector for quantification.

qPCR Transcriptomic Analysis

Extracted RNA was reverse transcribed utilizing a Superscript Reverse Transcriptase (10 μL reactions with 1 μg purified RNA as template). Immediately following completion of the reverse transcription, 1 μL of each reaction was transferred to a qPCR plate containing 19 μL of
the Sso Superfast qPCR reaction mixture (SYBR Green Dye). Gapdh (Athe_1406) was utilized as a reference gene for comparison.

Acknowledgements
This work was supported in part by the US Department of Agriculture (2018-67021-27716) and the U.S. Department of Energy BER Award DE-SC0019391. CT Straub acknowledges support from a US DoEd GAANN Fellowship (P200A160061).
Figure 7-1. Conversion of lignocellulose to acetone, carbon dioxide, and hydrogen in metabolically engineered C. bescii  Shown in blue/grey is the native metabolism of C. bescii which deconstructs lignocellulosic biomass and converts hexose and pentose saccharides through the Embden-Meyerhof-Parnas Pathway (Central Glycolysis) and the non-oxidative branch of the pentose phosphate pathway (Non-Ox PPP). Shown in green/yellow are heterologous steps: Thl = thiolase from C. subterraneous subsp. tengcongensis; Ctf from T. melanesiensis; Adc from C. acetobutylicum.
Figure 7-2. Analysis of potential in situ separation of acetone from fermentation with an extreme thermophile at $T = 70^\circ C$. Increased temperature fermentation permits \textit{in situ} separation potential with any fermentation products with normal boiling points below that of water (i.e., isoprene, acetone, methanol, and ethanol).

*CO$_2$ liquid fraction in water is the fraction present as dissolved CO$_2$ and does not include bicarbonate (HCO$_3^-$) or carbonate (CO$_3^{2-}$).
Figure 7-3. Conversion of microcrystalline cellulose to acetone by metabolically engineered *C. bescii*. Acetone (blue circles), acetate (red triangles), and net acetyl-CoA equivalents (black squares) resulting from fermentation on 5 g/L crystalline cellulose (Avicel). Added acetate levels from 0 to 125 mM at $T = 70^\circ$C, pH 6.5 (A), starting medium pH levels at $T = 70^\circ$C with 50 mM acetate added, and (B) varying growth temperatures with initial medium pH 6.5 and 50 mM acetate added to medium.
Figure 7-4. Transcription of heterologously expressed genes for synthetic acetone pathway in C. bescii relative to gapdh (Athe_1406). Three acetone pathway enzymes encoded by four genes (thl, ctfA, ctfB, adc) were expressed on a replicating plasmid in C. bescii. The plasmid also contains the high temperature kanamycin resistance gene for maintained selection. Transcript levels were obtained via qPCR from cultures grown on defined media (5 g/L cellobiose) and harvested in mid-exponential phase.
References


227
CHAPTER 8: Analysis of a Modification in Pyrococcus furiosus Central Metabolism and Its Implications for Metabolic Engineering

Christopher T. Straub\textsuperscript{1}, Gerritt Schut\textsuperscript{2}, Jonathan K. Otten\textsuperscript{1}, Lisa M. Keller\textsuperscript{1}, Michael W.W. Adams\textsuperscript{2}, and Robert M. Kelly\textsuperscript{1}

