PRODUCTION OF BUTYRIC ACID BY THE CELLULOLYTIC ACTINOBACTERIUM *THERMOBIFIDA FUSCA*

by

KYLE MERKLEIN

B.S., Kansas State University, 2011

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biological and Agricultural Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Dr. Mei He

Abstract

Thermobifida fusca, an aerobic moderately thermophilic, filamentous soil bacterium is capable of producing butyric acid. Butyric acid is a 4-carbon short chain fatty acid that is widely used in the chemical, food, and pharmaceutical industries. Currently, butyric acid is primarily produced through petroleum-based chemical synthesis, but could be a candidate to be produced by fermentation. By producing through a fermentation platform, production of butyric acid can be shifted from a non-renewable to a renewable source. In an effort to make *T. fusca* produce a high yield of butyric acid, multiple fermentation parameters were explored and optimized. The effect of different carbon sources (mannose, xylose, lactose, cellobiose, glucose, sucrose, and acetates) on butyric acid production was studied, where cellobiose produced the highest yield of 0.67 g/g C (g-butyric acid/g-carbon input). The best stir speed and aeration rate for butyric acid production were found to be 400 rpm and 2 vvm in a 5-L fermentor. The maximum titer of 2.1 g/L butyric acid was achieved on 9.66 g/L cellulose. Fermentation was performed on ground corn stover as a substrate to evaluate the production of butyric acid on lignocellulosic biomass, and the optimized conditions resulted in a titer of 2.37 g/L butyric acid. The butyric acid synthesis pathway was identified involving five genes that catalyzed reactions from acetyl-CoA to butanyol-CoA in *T. fusca*. A study into the transcriptomics of *T. fusca* was begun by growing *T. fusca* under a variety of fermentation conditions, isolating the messenger RNA, and performing a sequence of the mRNA using whole

transcriptome shotgun sequencing. The results of sequencing of various samples were plotted to determine correlation across numerous fermentation parameters. This correlation based analysis determined that the carbon to nitrogen ratio has the largest overall impact on gene transcription of *T. fusca* among all of the fermentation parameters studied. Overall, the work from this study proves that production of butyric acid is possible from a renewable cellulosic feedstock.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would not be here without the help of many people. Thank you to the U.S. China Center for allowing the use of some lab equipment, such as the 5-L fermentor. Thank you to my graduate committee members, Dr. Lisa Wilken and Dr. Donghai Wang for taking the time to assist me in my program. Also, thank you to Dr. Wang for the use of the HPLC and Ke Zhang for setting up the samples for the HPLC. Finally, thank you to Dr. Yu Deng and Dr. Mei He – you both were great professors and advisors, and I am a better person because of your instruction.

Dedication

I dedicate this to my family, who makes me who I am. Thank you to my parents for always teaching me to do my best in everything that I do, because without that attitude I would not be here today. And finally, thank you to my wife, Dr. Taylor Merklein. You supported me going back to school, helped me through all the stressful times, and were an amazing supporter. I am so lucky to have you by my side.

Chapter 1 - Introduction to Biomass Utilization

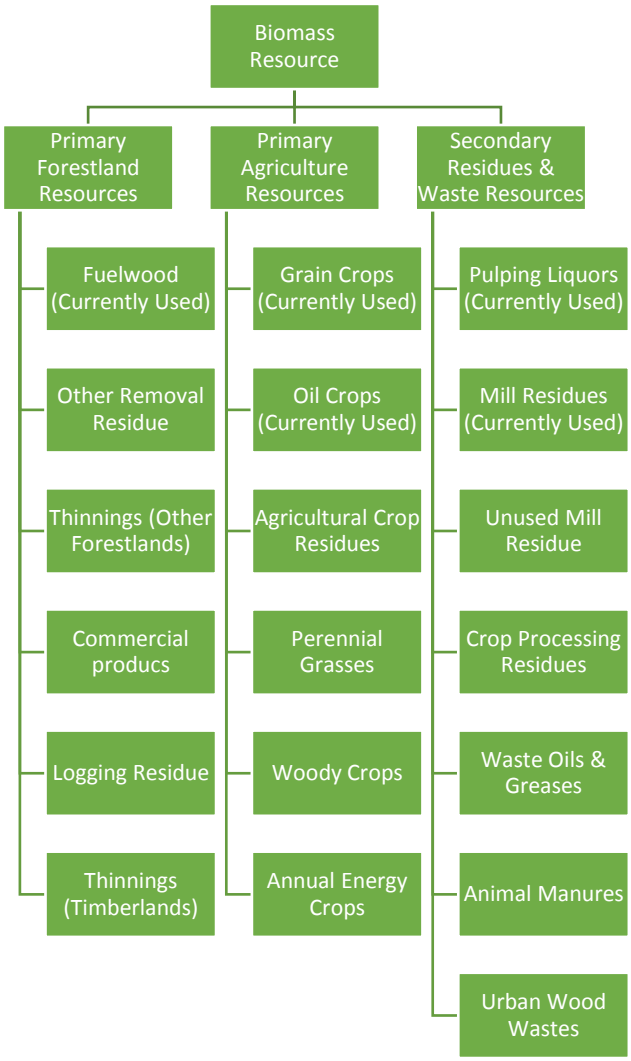
Much interest recently has been in the utilization of biomass substrates for biofuel production. This first chapter introduces the current research field derived from this interest, and it is primarily in a focus of production of bioethanol. The following chapters explore the significance of this research and optimization of fermentation conditions in an application of producing bio-based chemicals.

Introduction to Biomass

The term biomass refers to the general organic matter created by plants during the photosynthesis cycle, which converts and stores the energy from sunlight into chemical bonds. Biomass serves as a potential energy source for human use because there is energy stored within the chemical bonds that compose biomass. Humans can convert this potential energy stored in the bonds into a more useful form of energy. Utilization of this potential energy source is drawing great interest, due to its sustainability as a renewable energy. The simplest method, and historically most common, is to use this energy through a simple combustion reaction. Much of the current research focus is converting biomass into a form with a higher energy value for use as a substitute for petroleum fuel.

Numerous forms of biomass utilization have been investigated. First-generation biofuels, such as fatty-acid based methyl ester and corn based ethanol, have seen success on the industrial scale. However, these first-generation biofuels are generally produced from food sources, which have sparked extensive debates about the use of a food source as a fuel. These first-generation biofuels also raise questions on their impact to biodiversity and land use. Finally, another common claim against these biofuels is that they are not maximizing the abatement of greenhouse gas emissions [1].

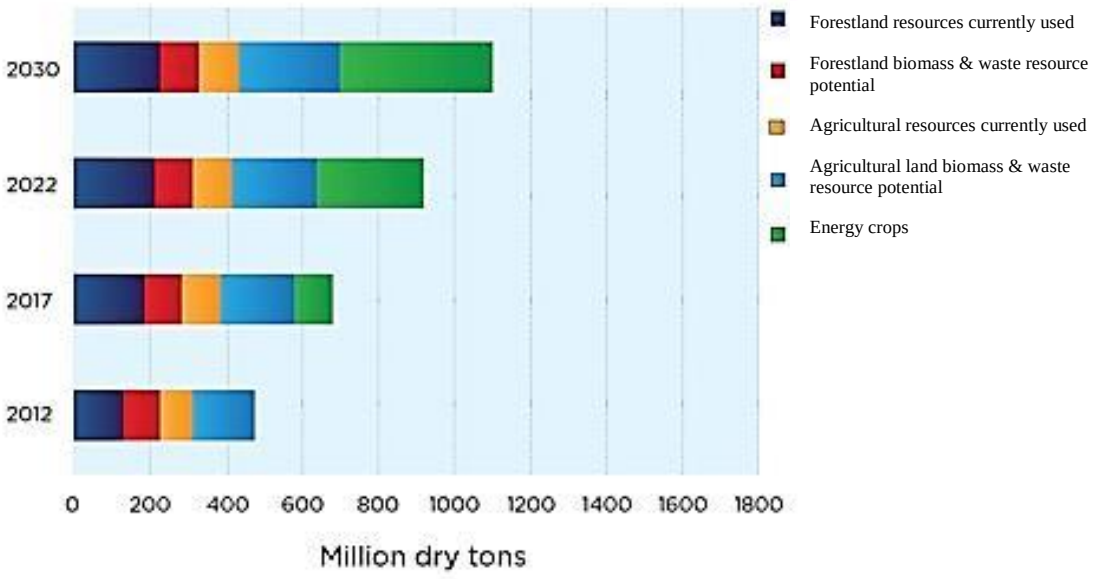
Figure 1-1.1 An overview of the biomass sources both currently used and those that have potential bioprocessing use. Adapted from [2].



Because of the reasons identified above, this chapter will focus on deriving fuels from lignocellulosic biomass (LCB). Fuels derived from LCB are known as second-generation biofuels [1], but these fuels are currently not produced at the industrial scale because they are not cost competitive with first-generation biofuels. As seen in the figure above, there are many potential sources including the following major sources: forest residues such as woody plant species, energy crops bred specifically to produce bioenergy, agricultural residues such as corn stover, and cellulosic waste. Producing second-generation biofuels could alleviate many of the

problems raised with first-generation fuels. As can be seen in the figure above, none of these biomass sources are food sources, which is the primary concern of first-generation biofuels. Sustainability is also promoted, as several of the proposed biomass sources are currently waste products. An example of the amount of biomass currently available, as well as projections through 2030, is illustrated in the figure below.

Figure 1-2 Amounts of LCB based feedstocks available in the United States. These projections take many factors into account, including conservative growth estimates for all feedstock types as well as sustainability of harvesting each feedstock. Used with permission from [2].



Structure of Lignocellulosic Biomass

Plant cells have a cell wall (0.1 to 10 μm) that provides rigidity to the cell, as well as preventing attack by pathogens [3]. The cell walls are typically composed of three layers – the middle lamella, the primary cell wall, and the secondary cell wall. Further sublayers of the secondary cell wall are present only in certain tissues [4]. Plant species are classified into either monocots or dicots. Corn stover or switchgrass (or other herbaceous biomass sources) are examples of monocots, and poplar (or other woody biomass) is a dicot. The chemical

composition of the cell wall is different between the two, and that is an important point to consider when processing the two types of plants. Major components of the cell wall in either species are cellulose, hemicellulose and lignin. Minor components include proteins, lipids, soluble sugars, and minerals [5]. Ethanol producing microorganisms cannot use these primary components in LCB for fermentation. Instead, these components would need to be broken down into fermentable sugars such as glucose.

As mentioned above, LCB is composed of the primary components of cellulose, hemicellulose and lignin. In dicots, these primary components are bound together in a very tight structure. Monocots have the same primary components but they are more loosely bound than dicots. The binding level of the primary components is associated with the lignin content of each LCB source [6]. Standard compositions of various species of biomass are exhibited in Table 1 below.

Table 1-1 Example of the variability of feedstock composition based on different LCB feedstocks [7].

Biomass Source	Variety	Cellulose (Percent mass)	Hemicellulose (Percent mass)	Total Lignin (Percent mass)	Total Ash
Hybrid Poplar	DN-34	39.23	16.66	25.18	2.03
American Sycamore	Platanus occidentalis	38.6	17.78	24.1	1.24
Sugarcane Bagasse	Gramineae Saccharum var. 65-7052	39.01	24.91	23.09	3.66
Corn Stover	Zea mays	34.61	22.21	17.69	10.24
Switchgrass	Alamo	30.97	24.39	17.56	5.76
Wheat Straw (Triticum aestivum)	Thunderbird	32.64	22.63	16.85	10.22

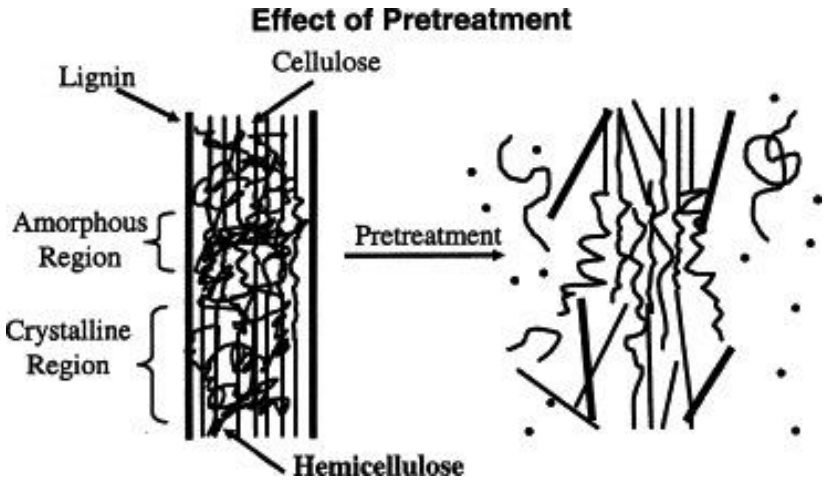
Each of the primary components of LCB is a very complex structure. Cellulose consists of linear chains of (1,4)-D-glucopyranose units, and those units are linked 1-4 in the β-configuration. The average molecular weight of cellulose is around 100,000 atomic mass units (amu) [6]. Hemicellulose is a mixture of polysaccharides, composed of sugars such as glucose, mannose, xylose, and arabinose, as well as methyglucoronic and galaturonic acids. Hemicellulose is a heterogeneous branched polysaccharide that binds tightly to the surface of each cellulose microfibril. Hemicellulose binding to cellulose is non-covalent bonding, and the average molecular weight of hemicellulose is <30,000 amu [6].

Lignin is a high molecular weight group of amorphous compounds. Lignin is composed of three-carbon chain attached to rings of phenyl-propanes, which is a six carbon ring. There are three possible structures of lignin, termed I, II, and III which correspond to zero, one, or two methoxy groups attached to the rings. The structures vary based on the source of the polymer; structure I is found in herbaceous plants, structure II is in the wood of conifers, and structure III is found in deciduous wood [6]. Biomass is not only composed of varying amounts of cellulose, hemicellulose, and lignin, but also is composed of different types of lignin. Therefore, any attempt to utilize biomass in a processing environment must be robust enough to accommodate the differing structures of biomass.

Pretreatment Technologies for Lignocellulosic Biomass

The first goal in pretreating lignocellulosic biomass is to disrupt the recalcitrant structure of the biomass. By disrupting the structure, enzyme access will be improved for downstream processing steps. The figure provides a simplified view of the pretreatment process.

Figure 1-3 The ultimate goal of pretreatment methods is to break LCB down from the recalcitrant structure to its individual components of lignin, cellulose, and hemicellulose. Used with permission from [8].



Because there are many methods of downstream processing, there are also many methods of pretreating LCB material. For instance, enzymatic hydrolysis may require the removal of hemicellulose since no hemicellulases are typically present, while consolidated bioprocessing utilizes organisms that naturally secrete hemicellulases and can effectively break hemicellulose down to its monomeric sugars. To prepare for the downstream processing, it is impossible to identify an industrial "gold standard" for pretreatment. Rather, different pretreatment strategies will be used, depending on what the needs are for the approach to biomass utilization. The overall strategies can be grouped as either a physical pretreatment, chemical pretreatment, and physico-chemical pretreatment.

Physical Pretreatment

Physical pretreatment of LCB consists of physically altering the structure of cellulose through one of two primary methods: the mechanical method that consists of chipping, grinding, and milling; and the extrusion method [9].