\textsuperscript{1}Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC 27695

\textsuperscript{2}Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

To be submitted to: Metabolic Engineering
Abstract

The modified Embden-Meyerhof glycolytic pathway found in hyperthermophilic archaeb, such as *Pyrococcus furiosus* (*T*<sub>opt</sub> 100°C), has many implications with regard to establishing such organisms as metabolic engineering platforms. The key difference in this modified pathway occurs at the conversion from glyceraldehyde-3-phosphate (GAP) to 3-phosphoglycerate (3-PG) where the typical intermediate 1,3-bisphosphoglycerate (1,3-BPG) is ‘skipped’. The resulting bypass and absence of the typical ATP yielding step performed by phosphoglycerate kinase (Pgk) has dramatic impacts on the energy yield, reducing power, and the kinetics of carbohydrate metabolism. In order to understand such perturbations at this critical juncture, two enzymes responsible for this bypass reaction in *P. furiosus*, glyceraldehyde-3-phosphate ferredoxin oxidoreductase (*gapor*) and NADP<sup>+</sup> dependent non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (*gapn*) were examined independently in their capability to support this step. Unlike previous work with *Thermococcus kodakarensis*, *P. furiosus* was viable with one of these enzymes without the presence of the other, albeit with differences in native fermentation product profiles. *P. furiosus* was viable in the gluconeogenic direction (growth on pyruvate or peptides plus sulfur) in a Δ*gapn*Δ*gapor* strain but was non-viable upon attempts to grow under glycolytic conditions (grown on maltose or cellobiose). Also, bio-based chemical production in engineered strains of *P. furiosus* was examined, using ethanol as a proxy for heterologous products (e.g., 1-butanol, isobutanol, fatty acids, 3-hydroxypropionate, etc.) that require reducing equivalents (e.g., NAD(P)H, reduced ferredoxin) generated from glycolytic metabolism. A high temperature NADPH-dependent primary alcohol dehydrogenase was identified to produce ethanol via the previously established aldehyde oxidoreductase (AOR) pathway, but at a temperature closer to its growth optimum. The insertion of a single gene encoding this enzyme, AdhA (Tte_0696) from *Caldanaerobacter subterraneus*, demonstrated high selectivity of ethanol over acetate (> 8:1) and enabled ethanol production up to 85°C, the highest temperature for bio-ethanol production reported to date. Further examination of these results provides clarity on the suitability of *P. furiosus* as a potential host for industrially relevant commodity products.
Introduction

Utilization of anaerobes to convert renewable feedstocks into bio-based chemicals requires striking a balance between carbon diverted to meet cellular requirements (i.e. proteins, lipids, and nucleotides) and carbon channeled to fermentation products as a necessity for energy metabolism due to the absence of inorganic electron acceptors such as O\textsubscript{2}. In an ideal case for industrial production of commodity products, the organism obtains sufficient energy, carbon, and other nutrients to maintain metabolic activity while minimizing undesirable by-products such as carbon dioxide, acetate, and cell mass. In many fermentative anaerobes, carbohydrates are processed through the canonical central glycolytic pathway, often in conjunction with a pathway to assimilate pentose sugars toward glyceraldehyde-3-phosphosphate which is further processed in the downstream portion of Embden-Meyerhof (E-M) glycolysis (Müller, 2008). The full E-M pathway results in a net energy gain (2 ATP per glucose equivalent) and the resulting pyruvate and electrons (in the form of reduced cofactor NADH) can be directed to “build” waste fermentation product, such as lactate, ethanol, butanol, and others (Müller, 2008). The cell will require some of the carbon to be catabolized, but for biotechnological purposes significant amounts of carbon should be incorporated into target bio-products.

One such potential fermentative anaerobe with broad substrate utilization is the extreme thermophile, \textit{Pyrococcus furiosus} (\textit{T\textsubscript{opt} = 100°C}). Its native capacity to metabolize renewable marine feedstocks, such as chitin (Gao et al., 2003) and laminarin (Gueguen et al., 1997), as well as \textalpha- and \textbeta- glucans (Kengen et al., 1993) and peptides (Fiala and Stetter, 1986), presents an opportunity for metabolic engineering to produce bio-based chemicals of interest. A naturally competent strain of \textit{P. furiosus} (\textit{T\textsubscript{opt} = 100°C) has led to development of a facile genetic system (Lipscomb et al., 2011), permitting the creation of strains producing ethanol (optimally at 72°C) (Basen et al., 2014), 1-butanol (at 70°C) (Keller et al., 2015), lactate (Basen et al., 2012), and 3-hydroxypropionate (72°C) (Keller et al., 2013; Lian et al., 2016). However, no metabolically engineered products have been demonstrated at or above 80°C.