Mechanical Method

The goal of mechanical pretreatment is to reduce the crystallinity and particle size of the lignocellulosic biomass so that the surface area is increased and the degree of polymerization is reduced [9]. The technologies available for mechanical pretreatment are very mature, and are already used in a variety of bioprocessing areas, such as corn milling for grain ethanol production. It is well known that a combination of chipping, grinding and milling can be used to obtain a variety of final particle sizes. After chipping, the size of particle is typically in the range of 10-30 mm. After milling or grinding, the particle size is in the range of 0.2-2 mm [10]. Many different types of milling exist and are suitable for different applications. For wet material, the colloid mill, fibrillator and dissolver are suitable. For dry material, the roller mill, cryogenic mill and hammer mill are used [11]. The ball mill can be used for either dry or wet material.

However, despite being economically feasible for some industries, it is unlikely that this technology will be used for LCB pretreatment, as it takes a large amount of energy to get LCB into a small enough particle size for enzymatic hydrolysis. The energy consumption for mechanical size reduction has been calculated in most cases (depending on final particle size and size reduction ratio) to be higher than the theoretical energy content available in the biomass [12]. It should be noted that use of milling as the sole pretreatment strategy is not economically feasible; it is used to reduce particle size prior to other pretreatment methods.

Extrusion Method

The extrusion process of physical pretreatment is a means to subject LCB to a variety of heating, mixing, and shearing forces during passage through the extruder [9]. All of these forces produce desirable physical and chemical treatment methods for enzymatic hydrolysis. Extrusion processing was first evaluated as a pretreatment method in the 1980s for crop residues, sawdust and municipal waste [13]. It is believed that the screw speed and the barrel temperature are able

to disrupt the LCB structure, which causes defibrillation, fibrillation and shortening of fibers, all of which allows higher accessibility of enzymatic depolymerization [9]. Studies recently have focused on the many variables that can be altered in extrusion such as single or twin extrusion, barrel temperature, screw speed, grinding for extrusion pretreatment, additional of chemicals, as well as solids loading [13, 14].

Chemical Pretreatment

Chemical pretreatment methods work by using additional chemicals to break down biomass. All have been proven to be effective at improving the enzymatic hydrolysis of LCB in a laboratory scale.

Alkali pretreatment

Some bases can effectively solubilize lignin, as well as have minor cellulose and hemicellulose solubilization. It is believed that this occurs by breaking the ester bond between lignin, hemicellulose, and cellulose [9, 15]. This can cause less sugar degradation than acid pretreatment. Additionally, alkali pretreatment is believed to be more effective on agricultural residues than on wood materials [16].

Alkali pretreatment occurs at room temperature, and the duration of pretreatment can range from just seconds up to days. Ammonium, calcium, potassium and sodium hydroxides are effective bases for alkali pretreatment. Sodium hydroxide (NaOH) is effective as a pretreatment strategy because it causes swelling which increases the internal surface of cellulose as well as decreasing the polymerization and crystallinity [17]. This further causes lignin structure disruption. NaOH has been used to increase hardwood digestibility from 14% to 55% by reducing the lignin content from 24-25% to 20% [16]. Alkali treatment with NaOH has been compared to other pretreatment strategies. In comparing NaOH pretreatment of cotton stalks

with sulfuric acid, hydrogen peroxide, and ozone pretreatment, the NaOH pretreated stalks resulted in the highest cellulose conversion and highest level of delignification[18].

Lime, $Ca(OH)_2$, is a well-studied hydroxide for alkali pretreatment. Compared to other hydroxides, lime has a lower cost as well as fewer safety concerns. It also can be easily recovered from hydrolysates through a reaction with CO_2 [8]. Lime pretreatment has been effective at reducing the lignin fraction in LCB and also removes the acetyl group from hemicellulose [8, 19]. Lime pretreatment of corn stover [19] and poplar wood [20] at a temperature range from 85-150°C for 3-13 h has been successfully demonstrated.

Acid Pretreatment

The primary objective through acid pretreatment is to solubilize the hemicellulosic portion of the LCB. Treatment can occur either at low acid concentration and high temperature (dilute acid pretreatment) or high acid concentration and low temperatures (concentrated acid pretreatment) [9]. While both dilute and concentrated acids could be used for pretreatment, using concentrated acids is less desirable because it produces more inhibiting compounds, such as HMF and furfural, than diluted acids. It also can cause greater corrosion and acid recovery problems compared to diluted acid, which increase the cost on a commercial scale. Diluted acid pretreatments have been studied for a variety of reactor types. The pretreatment can occur at high temperature (180°C) for a short period of time, or a lower temperature (120°C) for a longer time period.

The most studied acid for pretreatment is diluted H_2SO_4, which may in fact be the most common chemical pretreatment method [11]. Experiments have shown the ability to remove nearly 100% of the hemicellulose with dilute sulfuric acid although it is not effective at dissolving lignin. It does disrupt lignin as well as increase the cellulose susceptibility for

enzymatic hydrolysis [21, 22]. The conditions that are optimal for hemicellulose removal may not be optimal for the pretreatment for enzymatic hydrolysis. Studies testing various parameters of the pretreatment of olive tree biomass showed different process conditions for highest overall sugar yield, highest hemicellulose removal, and highest enzymatic hydrolysis yield [23]. The major drawback of this method is formation of inhibitors. Compounds such as carboxylic acid, furfural, and phenolic compounds were formed, which may affect enzymatic hydrolysis and have a negative impact on microbial growth [17, 24].

Additional acids used for dilute acid pretreatment are the organic acids, such as fumaric or maleic acids. Initial studies have shown that less furfural was formed when using organic acids when compared to pretreatment using sulfuric acid [25]. However, as fewer studies have been performed with organic acids, it is unknown if they can achieve the same results as dilute sulfuric acid.

Ionic Liquids

Ionic liquids (ILs) used as solvents are a relatively new pretreatment method that is receiving a large amount of research. These solvents are salts composed of large organic cations and small inorganic anions, and are typically defined as molten salts with melting points less than 100^{o}C. They are chemically and thermally stable, non-flammable, relatively non-toxic, and have low vapor pressures [9].

The pretreatment using ILs is carried out by the dissolution of biomass in the solvent at temperatures between 90-130^{o}C, ambient pressures, and varying amounts of time (typically 1-24hr). Water is then added to precipitate the biomass, and is washed prior to enzymatic hydrolysis [28, 29]. The IL mechanism is the formation of a hydrogen bond between the anion of the IL and the cellulose, which breaks up the cellulose crystalline structure [30]. Additionally,

the chemistry of certain ILs lends to the ability of selectively removing lignin and hemicellulose from biomass [31-33].

The research into ionic liquids as a pretreatment strategy was first reported in 2002[4, 34], so there are still numerous problems to solve before this is an industrially viable pretreatment strategy. Specifically, a means to efficiently recycle the IL is a prerequisite, as is the toxicity of ionic liquids to enzymes and fermentation microorganisms. Initial studies have shown that ionic liquids can have a significant negative impact on cellulase activity [9]. Additionally, ILs are currently too expensive for industrial use because they are derived from multiple synthesis of nonrenewable resources, so a means to produce ionic liquids at a much lower cost is required for widespread use as a pretreatment technology[35].

A strategy that could accomplish the production of ILs at a lower cost is the synthesis of ionic liquids from carbohydrates. One promising approach in the formation of ILs from renewable resources is to use the lignin and hemicellulose fractions of biomass as a starting point to derive ILs [35]. This is described as a hypothetical closed-loop biorefinery where the cellulose portion of the LCB would be used for fermentation, while the lignin and hemicellulose portions would be depolymerized and converted to ILs, which would be used by the refinery as the pretreatment method. Compared to the "gold standard" IL, which could remove 44.5% xylan and 52.4% lignin, the biomass derived ILs had the ability to remove xylan (33.9-51.4%) and lignin (3.9-43.0%) through pretreatment at rates close enough to warrant further research. Additionally, the initial techno-economic analysis showed that the raw materials cost of the biomass derived ILs is in the range of 60-70% the cost of the "gold standard" used in the study, with potential to be much cheaper.

Physico-chemical Pretreatment

As the name implies, physico-chemical pretreatment methods combine the advantages of physical disruption to cellulose as well as chemical processing. Similar to the purely physical or purely chemical methods, all have been proven to be effective at improving the enzymatic hydrolysis of LCB.

Steam Explosion

Steam explosion is a treatment method where the biomass is subjected to pressurized steam for a time period of several seconds to several minutes, followed by sudden depressurization. The mechanism of pretreatment involves both mechanical and chemical effects. The chemical effects occur through a hydrolysis (autohydrolysis) of acetyl groups that are present in the hemicellulose [9]. This autohydrolysis takes place at high temperatures, which promote the formation of acetic acid from acetyl groups; additionally, the water from steam can act as an acid at high temperatures [9]. Mechanically, the reduced pressure causes the fibers of biomass to separate as a result of explosive decompression [9].

Multiple factors can affect the effectiveness of steam explosion, including particle size, temperature (T), residence time (t), as well as the severity factor. The severity factor (R_0) is characterized by the equation '$\log R_0 = \log(t*e^{((T-100)/14.75)})$' where "t" is minutes and "T" is degrees Celsius[36]. It should be noted, that the severity factor R_0 does not take into account any acids used as a catalyst in the process. Combining dilute acids with steam explosion have been shown to effectively pretreat softwoods, improve hemicellulose hydrolysis during the pretreatment, as well as further decrease retention time and temperature of the steam explosion. Therefore, a lower severity factor will be required for the process when using dilute acids [30]. A decrease in the retention time and temperature has been shown to reduce inhibitory

compounds formed, which helps improve the later enzymatic hydrolysis and fermentation steps [37]. With all treatment strategies, addition of diluted acid comes with potential downfalls such as added processing cost, added steps of neutralizing the acid and removal of salts, as well as additional concerns when it comes to equipment requirements [30].

Liquid Hot Water

Liquid hot water pretreatment is carried out with only water and no catalysts or additional chemicals. The water is both heated to elevated temperatures (160^{o}C-240^{o}C) while applying a high enough pressure to keep the water in a liquid state. This process causes alterations in the structure of lignocellulose, primarily through the removal of hemicellulose. After pretreatment, the slurry formed can be filtered into a solid cellulose enriched fraction and a hemicellulose rich liquid fraction [9]. This process is a hydrothermal pretreatment method that has been used for several decades as part of the pulp industry [11].

The liquid fraction consists primarily of oligosaccharides derived from hemicellulose, lignin (35-60% of starting material), with a minor amount of cellulose (4-15% of starting material) [30]. To prevent inhibitory compounds from effecting enzymatic hydrolysis and fermentation, it is desired to solubilize as much lignin as possible. It has been shown that the amount of lignin solubilized is related to the lower range of the process temperatures for liquid hot water treatment [38].

Optimization of the liquid hot water method for pretreating corn stover found that conditions of 190 ^{o}C for 15 minutes resulted in a 90% conversion of the cellulose in the subsequent enzymatic hydrolysis step [39]. Another study compared the performance of liquid hot water and steam pretreatment of sugarcane bagasse [40]. Both methods were shown to improve enzymatic hydrolysis, but the liquid hot water method had a better xylan recovery. It

was also found that the liquid hot water was comparable to dilute acid pretreatment in terms of conversion.

Ammonia Fiber Explosion (AFEX)

The AFEX pretreatment process involves biomass being treated with liquid anhydrous ammonia at high pressures and a temperature of 60-100^{o}C. After a period of time, ranging from a low of 5 minutes to a moderate time of 30 minutes, the pressure is released which causes a rapid expansion of the ammonia gas. This rapid expansion causes swelling and physical disruption of the LCB, as well as partial decrystallization of cellulose. The ammonia loading for AFEX is typically a 1:1 ratio of ammonia to dry weight biomass [41]. The AFEX process differs from other processes like steam explosion as it only produces a pretreated solid material as opposed to separate liquid and solid fractions.

AFEX results in a decrease in cellulose crystallinity as well as the disruption in lignin-carbohydrate linkages [42]. Because both cellulose and hemicellulose remains in the pretreated material, both cellulases and hemicellulases will be required during the enzymatic hydrolysis step. Since the pretreated material contains full amounts of lignin, it has been shown that AFEX is more effective on agricultural residues and herbaceous crops as compared to high lignin feedstocks like woody biomass [21]. An advantage of AFEX pretreatment is the lack of inhibitor formation, which as has been previously discussed, helps the steps of enzymatic hydrolysis and fermentation more efficient.

Enzymatic Hydrolysis

After pretreatment, the biomass is broken down into monomeric sugars through enzymatic hydrolysis. These monomeric sugars can then be used through standard fermentation pathways. Therefore, three individual steps must occur to convert the pretreated biomass into the

final product – enzymatic hydrolysis, fermentation of 5-carbon sugar, and fermentation of 6-carbon sugar. The order of these three individual steps is separated into different schemes of enzymatic hydrolysis.

The first scheme that exists is called Separate Saccharification and Fermentation (SSF). This scheme separates enzymatic hydrolysis, fermentation of 5-carbon sugars, and fermentation of 6-carbon sugars into separate reaction vessels. The second scheme is called Simultaneous Hydrolysis and Fermentation (SHF). This scheme combines the process of enzymatic hydrolysis and fermentation of 6 carbon sugars in a single vessel, while fermentation of 5 carbon sugars occurs in a separate vessel. The final scheme available is Simultaneous Saccharification and Co-Fermentation (SSCF). The SSCF process combines enzymatic hydrolysis, fermentation of 5-carbon sugars, and fermentation of 6-carbon sugars into one single vessel.

One common element for all three schemes is that they all require production of cellulases in a separate system. The cellulases produced for enzymatic hydrolysis is a primary factor in how much sugar is available for fermentation. Therefore, the focus of the remainder of this section will be on the factors affecting enzymatic hydrolysis, as well as an overview of the enzymes used to break down the recalcitrant structure of biomass.

Factors effecting Enzymatic Hydrolysis

As discussed earlier in the chapter, the effectiveness of enzymatic hydrolysis depends on breaking down the recalcitrant structure of LCB materials into a more manageable substrate for enzymatic hydrolysis. Therefore, pretreatment of LCB material is necessary prior to performing any hydrolysis or fermentation steps. Specifically, the areas pretreatment should address are the following [9]:

- Cellulose crystallinity

- Cellulose degree of polymerization
- Substrate's available surface area
- Lignin barrier
- Hemicellulose content
- Feedstock particle size
- Porosity
- Cell wall thickness
- Change in accessibility with conversion

Enzymes used for Hydrolysis of Lignocellulosic Biomass

Once the issues structure of LCB has been pre-treated it is conducive to enzymatic hydrolysis. Hydrolysis of LCB is accomplished through the use of at least the following different types of enzymes [48]:

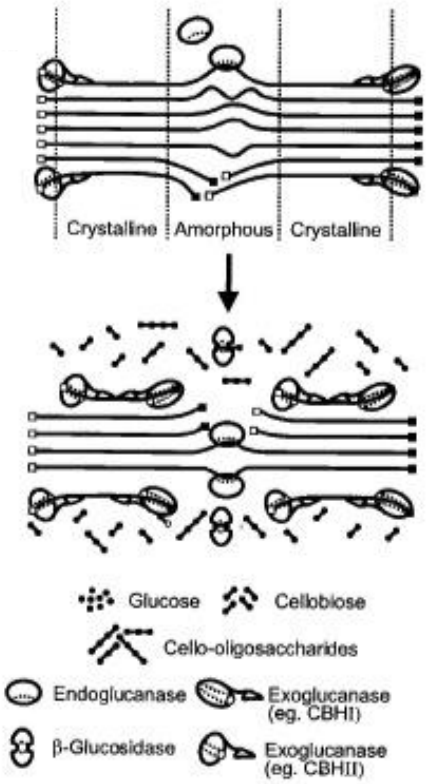
- Endoglucanase (EG) or 1,4-β-D-glucan-4-glucanohydrolase
- Exoglucanase, including 1,4-β-D-glucan glucanohydrolase (also known as cellodextrinase) and 1,4-β-D-glucan cellobiohydrolase (cellobiohydrolase; CBH)
- β-glucosidase (BGL) or β-glucoside glucohydrolase

The process of degrading cellulose begins with the EG hydrolyzing the β-glycoside linkages of the internal regions of cellulose, which yields oligosaccharides of varying degrees of polymerization. Additionally, this hydrolysis step generates new chain ends which are the access point for other enzymes. Exoglucanases hydrolyze in a processive manner, starting at the reducing or non-reducing ends of cellulose to yield either glucose or cellobiose. Additionally, exoglucanases can hydrolyze microcrystalline cellulose by peeling the cellulose chains from the

microcrystalline structure. Finally, BGL cleaves the soluble cellodextrins and cellobiose into glucose [48].

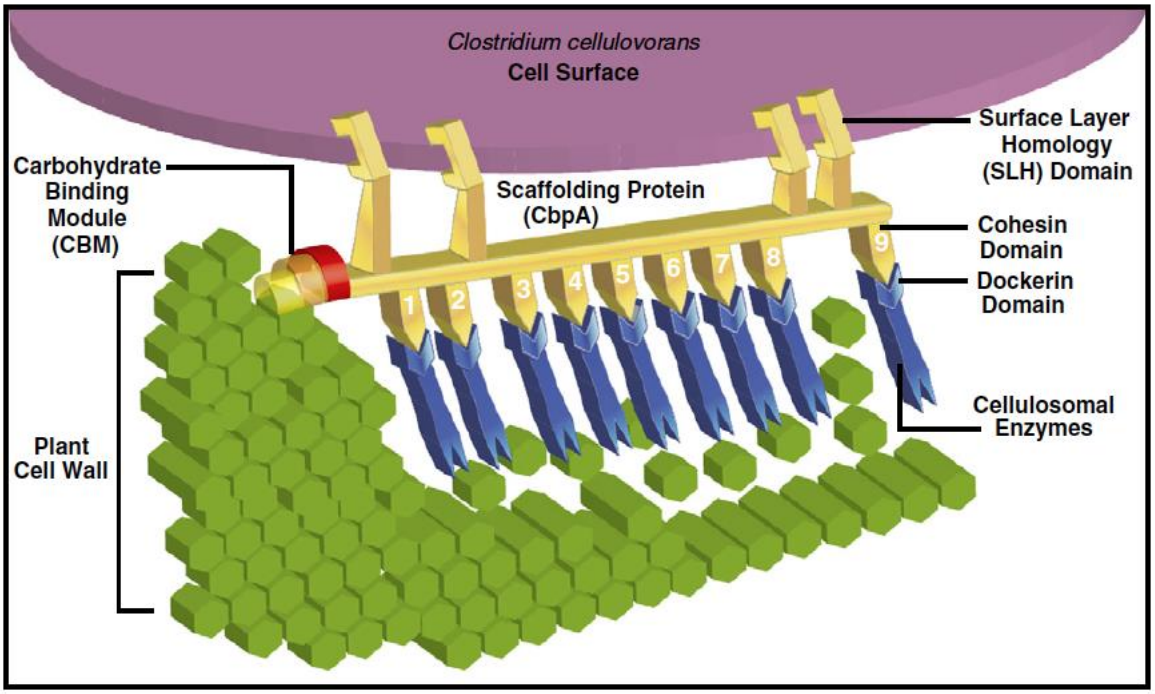
Fungi belonging to the genus Trichoderma, specifically *Trichoderma reesei*, are the primary microorganisms for commercial production of these cellulases since this one organism produces every required enzyme. *T. Reesei* is known to produce at least two exoglucanases (CBHI and CBHII), five endoglucanases (EGI, EGII, EGIII, EGIV, and EGV), and two β-glucosidases (BGLI and BGLII). The cellulolytic system of *T. Reesei* is a typical non-complexed cellulase system where cellulolytic enzymes are secreted and act in an uncoordinated, diffusion-driven manner [48].

Figure 1-4 Example of a non-complexed cellulase system where enzymes are secreted and act in an uncoordinated, diffusion driven manner. Used with permission from [49]



A complexed cellulase system, or cellulosome, is a multi-enzymatic complex that is found to be produced by some anaerobic bacteria. These multi-enzymatic complexes can efficiently degrade plant biomass, by assembling different cellulose degrading enzymes on a structural scaffoldin subunit. These enzymes are connected to the scaffoldin subunit through strong non-covalent protein-protein interactions between the docking modules (dockerin) and complementary modules (cohesins). Additionally, scaffoldin contains a carbohydrate-binding module, which allows the entire complex to bind to the cellulose surface [48]. Because complexed cellulase systems are attached to a specific organism through the structural scaffoldin subunit, they are not commercially produced for enzymatic hydrolysis.

Figure 1-5 Example of a complexed cellulase system, or cellulosome with the necessary parts required to bind to a plant cell wall for biomass degradation. Used with permission from [50]



Pretreatment strategies have the goal of solubilizing hemicellulose from the biomass, thus only leaving cellulose for hydrolysis. However, as the previous review of the pretreatment strategies indicated, hemicellulose typically is present in the pretreated biomass so it is important to identify the enzymes that can efficiently hydrolyze hemicellulose. In fact, because of the tangled nature of biomass, hydrolyzing the hemicelluloses can expose and liberate cellulose from the tangled biomass and further increase yields. The common hemicelluloses of β-glucan, xylan, xyloglucan, arabinoxylan, mannan, galactomannan, arabinan, galactan, polygalacturonan have hemicellulases that target them called β-glucanase, xylanase, xyloglucanase, arabinase, mannanase, galactanase, polygalacturonase, glucuronidase, and acetyl xylan esterase [51, 52]. For enzymes that work specifically on the structure of hemicellulose, glycoside hydrolases hydrolyze glycosidic bonds, carbohydrate esterases hydrolyze ester bonds, polysaccharide lyases

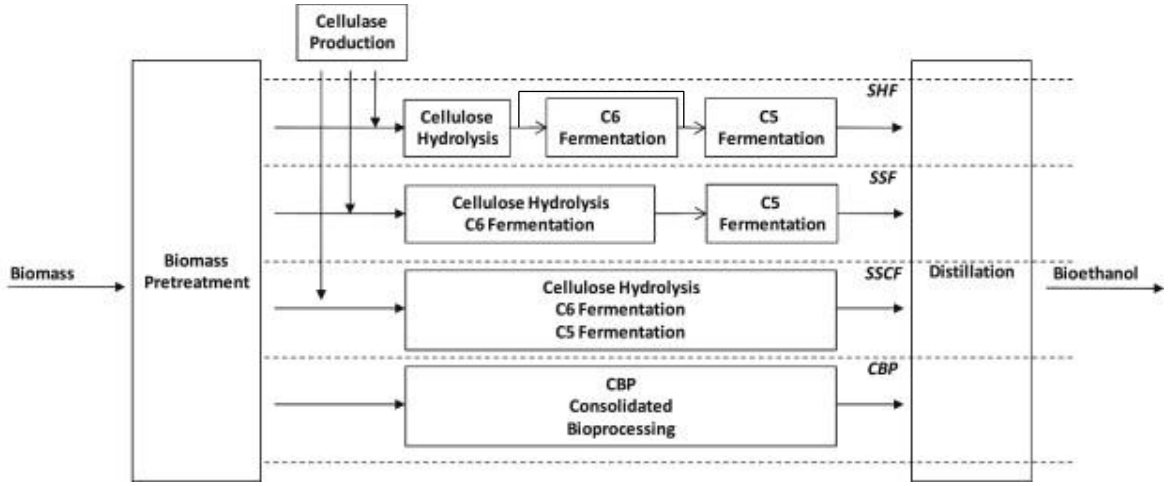
cleave glycosidic bonds. The enzyme *endo*-hemicellulase cleaves internal/backbone glycosidic bonds, and other glycosidases mainly remove substituent groups or side chains [53]. Different hemicellulase combinations are required for different feedstocks, because the composition of hemicellulose varies widely across plant type. Taking this a step further, synergisms have been found amongst hemicellulases and between hemicellulases and cellulases [54-59]. This suggests that complex data processing and bioinformatics tools must be applied to correctly apply these synergisms to an industrial process. Detailed mechanistic models have been created for enzymatic hydrolysis of cellulose [60-63]. Some mechanistic models have been created, although less advanced than the cellulose modeling, to understand the enzymatic hydrolysis of hemicelluloses [64-67]. However, only one such study to date was found for a mechanistic model that considers simultaneous enzymatic hydrolysis of cellulose and hemicellulose [68]. Developing the synergism of cellulases and hemicellulases could be extremely valuable in enzymatic hydrolysis, or through consolidated bioprocessing (CBP).

Consolidated Bioprocessing

The previous sections have described the processes involved in biomass processing: pretreatment of the biomass to help remove lignin and solubilize hemicellulose, saccharolytic enzymes (both cellulases and hemicellulases) and their use in the enzymatic hydrolysis process, and the fermentation of the sugars produced. While these processes all are highly connected, they exist as separate steps and require separate process parameters or vessels to maintain optimal efficiency. Combining the production of cellulases, the enzymatic hydrolysis process, and fermentation into a single step is consolidated bioprocessing (CBP) [49]. Because cellulolytic microorganisms can produce their own cellulases, the commercial production of cellulases is eliminated in CBP. Additionally, the cellulolytic bacteria can excrete their

cellulases and ferment sugars in a single step rather than in the multiple steps of enzymatic hydrolysis and fermentation. A diagram showing the differences between the schemes for biomass utilization is illustrated in Figure 1-6.

Figure 1-6 Overview of schemes to convert LCB into Biochemicals. The first three routes of SHF, SSF, and SSCF all require external production of cellulases, whereas CBP does not. Also, SSCF and CBP are the only schemes that can have hydrolysis, Fermentation of 6 carbon sugars, and Fermentation of 5 carbon sugars all occur in the same vessel. Used with permission from [69].



The primary advantage of CBP is the potential to both decrease the cost and raise the efficiency of biomass processing. Cellulase costs have remained relatively unchanged over the years, and still is a large cost in the biological conversion to biofuels [70]. As mentioned earlier, there is a potential synergy between cellulases and hemicellulsases. However, as the cellulase cost is already a large cost in the enzymatic hydrolysis step, it would not seem feasible to also add cost by using hemicellulases. Therefore, it is unlikely that this synergism will be utilized in standalone hydrolysis processes since it won't provide enough benefit to offset the added cost. Because CBP uses an organism with both cellulases and hemicellulases, this synergism could be leveraged with no additional cost.

Having enzymatic hydrolysis occur in one step through CBP eliminates the need for additional process equipment since CBP occurs in one step rather than two or three like other enzymatic hydrolysis/fermentation schemes. This reduction in steps lowers capital cost, both in fixed and variable operating costs in the industrial scale.

In addition to the cost reduction, there is a potential that CBP will increase the efficiency of the system. This can be realized through an enzyme-microbe synergy that has evolved over time in CBP organisms. Additional efficiency can also be seen by eliminating potential contaminants to the process, as non-cellulolytic organisms will compete with cellulose-adherent cellulolytic microorganisms during the hydrolysis step [71].

Cellulose Utilization in CBP

There are two mechanisms for crystalline cellulose utilization in cellulolytic organisms. This cellulose utilization is either directly adhering to cellulose or by producing non-complexed cellulases. Anaerobic bacteria typically follow the former mechanism, while aerobic bacteria and fungi tend to follow the later mechanism [49]. Some microbial species, however, exhibit different combinations of complexed and non-complexed cellulases [49].

After primary hydrolysis, which breaks cellulose down into soluble oligomers, the strictly anaerobic, cellulolytic bacteria utilize cellobiose and cellodextrin phosphorylases (CbP and CdP) to catalyze an ATP-independent phosphorolysis reaction. Further intracellular metabolism creates glucose-6-phosphate, which is the entry point in the Embden-Meyerhoff pathway. From the Embden-Meyerhoff pathway, there are a wide variety of end products that can be produced. The end product distribution appears to be a result of metabolic flux at several levels: mass action effects involving the concentrations of intermediates, fermentation products, and electron carriers; enzyme activity; and enzyme synthesis [49].

Substrate Pretreatment for CBP

All of the pretreatment methods reviewed earlier were initially designed for separate enzymatic hydrolysis and fermentation steps. However, since CBP still performs steps of enzymatic hydrolysis and fermentation, the pretreatment methods are still applicable for CBP.

Microorganism development for CBP should evaluate the strengths and weaknesses of a pretreatment method, and vice versa. Specifically, there are a number of points to analyze when comparing pretreatment/CBP compatibility [49].

The first is evaluating the effect pretreatment has on cellulose crystallinity. Because cellulose crystallinity is not affected by some pretreatment methods, so it is important for the cellulolytic enzyme system of the CBP microorganism to be capable of utilizing crystalline cellulose [49]. If it is not, then a large source of the potential sugar that will be unused in CBP.

Another area to evaluate is the ability of the CBP microorganism to hydrolyze soluble or insoluble hemicellulos. This is a minor consideration, as most processing methods offer a significant solubilization. However, strong alkali pretreatment methods may have low solubilization levels, which may pose challenges for microorganism utilization. It has been shown that residual insoluble hemicellulose present after pretreatment have been hydrolyzed by commercial cellulase preparations without any specific enzyme for hemicellulose hydrolysis [49]. However, the synergy of hemicellulases and cellulases has also been shown. Therefore it may or may not be necessary for CBP organisms to produce hemicellulose-specific enzymes; this should be examined further for each microorganism and each process case [49].

Another consideration is to understand the ability of the candidate microorganism to convert all sugars derived from hemicellulose. It is important for the microorganism to utilize both monomeric sugars, such as xylose, as well as soluble oligomers. Maximizing use of all potential sugars is a goal to increase productivity of CBP.

Finally, a microorganism should remain metabolically active in the presence of inhibitory compounds created during pretreatment. Because detoxification measures would add to overall cost of processing, it is desired to keep this to a minimum. The amount and types of inhibitors

produced depend on the process conditions and configurations. Degradation of lignin is not a goal in organism development, although decreasing the cellulose binding to lignin could be potentially valuable [49].