\textit{P. furiosus} has many potential advantages as a platform organism for metabolic engineering. Its low redox potential ferredoxin (-500 mV) (Park et al., 1991) can provide reducing power to redox reactions that are typically thermodynamically unfavorable with “normal” electron carrying cofactors of typical redox potential. This was demonstrated in a strain of \textit{P. furiosus} engineered to reduce acetate to acetaldehyde and further to ethanol (Basen et al., 2014). An
additional advantage is a soluble hydrogenase, capable of interconverting H₂ and NADPH, allowing many practical applications for utilization of H₂ as a source of electrons in product synthesis (Mertens et al., 2003).

A potential disadvantage of note is the energy yield from the modified E-M pathway present in *P. furiosus*. The lack of net ATP yield from glucose to pyruvate is a result of the pathway which converts glyceraldehyde-3-phosphate (GAP) to 3-phosphoglycerate (3-PG). Rather than utilizing a two-step reaction to transfer electrons to NADH followed by an ATP formation step, the single step produces reduced ferredoxin (Fd_red) without substrate level phosphorylation (Kengen et al., 1994) (Figure 8-1). However, energy conservation occurs via the membrane-bound hydrogenase which utilizes reduced ferredoxin to simultaneously reduce protons to H₂ and create a proton gradient with subsequent ATP generation (Sapra et al., 2003) (Figure 8-2a).

Since *P. furiosus* does not natively produce fermentation products of industrial interest, it must be engineered to do so. However, because of its high growth temperature, enzymes are often recruited from less thermophilic organisms to construct biosynthetic pathways to function in *P. furiosus*. So, even though *P. furiosus* has been engineered to produce ethanol (Basen et al., 2014), it must be grown at temperatures far below its optimum. If the temperature for bio-based chemical production in *P. furiosus* can be raised, a key ancillary benefit would be operation closer to the boiling point of the target, thereby facilitating recovery processes.

**Results and Discussion**

The first objective pursued here was investigating the essentiality of GAPOR and GAPN in the glycolytic pathway natively found in *P. furiosus* as a first step in re-designing glycolysis to enhance bioenergetic yield with an eye towards metabolic engineering. The second objective was to determine how such changes might impact heterologous pathways which require the electrons produced during glycolysis, such as the previously demonstrated aldehyde oxidoreductase (AOR) pathway at 72°C (Basen et al., 2014) (Figure 8-2b). To demonstrate this effect at a temperature closer to its optimum, an alcohol dehydrogenase capable of supporting high yield ethanol production above 72°C was sought, a point at which the doubling time of *P. furiosus* is only 15% of its optimal (Fiala and Stetter, 1986). Thus, a more thermophilic alcohol dehydrogenase was needed to support ethanol production at or above 80°C, a point at which the doubling time of *P. furiosus* is 60% that of its optimum.
Gapor or Gapn Alone Can Support Glycolytic Growth in *P. furiosus*

The essentiality of Gapor and Gapn in the *P. furiosus* glycolytic pathway was examined by deleting each of the genes and assessing the resulting phenotype. First, it was desired to delete both *gapor* and *gapn* to determine if there was another pathway supporting the conversion of GAP to 3-PG present in *P. furiosus*. Comparison of the genome of wild type *P. furiosus* and COM1 revealed numerous changes (Bridger et al., 2012). In fact, there is a 300 bp deletion in the the COM1 *gapn* gene resulting in a lack of an in-frame start codon; thus COM1 is essentially Δ*gapn*. Deletion of *gapor* from COM1 resulted in a Δ*gapor* Δ*gapn* (RK300) strain. The absence of these genes did not affect gluconeogenic growth (grown on pyruvate or peptides plus sulfur) but growth was not observed on either maltose or cellobiose, even after multiple adaptation strategies. Thus, it appears that there are no other pathways in *P. furiosus* that than can bridge from glyceraldehyde-3-phosphate to 3-phosphoglycerate.

Given that COM1 lacks Gapn, it can be concluded that Gapor alone is sufficient for the glycolytic pathway. This differs from findings with the closely related archaeon *Thermococcus kodakarensis* in which Δ*gapn* were not viable (Matsubara et al., 2011). It was further postulated that Gapn plays a key role in providing NADPH for *T. kodakarensis* and this was the reason that the strain was not viable for glycolytic growth. However the close relationship between the *T. kodakarensis* and *P. furiosus* along with the presence of annotated ferredoxin:NADP reductases in the *T. kodakarensis* genome would seem to suggest that *T. kodakarensis* can easily transfer electrons from reduced ferredoxin to generate NADPH (Fukui et al., 2005).

In order to determine if Gapn alone could support growth in the glycolytic direction, *gapn* was constitutively expressed under control of the S-layer protein (PFC_06220) promoter while also deleting *gapor*, resulting in strain RK301. Surprisingly, this strain was able to support glycolytic growth, albeit not to the same cell densities as the COM1 parent strain. With Gapn functioning as the sole link between GAP and 3-PG, significant quantities of NADPH would replace reduced ferredoxin as electron carriers. *P. furiosus* does have an NADPH-dependent hydrogenase (soluble hydrogenase I – SHI) as a means to dispose of electrons and recycle NADP⁺ (Bryant and Adams, 1989). The redox potential of NADPH ($E_0 = 0.32$ mV) compared to the *P. furiosus* reduced ferredoxin ($E_0 \sim 0.55$ mV at 90°C) makes it likely that H₂ build-up in the closed serum bottle used for cultivation would affect growth. The lower cell densities may be result of H₂ build-up eventually ceasing this route of electron disposal.
Of note, this strain (RK301) was created by growth and plating on defined medium with pyruvate as the carbon source. There was a potential that constitutive expression of gapn would result in a futile bioenergetic cycle during gluconeogenesis (Figure 8-3). However, these strains grew more rapidly on pyruvate than either the parent COM1 or the Δgapor Δgapn indicating that the high levels of GAPN may provide some advantage for gluconeogenic growth. Allosteric regulation of gapn may also play a role, as the Thermoproteus tenax Gapn (TTX_RS05635 – WP_014127065.1) was stimulated by high levels of glucose-1-phosphate and fructose-1-phosphate, compounds present in high quantities during glycolysis (Lorentzen et al., 2004). Given that the P. furiosus Gapn (Wild type: PF_RS03810 - WP_011011884.1) is highly homologous to the T. tenax Gapn, with conserved amino acid residues in the known allosteric regulation sites, this may be preventing such a futile cycle via inhibition of Gapn activity.

Utilizing a Redox Balanced Ethanol Pathway in P. furiosus

While utilization of either Gapor or Gapn does have an effect on growth and amount of fermentation products, it would not be expected to chance the types of compounds with the electron disposal via H2. Many heterologous pathways require utilization of the reduced electron carriers (e.g., NAD(P)H or Fdred) to generate a more reduced final fermentation product (ethanol, butanol, lactate, alkanes, etc). The utilization of either Gapor, generating reduced ferredoxin, or Gapn, generating NADPH, could potentially affect the heterologous pathway. To explore this concept, a heterologous pathway which consumes most of the electrons generated from central glycolysis and Pfor must be identified. To utilize the previously reported AOR pathway at temperature more favorable to P. furiosus growth, a higher temperature primary alcohol dehydrogenase was sought.