Organism Development for CBP

The ideal microbial culture for CBP combines the properties of both substrate utilization and product formation. The organism must have efficient substrate utilization, both in terms of the production of a hydrolytic enzyme system that allows high rates of hydrolysis, as well as using being able to use the hydrolysis products for growth. Product formation refers to the ability of the microorganism to selectively produce the desired end product. This is accomplished through high product selectivity through its native metabolism, or through the organism's ability to withstand high product concentration in the growth medium. Additionally, the CBP process must occur in an amount of time that is industrially feasible, eliminating slow growth microorganisms.

Developing an organism for CBP is separated into two distinct strategies. The first strategy is to identify a native cellulolytic organism and develop its ability to produce ethanol (or other fuel product). A second strategy is to find a natural ethanol-producing organism and develop a cellulase system for it to degrade cellulose. This strategy is called the recombinant cellulolytic strategy [49].

Native Cellulolytic Strategy

One possible strategy in CBP is to make use of naturally occurring cellulolytic microorganisms as a starting point for organism development. These cellulolytic organisms are isolated from nature in areas where cellulose must be used by microorganisms to survive, such as in the soil or in the rumens of animals. Because these organisms have developed the ability to

use cellulose and/or hemicellulose, they are a good starting point for strain development for CBP. Additional bioprospecting and characterization of cellulolytic organisms will lead to improvements in the CBP field and provide flexibility for targeting different compounds. The table in Appendix A shows a (non-inclusive) list of cellulolytic organisms that have been identified to date, relevant growth information, and the genome accession number if applicable [72].

Metabolic Engineering of Cellulolytic Organisms

To make CBP economical it is necessary to have an organism that is able to produce a high product yield. The most promising methods to develop a high yield organism is through metabolic engineering. While native cellulolytic bacteria typically produce large amounts of mixed organic end products, it is believed that these bacteria will be able to metabolize cellulose to a near exclusive end product[49]. This is supported by initial experiments with cellulolytic bacteria, as well as the large amounts of experience producing an exclusive end product through metabolic engineering of non-cellulolytic microorganisms.

A major challenge in the development of native cellulolytic organisms is the lack of universal gene-transfer and genetic engineering techniques. However, much work in recent years has gone into developing these techniques for some organisms, including *Thermobifida fusca*[73, 74] and *Clostridium thermocellum* [83, 84]. While these gene transfer techniques are generally more challenging than for non-cellulolytic bacteria, it does still provide the ability to further develop these cellulolytic bacteria. Using the knowledge gained from non-cellulolytic bacteria, metabolic flux can be altered by blocking undesirable pathways by generating gene deletions (knockout strains) or through overexpression of genes associated with desirable pathways. When performing initial attempts to redirect flux, it is common for unintended

consequences to occur. These are a result of either a metabolic imbalance of organic intermediates [75, 76], electron carriers [77], or interactions with metabolic control systems [78]. When engineering a microorganism for a certain function, like producing ethanol, it is important to fully evaluate the consequences of genetic modifications. Some of the changes may provide no value other than knowledge gained. Other engineering efforts may not initially improve product formation, but paired with an additional compounding change, may provide a significant improvement in product formation.

Because of the challenges in metabolic engineering, including lack of accuracy and a slow process, systems biology tools have been developed to rationally design microorganisms. Rational design strain optimization protocols have been developed, and they include many computational tools that help to model the cell metabolism and to predict strategies for metabolic engineering [79-82]. These models help process all available knowledge in a way that are intuitively difficult to analyze, and they take into account many factors including experimental data, all known biochemical pathways, and genomic information including enzyme functionality, stoichiometry, and much more possible information. By employing these models for metabolic engineering, the genetic manipulations that do occur have a much higher chance of being successful. Additionally, the advantage is that experimental results (even failed designs) can continue to improve modeling predictions to further increase the likelihood of successful metabolic engineering. As more tools are developed for synthetic biology, transcriptomics, metabolomics and systems biology, these models should become ever more robust.

Having computational models for rational design of microorganisms is only one challenge in metabolic engineering. Modeling an organism and identifying possible gene targets is useless if the organism under development is unable to be genetically modified. The genetic

manipulation platform for *Clostridium thermocellum*, a gram-positive, anaerobic, thermophilic bacterium, was noted in a preliminary report with protoplast transformation [83]. However, the first reported platform was reported through use of an electrotransformation (ET) protocol using a custom designed pulse generating system as well as custom designed cuvettes for electroporation [83]. The protocol for this ET was optimized. It was determined that chilling during cell collection and washing prior to electric-pulse application as well as the duration of the postelectropulsing recovery time had the largest effects. Additionally, the electric-pulse parameters had a large impact of transformation efficiency. However, as this was the first round of genetic manipulation via ET, it was noted by the authors that they expected that this method would become obsolete with a better understanding of the bacteria, as well as better technology. This proved to be the case as a new protocol for ET of *C. thermocellum* was developed through creation of a gene deletion system based on the *pyrF* gene [84]. This new platform reduced the need to have custom cuvettes, as well as simplifying the process to obtain competent cells. The custom pulse generator still worked best for this system, however, it was noted that commercial generators could obtain successfully modified cells.

The microbe *Thermobifida fusca* has also had a genetic manipulation platform developed to allow for future metabolic engineering efforts. *T. fusca* had been studied and characterized for it cellulolytic capabilities, and also had its genome sequenced. Initial genetic manipulation for this bacteria occurred through a non-rational design protocol. These efforts were made by continually altering the growth conditions of *T. fusca*, which induced stress, leading to natural evolution of the bacteria over time. Desired traits were imparted to the bacteria, but it was clear that this procedure would have little utility in future strain development. The efforts were made for rational design of the bacteria through chromosomal DNA manipulation. The high

exonuclease activity in *T. fusca* required atypically long flanking regions of homology in the disruption cassette. This disruption cassette was introduced as a plasmid, however since *T. fusca* does not possess a functional methyl-directed mismatch repair system, the plasmid preparation and propagation were conducted using a nonmethylating host strain. For transformation, *T. fusca* protoplasts (cell wall removed) were prepared. After transformation, because of the fragility of the protoplast, there was a period for protoplast regeneration in a non-selective medium. After regeneration, a selective media was used to screen for positive transformants. In the metabolic engineering work using the platform, *T. fusca* was manipulated to replace the celR gene with the adhE2 gene in the chromosome of *T. fusca* in an attempt to improve production of 1-propanol, which is a higher alcohol that could be compatible with existing infrastructure for ethanol. Attempts were successful as the transformed cells produced 1-propanol, although maximum titers were still low with a maximum titer of 0.48 g/L on raw switchgrass. However, having the platform for a thermophillic bacteria, combined with systems biology tools could lead to production of much higher yields of various chemicals.

The information above described the overall strategy as well as the information needed to fully understand metabolic engineering, which can be applied at a basic level for ethanol production. Based on the branched catabolism of a typical ethanol forming cellulolytic anaerobe in both the carbon and electron centered perspective, a few possible targets for gene knockout arise. Some specific knockout targets to redirect flux away from acetic acid formation are hydrogenase, acetate kinase and phosphotransacetylase. It has been noted that the basic requirements of fermentative metabolism, mainly the generation of net ATP and regeneration of reduced electron carriers, can be met by producing ethanol, acetic acid, or lactic acid either singly or in combination. All of this information is helpful to understand, but these are just

starting points for metabolic engineering efforts. The systems biology based model described previously should be employed for this purpose in order to achieve full optimization of metabolic engineering efforts.

Product Inhibition of Cellulolytic Bacteria

An additional consideration when selecting and developing a microorganism for CBP is the ability to maintain growth in high concentrations of product. While the microorganism should continue to grow in high concentrations of final product, in this case ethanol, it is also necessary for the organism to be able to withstand the accumulation of other products and inhibitory compounds.

It is believed that microorganisms are inhibited by ethanol through end product inhibition of glycolytic enzymes, as well as the damage to the cell membrane [85, 86]. Among cellulolytic organisms, *C. thermocellum* has been investigated extensively in terms of end product tolerance. In *C. thermocellum,* the inhibition is believed to be a blockage in glycolysis associated with ethanol-induced changes in the cell membrane [87-89]. Additionally, increased temperature has been observed to markedly decrease ethanol tolerance in *C. thermocellum* and other organisms [90, 91]. This occurs because both increased temperature and ethanol result in increased membrane fluidity.

C. thermocellum has been shown to grow at added ethanol concentrations up to 50 g/L [92, 93]. However *C. thermocellum* cannot natively produce ethanol at this high of concentration, and the maximum titer was about 25 g/L [92, 94]. This discrepancy is known as a titer gap, which is the difference between the maximum concentrations of a compound tolerated by a microorganism when added to the culture, versus the maximum concentration produced by the organism [70]. This titer gap has been decreased in non-cellulolytic microorganisms, such as

the production of ethanol with the industrial standard of *saccharomyces cerevisiae*, so it could be assumed that this will occur for cellulolytic bacteria as well. However, just like the challenges of metabolic engineering, this progress will initially be slow as the molecular platforms are less developed.

In additional to ethanol inhibition, various other compounds are likely to occur in an industrial setting that could cause an inhibition to growth. Similar to ethanol inhibition, the most studies to date have also occurred with *C. thermocellum.* In multiple studies, *C. thermocellum* was shown to have the ability to grow on pretreated LCB substrates that contained most or all of the lignin present prior to pretreatment, which suggests that soluble lignin does not have inhibitory effect [95-97]. However, two other common inhibitors have been shown to have a negative effect on *C. thermocellum.* The presence of hydrolysates that resulted from autoclaved corn stover [98] and steam-exploded aspen [99] were shown to inhibit growth. Additionally, the presence of acetate, which is a by-product of pretreatment as well as a fermentation product were shown to decrease the growth of un-acclimated cells by 50% in a near-neutral pH medium with an acetate concentration of 0.28 M [100]. The proposed theory of acetate inhibition is the magnitude of the transmembrane proton motive force is decreased. It is also suggested that anion accumulation is the explanation for the toxicity of cells to organic acids[101].

Similar to strain development, the reasons for product inhibition are extremely complex, and the development of bioinformatics and systems biology tools could be useful. Transcriptomics and proteomics can produce an extensive amount of data that could provide information for strain development and genetic engineering. The result of these strain development efforts could improve product tolerance. Genetic engineering has been used to create a mutant of *E. coli* that can produce ethanol in the presence of 15 mM of furfural [102,

103]. This mutant was produced using known experimental information; it is possible that, just like metabolic engineering, a model could improve furfural tolerance further.

Recombinant Cellulolytic Microorganisms

The recombinant cellulolytic strategy begins by using a microorganism that selectively produces the desired product. This microorganism then could be engineered to include a cellulolytic system to enable growth on cellulose. This cross species gene expression (also known as heterologous gene expression) can be applied to a number of different process goals, but in this case it is specifically discussed in terms of utilizing cellulose as a carbon source.

The heterologous expression of cellulase systems has seen evaluated in numerous yeast species including *Saccharomyces cerevisiae* [104], *Pachysolen tannophilus* [105], *Candida shehatae* [106], and *Pichia stipites s*[107]. Recent research efforts in bacteria include *Zymomonas mobilis*[108, 109], *Escherichia coli* [110, 111] and *Klebsiella oxytoca* [112, 113]. This section will focus almost exclusively on *S. cerevisae,* as that has both extensive research performed, as well as being the standard microorganism in current industrial ethanol production.

When designing a cellulase enzyme system, it is first necessary to identify what components are necessary to be added to the non-cellulolytic organism. As discussed in the previous section on enzymatic hydrolysis, a noncomplexed cellulase system must have a cellobiohydrolase attacking the reducing ends, a cellobiohydrolase attacking the nonreducing ends, an endoglucanase, and a β-glucosidase. Of these, the cellobiohydrolases should have a cellulose binding module (CBM). For complexed cellulases, the following are required: a scaffolding protein with a CBM; two cohesins and a domain that binds to a cell wall-anchoring protein; a cell wall-anchoring protein; and at least one exoglucanase and an endoglucanse, both containing dockerins capable of binding to the scaffolding protein. Additionally, either a β-

glucosidase or cellodextrin and cellobiose phosphorylases, together with the appropriate permeases, would be required [49, 114]. From a practical standpoint, the engineering of a non-complexed cellulase system appears to be much easier than that of the complexed systems.

Cellulase expression in *Saccharomyces cerevisiae*

The expression levels of cellulases in *S. cerevisiae* have improved significantly in recent years. It has been reported that two critical cellulases, Cel6A (CBH1) and Cel7a (CBH2) have achieved a maximum titer at values corresponding to 4-5% of total cell protein, which meets the calculated levels for growth on cellulose in an industrial process [114].

An approach that has been used for cellulase expression is to closely resemble the naturally occurring systems. *Trichoderma reesei,* has been analyzed extensively for its application in cellulase production for enzymatic hydrolysis. The functional enzyme classes in *T. reesei* from the Herpoel-Gimbert study [115] have been 100% expressed in *S. cerevisiae* while the enzymes identified in Nagendran *et al.* have been 80% expressed in *S. cerevisiae* [116]. Because of the intricacy of complexed cellulase systems, the expression has not been as successful with only 20% of *C. thermocellum* cellulases functionally expressed in *S. cerevisiae* [70].

Process and Challenges using *S. cerevisiae* in Consolidated Bioprocessing

After designing and expressing the cellulase systems, the next step is to test the microbial system on phosphoric acid swollen cellulose (PASC). Recent reports have achieved 75% of theoretical ethanol yield by optimizing the level and ratios of cellulase enzymes in *S. cerevisiae* [117], which shows a large improvement over the approximately 27% conversion of PASC first published in 2007 [118]. This report is promising in the improvements made, however much more improvement needs to occur for industrial success. Additionally, PASC is more of a

model substrate due to its extreme reactivity. Therefore these organisms would likely need to undergo further refinement to achieve high yields on industrially pretreated cellulose, which will be considerably more recalcitrant than PASC.

In addition to designing appropriate cellulase systems as well as achieving high cellulase expression levels in yeast, there are a number of other challenges facing the recombinant strategy. By adding a large amount of heterologous genes to yeast, long-term gene stability is an issue that should be addressed before industrial acceptance. Gene mutations over time could reduce the expression levels of the cellulases, which obviously would decrease recombinant organism's effectiveness in CBP.

Conclusion

Lignocellulosic biomass is a potential renewable feedstock for production of value added chemicals. However, biomass has a recalcitrant structure that cannot be used without additional processing. The useful components of cellulose and hemicellulose are intertwined with lignin and are bound together in a very complex arrangement. Liberating the cellulose and hemicellulose fractions of biomass occurs through a process called pretreatment. There are many total options of pretreatment, each with unique advantages and disadvantages as discussed in the chapter above.

Once the biomass is pretreated, the cellulose and hemicellulose must be further reduced to 5 and 6 carbon sugars for use in fermentation. This step occurs typically in a step called enzymatic hydrolysis, where external cellulases are added to the pretreated biomass. The enzymes break down sugar and fermentation and product removal would then occur.

A second strategy called consolidated bioprocessing was presented as a means to use pretreated biomass. In this strategy, a cellulolytic microorganism is used for the steps of

enzymatic hydrolysis and fermentation. This reduces the number of process vessels, and introduces potential synergies to the enzymatic hydrolysis and fermentation steps. Several options for microorganisms for use in CBP were presented in this chapter; however, much work still needs to be performed to fully develop these microorganisms. Future chapters will explore one of these microorganisms, *T. fusca,* more in-depth and evaluate it for potential use in consolidated bioprocessing.