Characterization of an extremely thermophilic primary alcohol dehydrogenase

A novel ethanol pathway was previously demonstrated in P. furiosus in which the acetate naturally produced by the P. furiosus glycolytic metabolism is converted to ethanol with two redox reactions involving the native aldehyde oxidoreductase (AOR) and an inserted alcohol dehydrogenase (Basen et al., 2014) (Figure 8-2b):

**Reaction #1:** Acetate + Fd_{red} → Acetaldehyde + Fd_{ox}  
(AOR)

**Reaction #2:** Acetaldehyde + NAD(P)H → Ethanol + NAD(P)^+  
(AdhA)
This pathway operates optimally at 72°C (with ethanol above control levels detected up to 78°C). The inability of this pathway to operate at higher temperatures is likely due to the lower optimal growth temperature of the organism (Thermoanaerobacter sp X514 (T_{opt} 60°C) (Roh et al., 2002)) from which the alcohol dehydrogenase was recruited. While the doubling time of P. furiosus at 80°C is approximately 70 minutes compared to 38 minutes at its optimal 100°C, there is a precipitous drop-off as the temperature decreases; the doubling time at 72°C is over 250 minutes (Fiala and Stetter, 1986). Thus, identification of an enzyme capable of functioning in this pathway above 80°C would be advantageous to potential applications and be more consistent with P. furiosus metabolism.

Based upon the reported and characterized alcohol dehydrogenase enzymes from various members of the genus Thermoanaerobacter, a pBLAST search was conducted based upon amino acid sequence homology, seeking other high temperature organisms that may produce a more stable enzyme. A strong homolog (Tte_0696) was found in the Caldanaerobacter subterraneus subsp. tengcongensis (T_{opt} 75°C) genome (Fardeau et al., 2004), with 86% amino acid homology to the T. sp. X514 alcohol dehydrogenase. In addition, a study on a C. subterraneous membrane-bound hydrogenases assayed cell extracts for alcohol dehydrogenase activity and postulated that ORF Tte_0696 was responsible for the primary alcohol dehydrogenase activity (Soboh et al., 2004).

Expression and characterization of the histidine-tagged, purified enzyme confirmed the acetaldehyde to ethanol activity with NADPH as eight times higher than NADH (Table 8-1). Assays to determine alcohol specificity demonstrated much higher activity on ethanol rather than slightly longer primary alcohols, such as propanol and butanol and isobutanol (Figure 8-4). The enzyme showed no measurable activity on the simplest secondary alcohol, isopropanol, and no activity on the simplest tertiary alcohol, tert-butanol. Thus, it appeared that this enzyme was suitable for use to drive the ethanol production pathway.

**In vivo AOR Pathway with C. subterraneus AdhA**

The AdhA, under control of the S-layer protein promoter was inserted into the naturally competent strain of P. furiosus (COM1) downstream of the gapor gene, creating strain RK302. Growth on 5 g/L cellobiose at temperatures ranging from 75°C to 95°C for 72 h resulted in optimal performance at 80°C; at 90°C, ethanol production returned to background levels observed in the
parent strain (COM1) (Figure 8-5a). Insertion of this AdhA into *P. furiosus* allowed exploration of a pathway, at temperatures more in line with *P. furiosus* optimum, that utilizes electrons for reduction of metabolites rather than hydrogen production.

**Gapn and Gapor Effects on Ethanol Pathway in *P. furiosus***

Utilizing the AOR pathway, strains were created that could provide insight into the contributions and effects of Gapn and Gapor on the generation of redox cofactors. While the strain that contains Gapor produces ferredoxin, *P. furiosus* easily transfers these electrons to NADPH via a ferredoxin:NADP⁺ oxidoreductases (Ma and Adams, 2001; Nguyen et al., 2017). Thus, there is an ample supply of NADPH for the alcohol dehydrogenase. It is also possible that some of the reduced ferredoxin does generate H₂, but that this H₂ is utilized by soluble hydrogenase I to generate NADPH. As expected, a strain utilizing Gapor and an AdhA generated ethanol at a high ratio compared to acetate. Titers of ethanol over 1 g/L were generated (Figure 8-5b).

Replacement of Gapor with Gapn in the AOR pathway would generate a pathway that is redox carrier balanced. The NADPH generated from Gapn would be recycled by the AdhA and likewise, the ferredoxin produced by pyruvate ferredoxin oxidoreductase (Pfor) would be recycled by the ferredoxin dependent aldehyde oxidoreductase (AOR). Strain RK303 was created with *gapn* and the *adhA* expressed under control of the S-layer protein promoter with a *gapor* deletion. Strain RK303 produced only about half the amount of ethanol as strain RK302 (Figure 8-5b). This may be due to the concentrations of individual reducing equivalents present. The reduction of acetate to acetaldehyde with ferredoxin is much less thermodynamically favorable than the final step, reduction of acetaldehyde to ethanol with NADPH. Thus, a larger pool of reduced ferredoxin may be more advantageous for the ethanol pathway, even though some electrons must be transferred from reduced ferredoxin to NADP⁺. If Gapn generates all of the NADPH that is required and Pfor generates all of the reduced ferredoxin that is required for this pathway, any further transfer of electrons from reduced ferredoxin to NADP⁺ would result in an excess of NADPH. If any Fd:NADP⁺ reductases are present, as they may be constitutively expressed, there would be an excess of NADPH. This potential draining of the ferredoxin pool and excess NADPH would presumably run into the same H₂ limitation that was observed in RK301.

Utilizing Gapn in place of Gapor to directly generate NADPH required for synthesis of fermentation products may not be advantageous as multiple routes exist for *P. furiosus* to transfer
electrons from ferredoxin to NADPH (Ma and Adams, 2001; Nguyen et al., 2017). However, for a pathway that requires only NADPH for reducing equivalents, such as the 3-hydroxypropionate or fatty acid synthesis pathways, the direct generation of NADPH by Gapn may be more favorable.

Conclusions

The potential of *P. furiosus* as a platform for the production of renewable fuels and chemicals is primarily supported by its ability to natively metabolize renewable biopolymers, such as laminarin, chitin, polysaccharides, and peptides. From a metabolic engineering perspective, its many well studied redox handling enzymes provide for transfer of redox potential to either generate hydrogen gas or provide electrons in the form of NAD(P)H or reduced ferredoxin for synthesis of biomolecules. Likewise, its efficient genetic system and smaller genome (2 MB) provide a case for a more pliable and accessible platform organism. However, if acetyl-CoA is to be redirected to heterologous pathways that also consume the electrons provided by central metabolism, the organism will be unable to obtain its required energy from the central glycolytic metabolism, requiring another energy source. Here, we illuminate the effects of Gapor and Gapn on the production of the primary native metabolites, acetate and hydrogen, as well as the effects of such changes to an engineered pathway to ethanol.
Methods and Materials

Protein Expression

A codon optimized gene fragment of Tte_0696 (C. subterraneus AdhA) (Integrated DNA Technologies, Coralville, IA) was inserted into the pET-46 backbone containing a 6-histidine residue tag on the C-terminal end. Plasmid construction was performed via Gibson assembly and transformed into New England Biolabs (NEB) 5-α. Following plasmid purification, the sequence was verified by Sanger sequencing (Genewiz). The plasmid was then transformed into Novagen Rosetta\textsuperscript{TM} (DE3) protein expression strain. Protein was expressed with auto induction media (ZYM 5052) for 18-24 h. Cells were pelleted by centrifugation at 6,000 x g for 10 minutes, lysed via French Press and heat treated at 65°C for 20 minutes to precipitate thermolabile proteins. The resulting lysate was centrifuged at 12,000 x g for 30 min and the supernatant was sterile filtered (0.22 μm). The presence and purity of the protein of interest was checked via SDS-PAGE (Bio-Rad TGX stain free gel). The protein solution was concentrated utilizing Sartorius Vivaspin 20 (10,000 MW cut-off) protein concentrator conical. During the concentration, the protein was also buffer exchanged into a protein storage buffer (50% glycerol). The resulting concentrated protein was again checked via SDS-PAGE. Protein concentration was measured via Bradford assay.