Chapter 2 - Review of Consolidated Bioprocessing research using *Thermobifida fusca*

Thermobifida fusca (formerly *Thermomonaspora fusca*) is an aerobic, moderately thermophilic, filamentous soil bacterium that has been studied for extracellular enzymes that are thermostable, have high activity, and are stable across a broad pH range [119]. It is commonly found in heated organic materials, such as compost piles, rotting hay, or manure piles as a degrader of plant cell walls [120]. *T. fusca* was introduced in the previous chapter as a potential cellulolytic organism for consolidated bioprocessing. The work in this study uses strains of *T. fusca* derived in a previous study through laboratory evolution, as well as information on optimizing fermentation parameters for cellulase production. This chapter provides an overview of all relevant information regarding *T. fusca* that was studied prior to performing work.

Study of Fermentation Conditions for *T. fusca*

An initial study that was performed with *T. fusca* was to determine the fermentation parameters for *T. fusca* that were necessary for growth, and a high cellulase activity [121]. This study was performed using a wild type strain of *T. fusc*, and consisted of multiple phases. The first phase was to compare growth in a shaken flask and fermentor as well as cellulase and endoglucanase activity in each condition. It was determined that *T. fusca* exhibits markedly different growth behavior depending on if it is grown in a shaken flask or fermentor. The shaken flask showed a higher total cellulase and endoglucanase activity, but the bioreactor resulted in a higher specific cellulase and specific endoglucanase activity.

Another outcome from this study was to determine the activities in the bioreactor under different circumstances. It was determined that oxygen transport was important to the cell, and as such the stir speed and aeration rate should be optimized. Growth curves, cellulase, and

endoglucanase activity was measured for various conditions. It was determined that at high aeration rates (above 1 vvm) there was a tradeoff between overall cellulase activity (increased with stir speed) and endoglucanase activity (decreased with increasing stir speed). Too high of stir speeds, above 400 rpm, was found to be detrimental to cell growth.

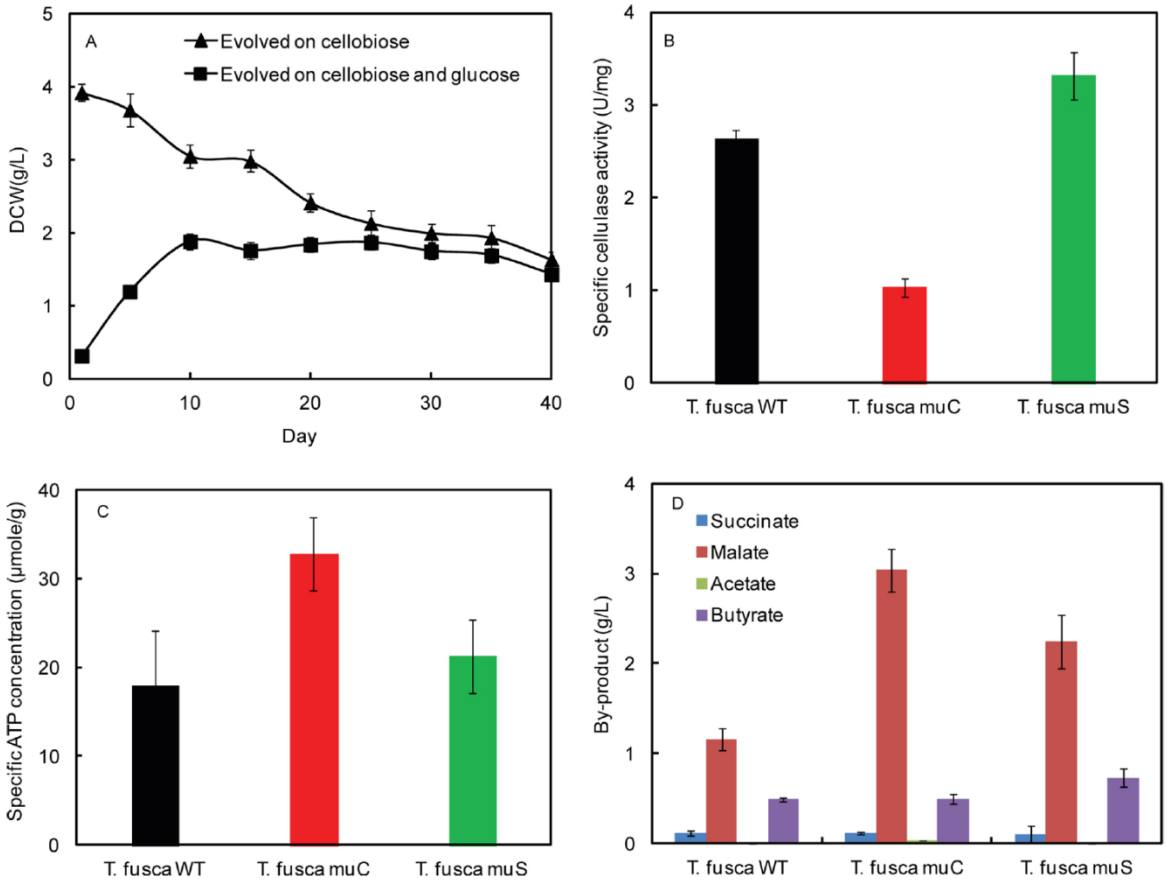
Laboratory evolved *T. fusca*

A second study used the wild type strain of *T. fusca* as a starting point in a laboratory evolution process for two different conditions [122]. The two conditions were both maintained at 55°C for 40 days on Hagerdahl medium. In the first condition, *T. fusca* was grown for the 40 day period with 0.1% cellobiose. In the other condition, *T. fusca* was grown for 40 days on Hagerdahl medium where the carbon source was switched every day between 0.1% cellobiose and 0.1% glucose. The strain that is the final product of the first condition is what is known as muC, which is short for mutant grown on cellobiose; the second strain from the second condition is known as muS, which stands for mutant grown by switching sugars. The cells were constantly maintained in exponential growth during the laboratory process by a serial passage of cells into fresh medium at specific times. The adaptive evolution was halted at 40 days because the cell growth and cellulase activity were stable. The muC strain was propagated for nearly 284 generations and the muS strain was propagated for 220 generations. The cell yields for the muC strain decreased from 3.92 g/L to 1.63 g/L, while the muS strain yields increased from 0.31 g/L to 1.43 g/L.

Characterization of the phenotypes of the two new strains was performed to understand the differences between the two strains. When tested for growth with different carbon sources, it was noted that the muC strain performed very poorly and exhibited slower growth than when grown on cellobiose. Meanwhile, the muS strain was able to maintain similar growth patters to

cellobiose or glucose which indicated the evolution was able to transform this species into more of a generalist strain that could grow on a wide variety of substrates. The specific cellulase activity of both strains, as well as the wild type (WT) was analyzed, along with the specific ATP concentrations. Finally, the by-product of each was measured of the primary metabolites to determine how the evolution impacted the metabolic pathways. The results of this previous study are shown in figure 2-1.

Figure 2-1 Laboratory evolved strains of muC and muS were compared in a number of different ways. These were also compared to the wild type (WT) strain to identify the changes derived from evolution. Used with permission from [122]



The comparison of the two phenotypes was continued even more in-depth by performing a genome re-sequencing of the two evolved strains. In total, it was determined that there were 18 SNPs in the muC strain and 14 SNPs in the muS strain. Some mutations could be linked to a

phenotypic change, such as a mutation that causes the muC strain to secrete acetate as an end product but not muS, or the ability of muS to switch between carbon sources. However, some mutations were ambiguous on the overall effect, as there were no observable changes and weren't apparent from genetic expression. However, the cells were characterized in-depth enough to be used in future experiments.

Conclusion

The muC and muS strains of *T. fusca* were laboratory evolved strains that had different growth characteristics than the wild type strain of *T. fusca.* The muC strain was developed as a specialty strain that became efficient at growth on cellobiose, while the muS strain was developed as a strain that could metabolize different types of sugar. Both strains were found to be similar to the wild type strain by being mixed acid fermenters, meaning they produced a variety of organic acids as end products of fermentation rather than selectively producing a single product. The next study evaluates the end products further, and attempts to selectively produce butyric acid as a final metabolite of *T. fusca.*

Chapter 3 - Butyric Acid Production using *T. fusca* in Consolidated Bioprocessing

While much work has been done researching the production of ethanol from biomass, another interesting and promising strategy is to produce organic acids and other value added chemicals from biomass. Assuming fossil fuels will diminish over time, it is important to not just replace the as a transport fuel, but also all of the other products derived from these fossil fuels.

Butyric acid, a 4-carbon fatty acid, is widely used in the chemical, food, and pharmaceutical industries. It is currently produced industrially through chemical synthesis of petroleum products. Historical attempts have been made to produce butyric acid through the acetone-butanol-ethanol (ABE) fermentations, which was the second largest industrial fermentation process in the world. However the ABE fermentation process is carried out by *Clostridium acetobutylicum* and *Clostridium beijerinckii,* which are both anaerobic organisms, and are challenging to use for industrial fermentation because they typically have a low growth rate and energy level compared to aerobes [123].

These ABE fermentation systems also generally use glucose as the primary substrate. Of these *Clostridium* based platforms, only *C. butyricum* and *C. thermocellum* is able to directly use complex carbon sources like lignocellulose while the others are only capable of utilizing simple sugars [130]. These studies have been researched extensively and thus there are a variety of results available. Of these studies, one has used *C. tyrobutricum* to ferment a hydrolysate from pretreatment, while the others directly use a refined simple sugar as the substrate [131]. These studies are shown in the table below. The final concentration of butyric acid ranged from 7.3 g/L all the way up to 62.8 g/L in these studies. In addition to traditional studies where fermentation

occurred in a batch or fed-batch system, several of these studies performed novel process engineering efforts to help improve yields. Continuous fermentation, extractive batch fermentation, and immobilized cell fermentations all were performed to try and remove or neutralize the butyric acid. These studies all noted that butyric acid is toxic to the cells, and that the best way to improve final concentration was to remove the butyric acid.

Table 3-1 Summary of previous research into butyric acid fermentation. Used with permission from [131].

Strain	Sugar	Final conc. (g/L)	Fermentation mode
C. butyricum ZJUCB	Glucose	12.25	Batch
		16.74	Fed-batch
C. butyricum S21	Glucose	7.3	Batch
	Sucrose	10	Extractive batch
	Sucrose	20	Pertractive fed-batch
C. butyricum S21	Lactose	18.6	Batch
C. thermobutyricum ATCC 49875	Glucose	10.04	Batch
		19.38	Continuous
C. beijerinckii	Lactose	12	Batch
C. populeti ATCC 35295	Glucose	6.3	Batch
C. tyrobutyricum JM1	Glucose	13.76	Batch
C. tyrobutyricum CIP 1–776	Glucose	45	Batch
	Glucose	62.8	Fed-batch
	Glucose	28.6	Fed-batch
C. tyrobutyricum ATCC 25755	Glucose	24.88	Fed-batch (immobilized cells)
	Glucose	43.4	Fed-batch (immobilized cells)
	Glucose	53	Fed-batch (immobilized cells)
C. tyrobutyricum CNRZ 596	Glucose	44	Batch
	Glucose	16.8	Continuous
	Glucose	33	Continuous (cell recycle)
C. tyrobutyricum ZJU 8235	Jerusalem artichoke hydrolysate	27.5 g/L	Batch
		60.4 g/L	Fed-batch (immobilized cells)

The cell productivity of the above studies was not reported, so only the production of butyric acid in g/L is known. A wide variety of carbon sources were used in the ABE fermentation research. Of the ABE fermentation research, no studies were found where butyric acid was produced direct from cellulose as part of a consolidated bioprocessing scheme. The

authors of the review mentioned CBP as part of future work, but no published results were found [130].

The objectives of this study were to identify the carbon source that *T. fusca* used most efficiently for butyric acid production, determine the optimal stir speed and aeration rate for butyric acid production, as well as produce a first report of using the consolidated bioprocessing scheme to produce butyric acid direct from an LCB feedstock. The final objective was to determine the metabolic pathways, so that future genetic engineering work could be performed on *T. fusca* to increase butyric acid production.

Materials and Methods

Materials

All materials for this study were purchased by Kansas State University or were donated for the study. All chemicals used were ordered from Sigma Aldrich. The equipment was both purchased and donated by Fisher Scientific as part of a new lab startup program. The corn stover used in this study was donated by Paul Merklein, a farmer in Stuttgart, Kansas.

Culture Conditions

The muS strain was grown in Hagerdahl medium in either shaken flasks or a fermentor, depending on the condition being tested. In the shaken flasks, 50 mL pre-cultures were grown at 55°C and 250rpm for 12-24 h in a 500 mL Erlenmeyer flask. The growth cultures were then inoculated in a 300 mL volume in a 500 mL Erlenmeyer flask by using 5% of the pre-culture. The culture was grown at 55°C and 250rpm for a range of 42-48 h, with each condition being halted when the cell density began to decrease.

For the fermentor experiments, 200 mL of pre-culture was grown for 12-24 h in a 500 mL Erlenmeyer flask and used to inoculate the bioreactor (BIOSTAT B) with a 5 L working volume. The fermentor had a jacketed glass vessel with stainless steel top-plate and baffle inserts. The experiments were conducted with four baffles and two six-bladed disk impellers for mixing. Cells were cultured in the bioreactor at 55°C for 42-68 h, with each condition being halted when the cell density began to decrease. Stir speeds and aeration rates varied depending upon the specific experiment.

Cell density and by-product measurement

The growth physiology of *T. fusca* causes challenges to accurately measure the culture density during fermentation, because they are filamentous cells that aggregate. Because of this,

it is important to have other means to determine growth during fermentation. A 1 mL culture was centrifuged in an Eppendorf 5424R Microcentrifuge at 10,000 x g for 5 minutes, and then re-suspended in fresh media to wash, followed by another round of centrifuge at 10,000 x g for 5 minutes. Sediments were dissolved in 200 µL 50mM Tris-HCl buffer (pH 6.8) containing 2% SRS, 0.1M DTT and 50% glycerol. Samples were then pulsar sonicated using a QSonica Q125 Pulsar Sonicator for 10 minutes in an ice bath at 70% strength. A Bradford protein assay [124] was performed on the supernatant (using Thermofisher Genesys 10 S) after centrifuging at 10,000 x g for 5 minutes. A standard curve was generated (not reported) plotting the dry cell weight against the overall protein content, and it was found that the dry cell weight is proportionally related to the overall protein content.

Byproducts were detected using an HPLC system equipped with Bio-Rad HPX-87H ion exclusion column. The mobile phase was 0.005Mol/L H_2SO_4 at the rate of 0.6 mL/min and IR and UV detectors were used. The yield of butyric acid is defined as grams butyric acid produced per gram carbon entering the cell.

Pre-treatment of corn stover

For the experiment where corn stover was used, the pretreatment occurred through mechanical milling by using a burr coffee grinder. No characterization, such as average particle size, was performed prior to use in fermentation.