Enzyme Assays

The alcohol dehydrogenase was assayed at 75°C utilizing a Perkin Elmer Lambda 25 UV/Vis spectrophotometer. The reaction progress was monitored via absorbance at 340 nm, a wavelength at which NAD(P)H strongly absorbs, but NAD(P)\textsuperscript{+} does not. The reaction was performed in quartz cuvettes containing 200 μL of total reaction. Concentration of the reagents was buffer (50 mM Tris-HCl pH = 7.5), protein (variable), NAD(P)(H) (2 mM), and the substrate (aldehyde or alcohol) (12 mM). Reactions without substrate were utilized as controls to account for thermal degradation of the co-factor and/or protein.

Strain Generation

All strains of \textit{P. furiosus} were generated in the naturally competent COM1 (ΔpyrF), as described by (Lipscomb et al., 2011). Briefly, the target region for integration was identified and 500 bp flanking regions segments were amplified from gDNA. A copy of the pyrF gene, under control of the glutamate dehydrogenase (gdh) (PFC_07240) promoter (P\textsubscript{gdh}), was placed...
downstream of the 5’ flanking region. Any additional genes to be inserted were placed downstream
of the pyrF cassette under the control of the surface layer protein promoter (P_{slp}). An apramycin
resistant plasmid was constructed via Gibson assembly and cloned into NEB 5-α competent cells.
Following plasmid production and purification, the plasmid was sequenced via Sanger sequencing
(Genewiz). A PCR product was generated from the plasmid template with primers at the ends of
the 500 bp flanking regions. The plasmid was digested with enzyme DpnI (New England Biolabs,
product #R1076S) to remove any residual plasmid template and purified by a Qiagen PCR
purification kit (Cat. No. 28104).

Transformation was accomplished by growing the COM1 parent strain (from freezer stock)
overnight on defined medium (described in growth conditions) at 85°C. It was removed and
allowed to cool for 1-2 h at room temperature. Approximately 2 μg of the PCR product was mixed
with the cooled competent cells and immediately plated (glass) on defined media lacking a uracil
source. The glass plates were placed inside a pressure vessel and degassed with 20/80 CO_{2}/N_{2} to
bring to anaerobic conditions. The pressure vessel was placed in an oven for 3-5 days. After
allowing the pressure vessel to cool for 2 h, the vessel was opened and colonies were picked into
defined uracil deficient liquid media and brought to anaerobic conditions. Upon culturing for 24-72
hours at 85°C, 5 mL samples were taken and the pellet was centrifuged for extraction of
genomic DNA utilizing the ZymoBIOMICS DNA Miniprep Kit (Cat. No. D4300). The resulting
.genomic DNA was screened with PCR for integration. Further PCR products produced with
primers outside of the flanking regions were verified with Sanger sequencing (Genewiz). Positive
cultures were plate purified, as described by (Lipscomb et al., 2011), on uracil-deficient, defined
media and screened again. For strains gene deletions, primers which would be present inside the
gene were utilized to check for its potential presence.

*P. furiosus* Growth Conditions

*P. furiosus* was grown anaerobically in serum bottles on defined media (5 g/L cellobiose,
maltose, or pyruvate) with added amino acids (1x) on media, as previously described (Lipscomb
et al., 2011), at temperatures ranging from 75°C to 95°C.
Fermentation Product Analysis

Cultures were allowed to cool to room temperature. A 5 mL sample was taken and centrifuged to pellet the cells. The remaining liquid was sterile filtered. One mL was transferred to a vial for gas chromatography analysis (Shimadzu GC-2014) to measure ethanol and acetoin. For HPLC, 1 mL samples were acidified with 1 μL 50% sulfuric acid. A Rezex-ROA column (300 mm x 7.8 mm; Phenomenex) was utilized for separation with a Water Model 2414 Refractive Index detector for quantitation.