Addition of precursors of butyric acid into washed cell system

T. fusca was grown in 500 mL shaken flasks to mid-log phase, when cells were obtained and washed by an equal volume of fresh media without a carbon source three times, for 5 minutes each time. The cells were then transferred to 500 mL shaken flasks with fresh media without a carbon source to grow at 55°C and 250 rpm. After adding cells to the media, 0.1% of

butanoyl-CoA and butanoyl-P were added into the flasks, respectively, and allowed to culture for an additional 12 h. After cultivation, the concentration of butyric acid in the supernatant was determined.

Enyzme activity

Determination of cellulase activity was measured according to the published protocol [121].

To determine Butyryl-CoA transferase activity, the assay mechanism involved the condensation of the formed acetyl-CoA with oxaloacetate, and the subsequent liberation of CoA-SH under the influence of the citrate synthetase. The CoA-SH reacted with 5,5'-dithio-bis-(2-nitrobenzoate) to form the yellow thiphenoate anion, 2-nitro-5-mercaptobenzoate. The enzyme activity was determined by following the formation of this colored product at 410 nm at 25°C and a pH of 7.

The enzyme assay was adapted for microtiter plate measurements. A 50 mM solution of Ellman's reagent was freshly prepared and kept on ice until use. 20 µL of the enzyme solution at the appropriate dilution was placed in a well of a 96-well plate (Fisher Scientific Nunc). The reaction was started by mixing thoroughly 4 µL of Ellman's reagent to the reaction mix and adding immediately to the enzyme solution in the well. A master mix was prepared for several reactions and a volume of Ellman's reagent was added accordingly before dispensing 180 µL of the mix into each well that contains the enzyme of the microtiter plate. The increase in absorbance was followed at 410 nm every 30s over 5 min with a thermo-regulated microtiter plate spectrophotometer set at 25°C and data was recorded. The plate was briefly shaken before each reading. The negative control contained all reagents but butyryl-CoA.

RNA Preparation and real-time PCR

Molecular-level differences in the cultures of *T. fusca* were studied using real-time PCR. The samples of cells were harvested at mid-log phase. The cells were centrifuged at 10,000 x g for 5 minutes. The cell pellets were re-suspended in sufficient volume of Qiagen RNAprotect Bacteria Reagent, as prescribed by Qiagen. The cells were incubated at room temperature for 5 minutes, and then pulsar sonicated at 10% strength for 2 minutes. The RNA was then isolated using the Qiagen RNeasy Midi kits, using the protocol suggested by Qiagen. The real-time PCR measurements were performed in the ABI Prism 7900 Sequence Detection System, using the TaqMan One Step PCR Master Mix Reagents Kit. The cycling conditions were 48°C for 30 minutes; 95°C for 10 minutes; and 40 cycles of 95°C/15 seconds and 60°C/1 minute. The cycle threshold was determined to provide the optimal standard curve values. The genes were measured along with one housekeeping gene (Tfu_02001404), which was used as a control with all reported transcript levels normalized to the housekeeping gene. The probes were labeled at the 5' end with FAM (6-carboxyfluoresceine) and at the 3' end with TAMRA (6-carboxytetramethyrhodamine) [121]. All data was measured in triplicate.

Results and Discussion

Effect of carbon sources on butyric acid fermentation

It is known and described that *T. fusca* produces a variety of cellulases, hemicellulases, and xylanases which are able to degrade most lignocellulosic biomass [125]. While classified as cellulolytic, *T. fusca* has the ability to uptake a variety of carbon sources to support growth, and this was seen in the lab through the ability to grow on almost any source [122]. Prior to this study, however, the effect of different carbon sources on chemical productivity had not been explored. The sources of carbon explored were cellobiose, xylose, sucrose, glucose, mannose,

lactose, and acetate with the goal of producing butyric acid. These sources were chosen based on known growth of unofficial samples, as well as evaluations of the metabolic pathways on the Kyoto Encyclopedia of Genes and Genome (KEGG). In all cases 3-6 g/L of initial carbon sources were added into shaken flasks for culturing up to 43 h. As shown in figure 3-1, *T. fusca* consumed all of the cellobiose, xylose, sucrose and acetate, but only consumed 60% of the mannose, and 50% of the glucose and lactose. The cellobiose and sucrose were used quickest by *T. fusca*.

In terms of butyric acid production, cellobiose yielded the most with more than 1 g/L of butyric acid produced, followed by acetate, which produced 0.7 g/L. The lowest were 0.1 g/L of butyric acid produced on lactose and glucose. The curves of butyric acid production for each carbon source are shown in Figure 3-1.

The yield of butyric acid was calculated as gram of butyric acid produced per gram of carbon input into *T. fusca* (g/g-C). The highest butyric acid yield was found for growth on cellobiose (0.67 g/g-C) with the second highest yield on acetate (0.46 g/g-C). The lowest was glucose. These results can be seen in figure 3-1.

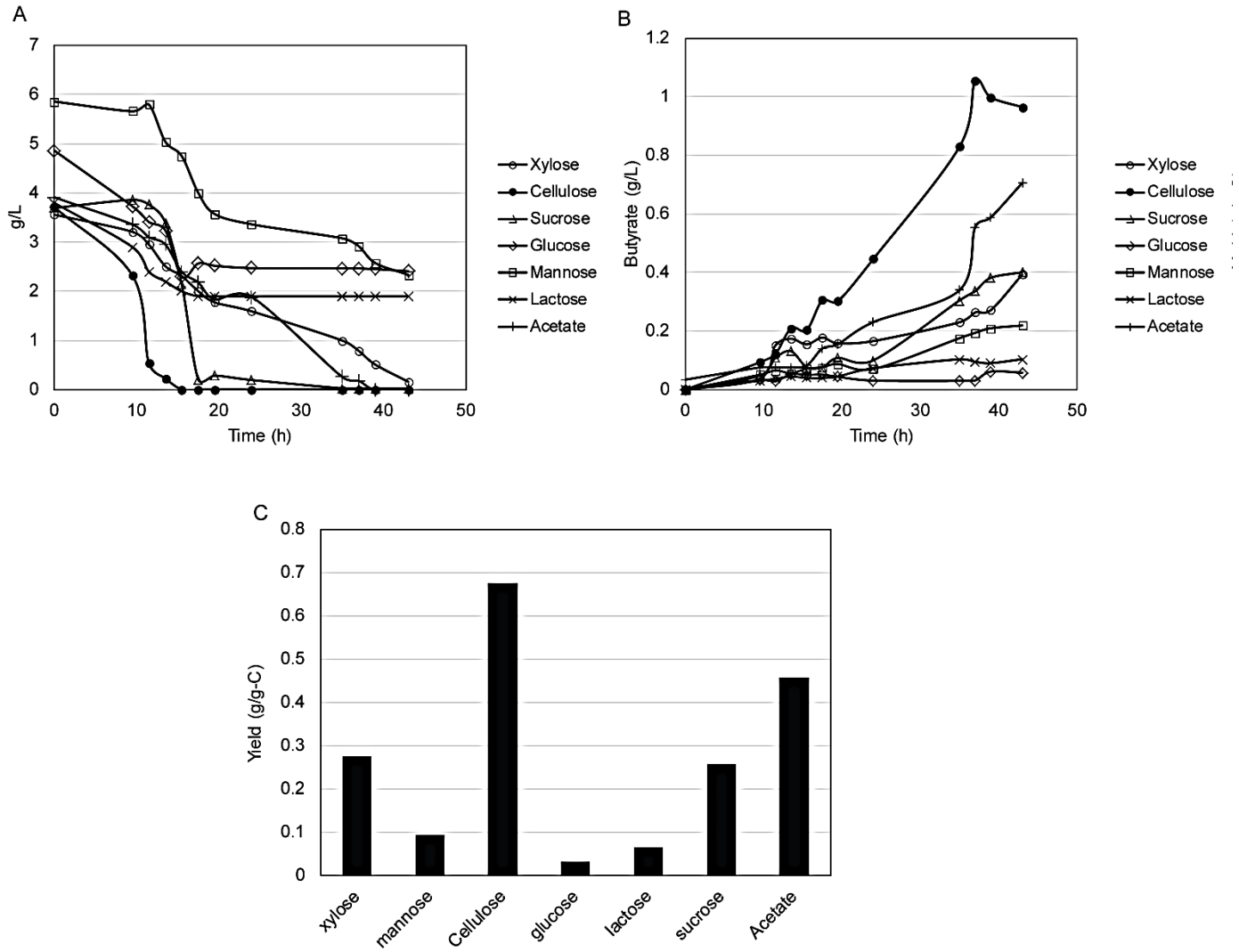
A major byproduct during the pre-treatment of biomass is acetate, and it has been noted as an inhibitor of microbial growth [126]. Testing in the lab indicated that *T. fusca* could be grown on more than 10 g/L of acetate, indicating less inhibitory effects than other microorganisms. Because of the growth on acetate, *T. fusca* not only could consume hydrolysates of biomass, but also could consume acetate as well for value added chemical fermentation.

The *T. fusca* muS strain used in this study is a mutant strain with a significantly slower growth rate compared to the wild-type strain. However, this strain was found to have higher

cellulase activities than the wild-type strain [122]. *T. fusca* has a high activity of beta-glucosidase, which directly degrades cellobiose to glucose [127]. Also, *T. fusca* can directly uptake cellobiose without degrading it to other sugars first (ABC Transporter). During fermentation, especially the muS strain that has high cellulase activity but slow growth, a lot of glucose was present from cellobiose degradation in the broth. Figure 3-1 shows no cellobiose available in the broth after 16 h, however, the HPLC results showed that glucose was present (data not shown). This indicates that *T. fusca* gradually consumed glucose as a secondary carbon source after cellobiose was unavailable.

Similarly, sucrose was degraded first to D-fructose and D-glucose by an alpha-glucosidase, which is the enzyme present in *T. fusca* that is known to degrade sucrose. Once the fructose was used, the glucose was consumed. While not studied any further, this order of sugar uptake could also be useful in industrial applications by co-culturing a sole glucose-favoring organism with *T. fusca* to consume the glucose that is the last carbon source by *T. fusca.* Because of the relatively high temperatures during fermentation, a co-culture may not be optimal; however, it would be advantageous to explore some sort of synergistic fermentation platform because the biomass could be degraded by *T. fusca*, which would prefer longer chain sugars, while giving a constant supply of short chain sugars to a different organism.

Figure 3-1 Effect of carbon sources on butyric acid production by *T. fusca*. Samples were performed in triplicate, and combined in equal parts prior to performing HPLC results. (A) The consumption of different carbon sources in g/L of carbon remaining; (B) production of butyric acid on different carbon sources in g/L butyrate produced; (C) the yield of butyric acid on different carbon sources (gram butyric acid produced per gram carbon input, g/g-C)



Optimization of fermentation Parameters

During the fermentation, *T. fusca* hyphae formed cell balls, which caused uneven distribution of *T. fusca* cells in the fermentation broth making substrate and oxygen transfer more difficult than most bacteria. To facilitate cell access to nutrients and oxygen, it as crucial to have

a reasonably high stir speed. However, the hyphae are very sensitive to high shear stress from agitation and too high of stir speed is detrimental to cell growth. It has been shown that stir speed, aeration rate, and oxygen supply rate are key parameters that affect fermentation of *T. fusca* [121].

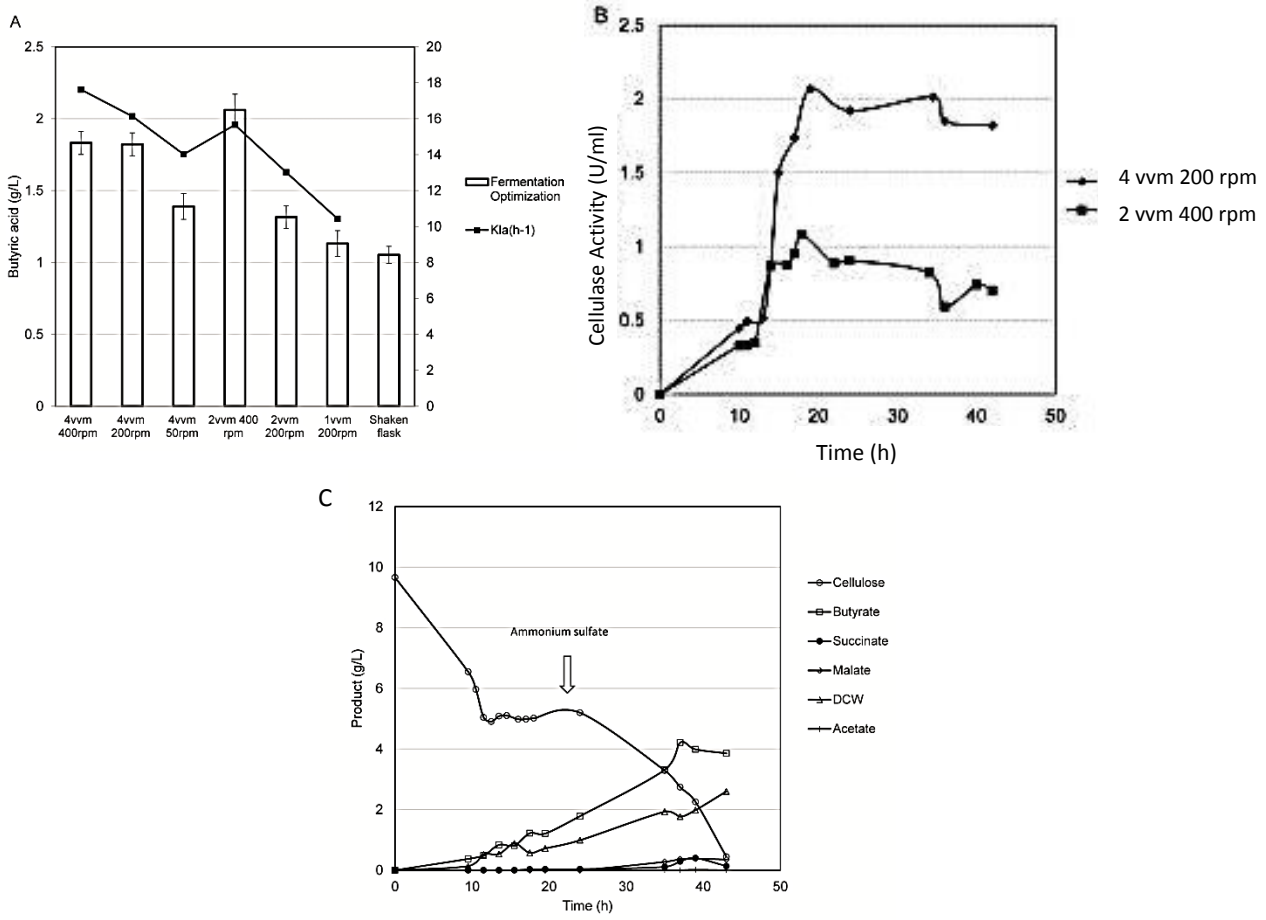
The previous study showed *T. fusca* is not tolerant of high stir speeds (>400 rpm), so the rates of 400, 200, and 50 rpm were used in this study. Oxygenation rates of 4, 2, and 1 vvm were used to study the effect of aeration on butyric acid production. These parameters are interesting to study because of the "sticky" fermentation of *T. fusca*, where the cells aggregate together. No matter how well optimized the fermentation broth is, if the cells can't obtain nutrients they will not be productive.