Acknowledgements

CT Straub acknowledges support from a US Dept of Education GAANN Fellowship (P200A160061).
Figure 8-1. Central glycolytic pathway comparison  The canonical Embden-Meyerhof (E-M) central glycolysis produces a net two ATP along with the reduction of NAD+ to NADH when converting one glucose to pyruvate. The modified E-M pathway found in many thermophilic archaea reduces ferredoxin without a net gain in substrate level phosphorylation. [Red minus signs indicate net input of phosphorylated adenosine moiety (ADP → AMP for modified E-M pathway and ATP → ADP in classical E-M pathway) and green plus signs indicate net gain of ATP via substrate level phosphorylation (ADP + P_i → ATP).]
Figure 8-2. Energy Conservation in *P. furiosus* (A), AOR pathway to ethanol in *P. furiosus* (B) Insertion of an NADPH-dependent primary alcohol dehydrogenase drive ethanol production from acetate at the expense of reduced ferredoxin to reduce acetate to acetaldehyde and NADPH to reduce further to ethanol. AOR is native to *P. furiosus*, AdhA is heterologously expressed.
Figure 8-3. Potential futile cycling during gluconeogenesis (growth on pyruvate) when gapn is expressed at high levels During growth on pyruvate, P. furiosus catabolizes glucose via the ATP consuming phosphorylation of 3-phosphoglycerate (3-PG) followed by the NADPH-dependent reduction of 1,3-bisphosphoglycerate (1,3-BPG). The presence of Gapn could potentially re-oxidize glyceraldehyde-3-phosphate (GAP) back to 3-phosphoglycerate, creating an energy burning futile cycle.
Figure 8-4. Alcohol specificity of *C. subterraneus* alcohol dehydrogenase Enzyme assayed at 75°C in the direction of alcohol oxidation to aldehyde with NADP⁺.
Table 8-1. Relative activity of *C. subterraneus* AdhA in the oxidation and reductive reactions with co-factors

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Relative Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde + NADPH → Ethanol + NADP⁺</td>
<td>100%</td>
</tr>
<tr>
<td>Acetaldehyde + NADH → Ethanol + NAD⁺</td>
<td>14%</td>
</tr>
<tr>
<td>Ethanol + NADP⁺ → Acetaldehyde + NADPH</td>
<td>20%</td>
</tr>
<tr>
<td>Ethanol + NAD⁺ → Acetaldehyde + NADH</td>
<td>2%</td>
</tr>
</tbody>
</table>
Table 8-2. Strains utilized in this study

<table>
<thead>
<tr>
<th></th>
<th>gapn</th>
<th>gapor</th>
<th>AdhA</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. furiosus</em> DSMZ 3638</td>
<td>Present</td>
<td>Present</td>
<td>-</td>
</tr>
<tr>
<td>COM1</td>
<td>*Natural Mutant</td>
<td>Present</td>
<td>-</td>
</tr>
<tr>
<td>RK300</td>
<td>*Natural Mutant</td>
<td>Δ</td>
<td>-</td>
</tr>
<tr>
<td>RK301</td>
<td>Expressed under P_{slp}</td>
<td>Δ</td>
<td>-</td>
</tr>
<tr>
<td>RK302</td>
<td>*Natural Mutant</td>
<td>Present</td>
<td>Expressed under P_{slp}</td>
</tr>
<tr>
<td>RK303</td>
<td>Expressed under P_{slp}</td>
<td>Δ</td>
<td>Expressed under P_{slp}</td>
</tr>
</tbody>
</table>
References