Figure 3-2 shows that when the aeration rate was set at 4 vvm, *T. fusca* produced almost the same amount of butyric acid (1.8 g/L) with stir speeds of 400 and 200 rpm. At 50 rpm, the butyric acid concentration was only 1.4 g/L, which may be due to the stir speed was not high enough to properly transfer oxygen and nutrients to the cells. Often, oxygen transfer rate is an important factor in scale up, and models exist to keep the oxygen transfer rate constant throughout the fermentation at any volume once an optimum is found. These models usually dependent upon the volumetric oxygen transfer coefficient, K_La, as the driving force [128]. The K_La was calculated for the fermentor experiments, and it was found to decrease as stirring speeds were reduced at 4 vvm.

At an aeration rate of 2 vvm, only 400 and 200 rpm were studied. A rate of 50 rpm was not used as it was proven to ineffective at a higher aeration rate. It was shown that 400 rpm produced 2.1 g/L, the highest yield in the study, and 200 rpm produced 1.3 g/L.

An additional fermentation was performed with rates of 1 vvm and 200 rpm, because a previous study found these conditions were optimal for cellulase production. However, *T. fusca* only produced 1.1 g/L of butyric acid in this case, which was the lowest amount in the fermentor, and was nearly the same level as the shaken flask used as control. This indicates the large variation fermentation parameters have on overall output of the cell – the optimum for one study (cellulase production) is the worst in a different study (butyric acid production).

Figure 3-2 Fermentation of butyric acid on cellulose. (A) Butyric acid production on different stir speeds and aeration rate in g/L of butyric acid produced as well as the K_La in h^{-1}; (B) cellulase activity under different fermentation conditions; and (C) Fermentation curve of producing butyric acid with each product listed in g/L produced



Cellulase Production

The cellulase expression places a large metabolic burden on cellulolytic organisms, and studies have suggested that as much as 20% of the total cell mass can be associated with cellulolytic components [125]. With a goal to use *T. fusca* as a producer of butyric acid from lignocellulosic biomass, then it is important to determine a proper tradeoff between cellulase production and butyric acid production; cellulase is needed for the cell to reproduce, but the end goal is production of butyric acid.

The top two conditions for butyric acid conditions were evaluated to determine the cellulase activities, as shown in Figure 3-2. While the yield of butyric acid was highest at 2 vvm and 400 rpm, the average cellulase activity (1.08 U/mL) was lower than the condition at 4 vvm and 200 rpm (2.06 U/mL). This further indicates high cellulase activity is not advantageous for butyric acid production.

Batch Fermentation on Cellulose

A batch fermentation on cellulose was performed over the entire fermentation range. An initial cellulose concentration of 9.66 g/L was used with an inoculum size of 5%. Within 12.5 h, almost half of the cellulose was consumed. However, after 12.5 h, the cells stopped consuming cellulose and the growth stopped. Butyric acid was produced during this 12.5 h period, but the production rate of butyric acid was significantly decreased. No succinic acid or malic acid was detected during this time period.

The hypothesis for why cellulose was no longer consumed was an imbalance in the carbon to nitrogen ration. At the beginning of the fermentation, there was a carbon/nitrogen ration of 6.4:1 and at the 12.5-hour point, the ratio increased to approximately 10:1. At 24 h, ammonium sulfate was added to the bioreactor to help balance the carbon/nitrogen ratio, which is critical for flux distribution. The cellulose continued to be consumed after addition, and the cell growth and butyric acid continued to increase. The final butyric acid titer was 2.1 g/L with other major products of malic acid and succinic acid below 0.4 g/L, and acetate below 0.1 g/L (data not shown).

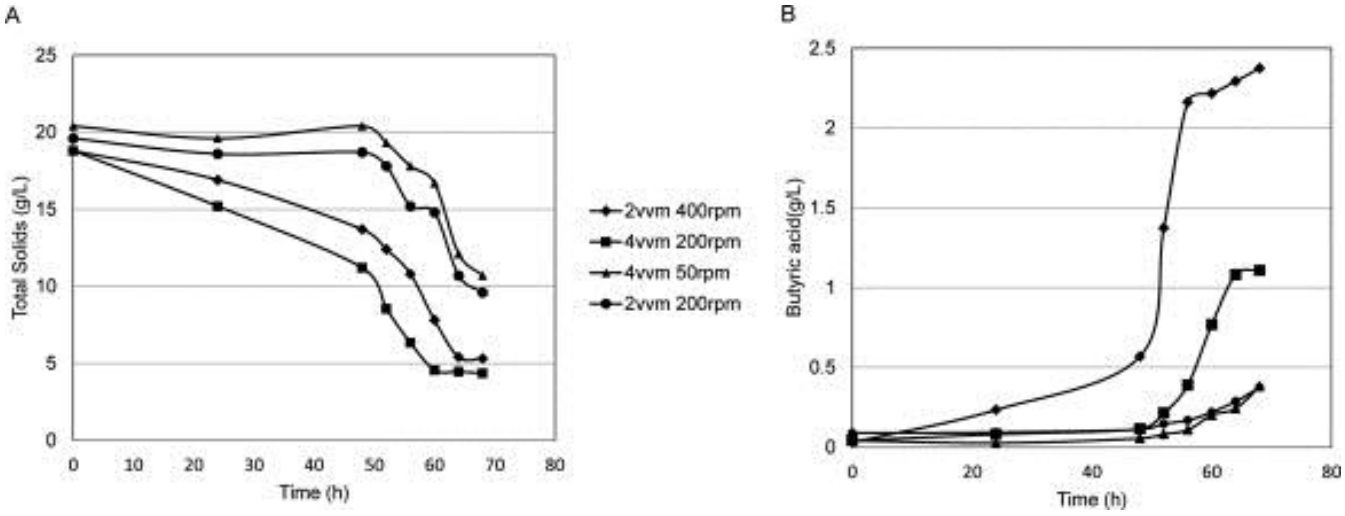
Batch Fermentation on Milled Corn Stover

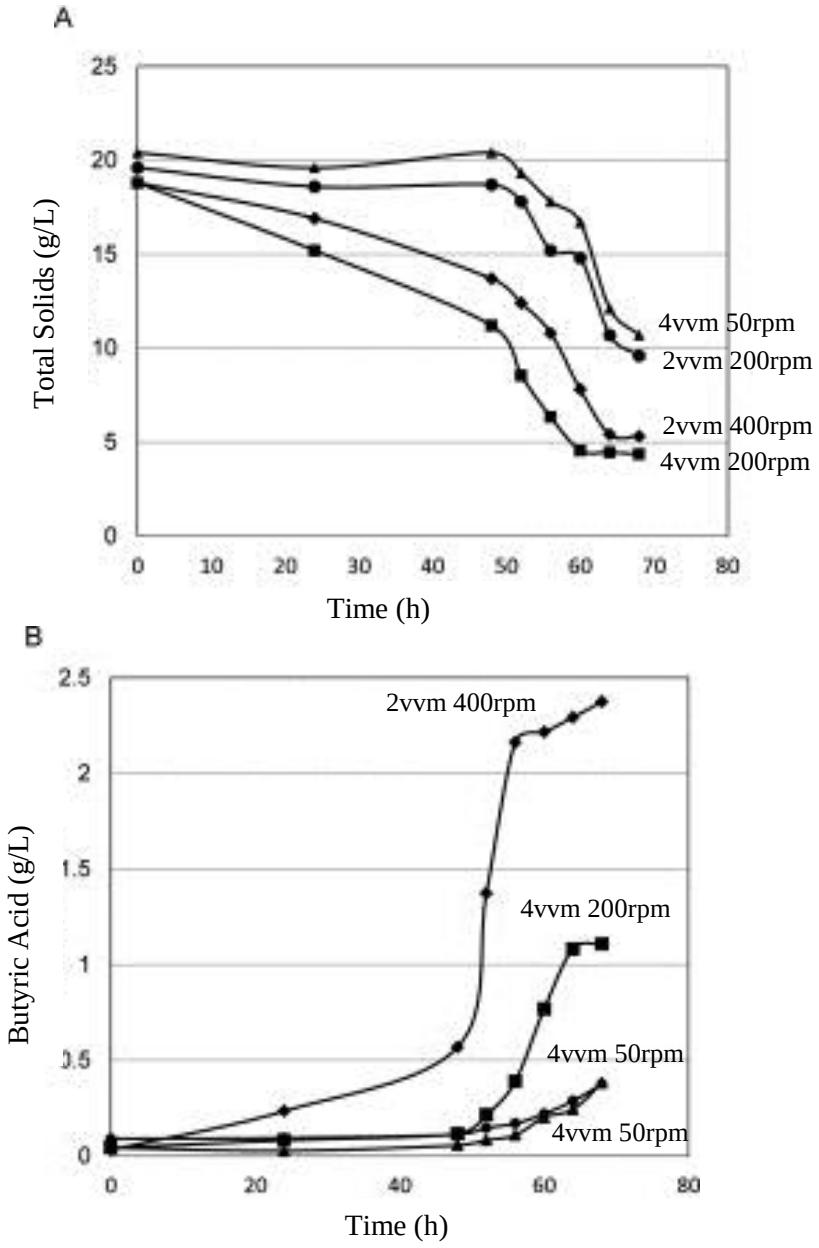
The final fermentation study evaluated the production of butyric acid from corn stover using the 4vvm and 200 rpm; 2 vvm and 400 rpm; 4 vvm and 50 rpm; and 2 vvm and 200 rpm

conditions. The highest titer of butyric acid was 2.37 g/L at the 2 vvm, and 400 rpm condition. The degradation of corn stover (initial concentration approximately 20 g/L) was significantly slower than pure laboratory grade cellulose with slow cell growth before 50 h and a total time of 60 h to reach stationary phase.

The consumption of LCB by *T. fusca* was significantly slower than pure laboratory grade biomass for a number of reasons. Biomass consists of a combination of cellulose, hemicellulose, xylan and lignin all intertwined. This causes steric hindrance for the cellulases, hemicellulose, and xylanases as discussed previously. This fermentation was just performed on milled corn stover, and the results would likely be very different if another form of pretreatment was used. Additionally, as discussed above, lignin is an inhibitor of cell growth, and it likely reduced *T. fusca's* growth, although that was not specifically determined in this study.

Figure 3-3 Fermentation of butyric acid on corn stover where all samples were performed in triplicate and then combined into a single sample prior to HPLC analysis. (A) Total solids consumption with the remaining solids listed in g/L; (B) butyric acid production on corn stover in g/L.





Identification of Pathways to Butyric Acid

After experimentally producing butyric acid through numerous fermentation systems, it was desired to determine the metabolic pathway for butyric acid production in *T. fusca*. Through analysis of the KEGG defined metabolic pathways, and comparing against the genome of *T. fusca*, it was determined that butyric acid in *T. fusca* was synthesized through the following route: acetyl-CoA was converted to acetoacetyl-CoA by acetyl-CoA acetyltransferase, and then

acacetoacetyl-CoA was converted to 3-hydroxybutanoyl-CoA by 3-hydroxyacyl-CoA dehydrogenase, which was catalyzed to crotonoyl-CoA by enoyl-CoA hydratase. The crotonoyl-CoA was catalyzed to butanoyl-CoA by butyryl-CoA dehydrogenase and then butyric acid was synthesized from butanoyl-CoA by butyryl-CoA transferase or phosphotransbutryrylase and butyrate kinase. There are five enzymes whose genes were annotated in *T. fusca* genome that are suggested to catalyze those reactions. However, no genes were annotated in the conversion of butanoyl-CoA to butyric acid. This proposed metabolic pathway is presented in figure 3-4.

Table 3-2 The expressions of the genes whose proteins catalyze the creations from acetyl-CoA to butanoyl-CoA

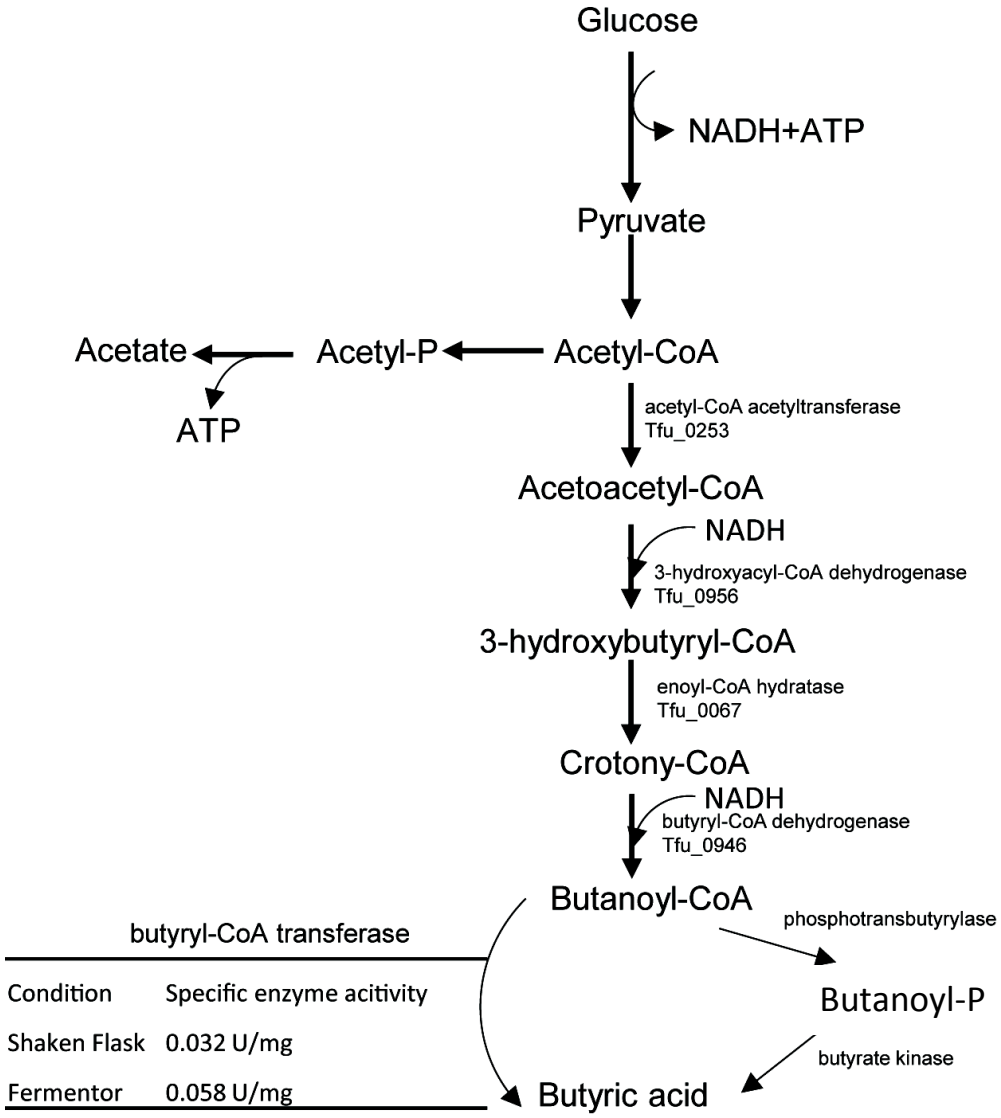
Condition	Tfu_0946 butyryl-CoA dehydrogenase	Tfu_0067 enoyl-CoA hydratase	Tfu_0956 3-hydroxyacyl-CoA dehydrogenase	Tfu_0253 acetyl-CoA acetyltransferase
Fermentor	0.820 ± 0.114	1.294 ± 0.052	1.300 ± 0.058	1.170 ± 0.06
Flask	0.459 ± 0.070	0.330 ± 0.017	0.468 ± 0.055	0.510 ± 0.01

The expression levels of the genes found in the *T. fusca* genome are shown in Table 3-2. Gene expressions of these genes were measured and mRNA transcripts were found to be present during fermentation in *T. fusca*. Gene expression was much higher in the fermentor than for shaken flasks for similar time points, which corresponds with higher yields in the fermentor than in shaken flasks.

In an effort to understand the route from butanoyl-CoA to butyric acid, butanoyl-CoA and butanoyl-lP were added into washed cells in media without carbon sources, respectively. No significant amount of butyric acid was produced from butanoyl-P whereas the addition of butanoyl-CoA yielded some butyric acid production. This indicated that no butyrate kinase exists in *T. fusca* but that there is a butyryl-CoA transferase that catalyzes this reaction as seen in figure 3-4. To confirm the existence of the butyryl-CoA transferase, an enzymatic assay was

conducted. The specific activity of butyryl-CoA transferase in the fermentor was again higher than that activity from a similar time point in the shaken flask.

Figure 3-4 Proposed pathway for butyric acid production in butyric acid. Based on a washed cell system, the genes are not present to convert Butanoyl-P to Butyric Acid, but the enzyme activity was present from Butanoyl-CoA to Butyric Acid.



Conclusion

This study showed that *T. fusca* can produce butyric acid directly from cellulose. It showed that the best stir speed and aeration rates were 400 rpm and 2 vvm. A batch growth from

a media starting with 9.66 g/L of cellulose produced 2.1 g/L butyric acid. To test the production of butyric acid on lignocellulosic biomass, milled corn stover was used to produce 2.37 g/L under optimal fermentation conditions. The pathway for butyric acid production was determined based on the genome, as well as confirmed experimentally.

Chapter 4 - RNASeq Analysis of *T. fusca*

Based on the information described in Chapter 3 about *T. fusca*, a plan for mRNA sequencing was created that would help to provide additional information regarding the gene expression throughout various fermentation conditions. It is well understood that environmental conditions greatly impact transcription in a cell, and the intent of this study was to analyze a broad range of conditions.

The goal for this study was to produce a database of transcriptomics results that could be used for future projects in genetic engineering or fermentation engineering. Some uses of this transcriptomics database could be to provide context on the environmental conditions that most impact fermentation, provide genes that could be knocked out or overexpressed to improve product yield, or assist in confirming metabolic pathways.

Study Protocol and Methods

Materials and Equipment

All materials for this study were purchased by Kansas State University or were donated for the study. All chemicals used were ordered from Sigma Aldrich. The equipment was both purchased and donated by Fisher Scientific as part of a new lab startup program. The chemicals and equipment, if not otherwise specified for this study are the same as for the previous study.

Fermentation Growth Curves

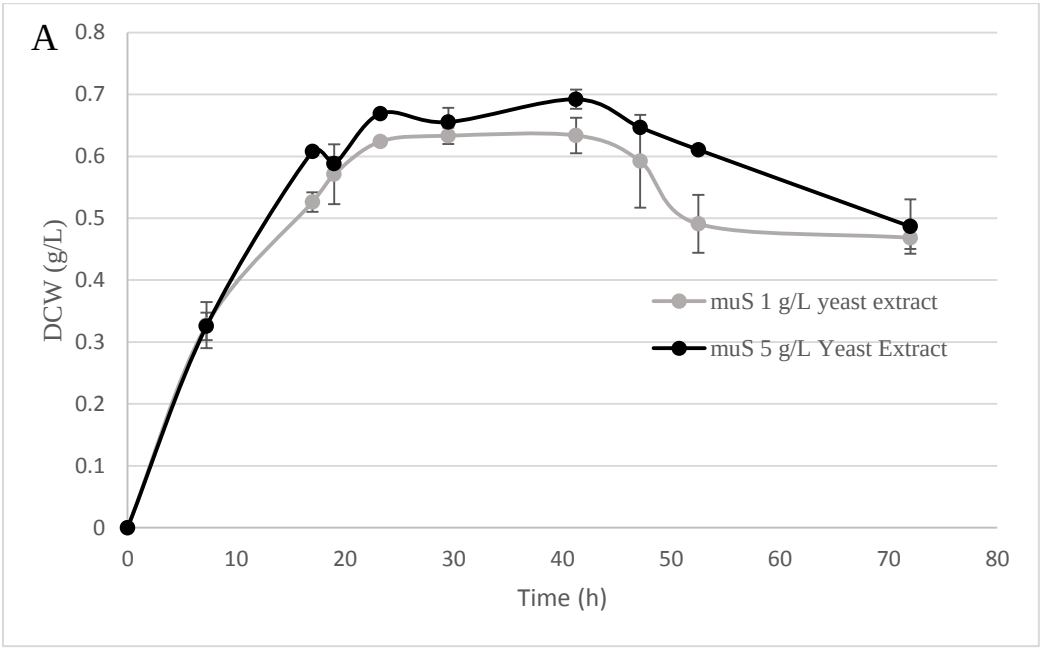
Before any samples could be taken, an introductory growth curve was plotted for the two strains when grown on cellobiose. While the exact growth curve will vary greatly for each condition, it was important to have a baseline so that the mid-log phase can be determined. The samples taken for isolation of mRNA should be harvested during the mid-log phase, because the

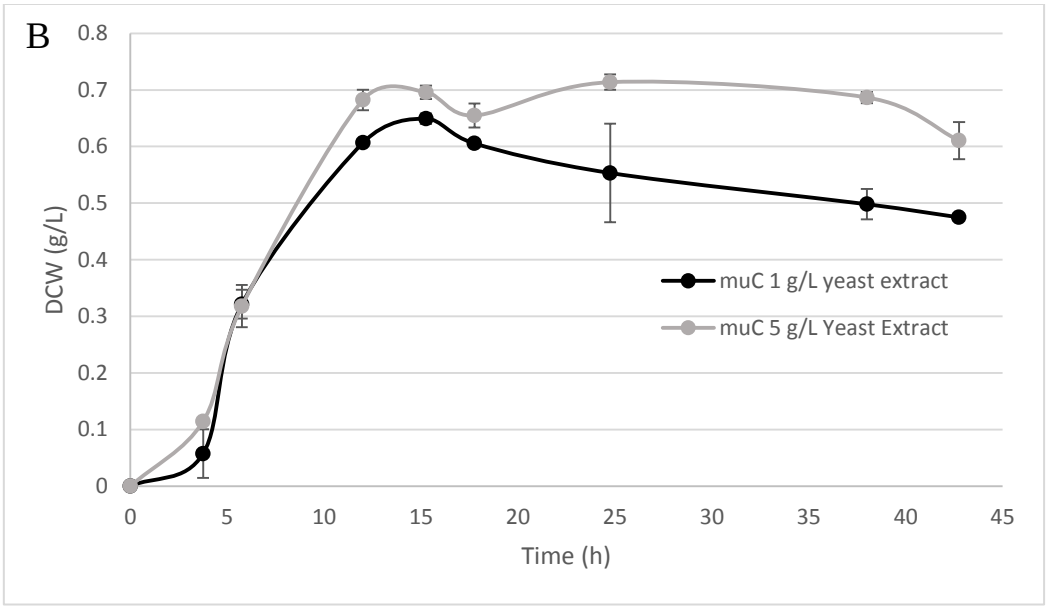
maximal growth that corresponds to the log phase is when the cell will be producing the most mRNA as the cell works on producing the protein and enzyme machinery necessary to grow. Since the goal of the mRNA analysis is to identify the genes that are transcribed by the cell, harvesting cells at the log phase align with that goal.

Ideally, an exact growth curve would be carried out for each specific condition in the mRNA study, so that the exact sample time could be planned out. However, due to resource limitation, only one growth curve could be performed. This curve is based on a simple shaken flask fermentation of 5 g/L cellobiose with either 1 g/L yeast extract or 5 g/L yeast extract added to a 2382 Hagerdahl medium that did not contain growth additives of biotin or thiamine. The fermentation occurred in 100 mL total fermentation volume in a 500 mL Erlenmeyer flask. The flask opening was covered with aluminum foil to allow a mostly aerobic environment. The inoculum size was 1% *T. fusca* and the fermentation was carried out at 55^{o}C and 170 rpm. These parameters were chosen because they were determined to be ideal for *T. fusca* growth based on informal studies performed in the lab (data not included).

To obtain a dry cell weight (DCW), a 1 mL sample was obtained and centrifuged at 10,000 x g in pre-weighed microcentrifuge tubes and the majority of the liquid decanted off. The cells were then placed in an oven to dry for 48 h, followed by weighing of the microcentrifuge tubes with the dry cells. The weight was then converted to the appropriate units of g/L. Additionally, when taking each sample, the optical density at 600nm was obtained, as the OD_{600} reading is the preferred sampling method during fermentation for instant monitoring. The data was plotted on a chart to give a corresponding standard curve, and it was found that the OD_{600} was proportionally related to DCW. Using this information, a growth curve for each strain in each yeast extract condition was obtained by plotting the OD_{600} versus time as seen in figure 4-1.

Figure 4-1 Fermentation growth curve of a shaken flask fermentation of two types of *T. fusca.* The error bars are from triplicate measurements with standard error displayed. (A) The slower growing, but higher butyric acid producing muS strain (B) The faster growing, lower butyric acid producing muC strain. DCW stands for dry cell weight.





Fermentation Parameter Growth Plan

Once these initial growth curves were obtained, a plan was determined for the samples that would be used for mRNA analysis, as shown in table 4-1. A number of factors went into determining the conditions for this study. One criterion to be evaluated was the differences in transcription in a differing amount of carbon source, in this case 5g/L and 50g/L. This was chosen to determine the differences in a high and low carbon to nitrogen ratios as the nitrogen sources rates would remain constant throughout. A second variable was the actual type of carbon source used, so each 5g/L and 50g/L condition was performed using either cellobiose or cellulose. A third area of study was how the transcription differs between the types of nitrogen sources. This study also evaluated a purely organic source, yeast extract; a purely inorganic source, ammonia; and a combination of the two. Additionally, table 4-2 shows an additional plan to study the addition of butyric acid to the cells and determine how the strains would respond to butyric acid. The additional acid was either 10 g/L or 15 g/L butyric acid, with samples taken at 0.5 h, 1 h, and 2 h after addition of the acid. Butyric acid was added for this study but, because the chemical compositions are similar, the response will likely also be similar to other organic acids. The constants throughout all experiments are that they were all performed with muC and muS, all were grown in a modified 2382 Hagerdahl medium (with the modifications previously described), and no biotine or thiamine was added to the medium. The situations are summarized in tables 4-1 and 4-2

Table 4-1 Effects of fermentation parameters on gene expression levels for biochemical production. The table shows the sample label that was assigned to each unique condition. The amount of yeast extract used in each is listed in the table, while the NH_4 amount is the standard for Hagerdahl medium.

T. fusca strain	Carbon Source and Amount	Nitrogen Source and Amount		
		Both yeast	Yeast extract (5	NH_4 (+ 0.2 g/L

		extract (5 g/L) and NH_4	g/L)	Yeast extract)
muS	5 g/L cellobiose	SA1	SB1	SC1
	5 g/L cellulose	SA2	SB2	SC2
	50 g/L cellobiose	SA3	SB3	SC3
	50 g/L cellulose	SA4	SB4	SC4
muC	5 g/L cellobiose	CA1	CB1	CC1
	5 g/L cellulose	CA2	CB2	CC2
	50 g/L cellobiose	CA3	CB3	CC3
	50 g/L cellulose	CA4	CB4	CC4

Table 4-2 The table below shows the samples taken to study the effect that organic acid accumulation has on the effect of the cell. The table shows what label was assigned to each unique condition. The cells were grown in a nutrient rich medium, and butyric acid was added to each sample in varying amounts.

Added Butyric Acid Test (5 g/L cellobiose + 5 g/L yeast extract + NH_4)			
T. fusca strain	Time Relative to Acid Addition	Butyric Acid Addition	
		Low Level Butyric Acid (10 g/L)	High Level Butyric Acid (15 g/L)
muS	1 h before	SL-1	SH-1
	0.5 h after	SL.5	SH.5
	1 h after	SL1	SH1
	2 h after	SL2	SH2
muC	1 h before	CL-1	CH-1
	0.5 h after	CL.5	CH.5
	1 h after	CL1	CH1
	2 h after	CL2	CB2

mRNA Isolation

For this study, a ZR Fungal/Bacterial RNA MiniPrep kit was used to isolate the mRNA from *T. fusca.* The first phase of isolation was to prepare the bacteria with a RNAProtect Bacterial Reagent from Qiagen. This reagent is designed to help prevent mRNA degradation that typically occurs during the isolation and transportation stages. It is especially recommended to use this if the cells will be first frozen at -80^{o}C before isolating the RNA, as was done in this protocol.

During isolation, RNase decontamination reagent was constantly used to decontaminate the area, and gloves were constantly changed and sprayed with the decontamination reagent. RNase and DNase free pipette tips and centrifuge tubes were used for all steps. Because RNA is very sensitive to RNase and degrades quickly, these precautions were very important to ensure the isolation occurred in an area as clean of RNase as possible. Once a workstation was decontaminated, the RNAProtect pretreated pellet of *T. fusca* was resuspended in a lysis buffer, and then transferred to a ZR BashingBead Lysis tube. Once the *T. fusca* was in the proper tube and buffer, it was vortexed on high for no less than 5 minutes. This step is critical to the success of the experiment because *T. fusca* is a gram-positive bacteria and the solution alone would not lyse the cell wall. Instead the physical force of the BashingBeads combined with the solution lysed the wall leaving the cellular contents suspended in the buffer.

After the cell was lysed, a series of steps were performed to further remove proteins and DNA from the cell, and leave only mRNA in solution. This was performed in the order provided by the protocol. The process used a series of suspensions of the cellular contents in a provided solution, followed by filtering through provided spin columns. Throughout the procedure, timing played a critical role during the process. The exact times that a spin column needed to be

centrifuged would impact the efficiency of the column, and timing and speed were defined by the protocol. A solution called the DNase I Mastermix had to incubate for at least 15 min to ensure proper removal of DNA from the samples.

The end result was mRNA isolated and suspended in 30 µL of sterile, RNAse free water and immediately sealed and placed on ice to limit degradation. Checking the absorption at 260 nm and performing a calculation to ensure adequate mRNA in the solution performed a simple quality control check. If the check performed in our lab did not show a high enough concentration, then the procedure was repeated using more *T. fusca* until there was an adequate amount. After the quality control check passed, the samples were frozen in an -80^{o}C freezer until time of transport, which was no more than 6 weeks after the first sample was extracted. The transport occurred via overnight shipping to the lab performing the RNASeq test, with the samples kept on dry ice to remain frozen. The receiving lab performed an additional quality control check to ensure the samples had a high enough concentration of mRNA after transport.

The exact steps performed are shown in the full protocol in Appendix B. This entire protocol was repeated for every sample collected.

RNASeq sequencing of mRNA

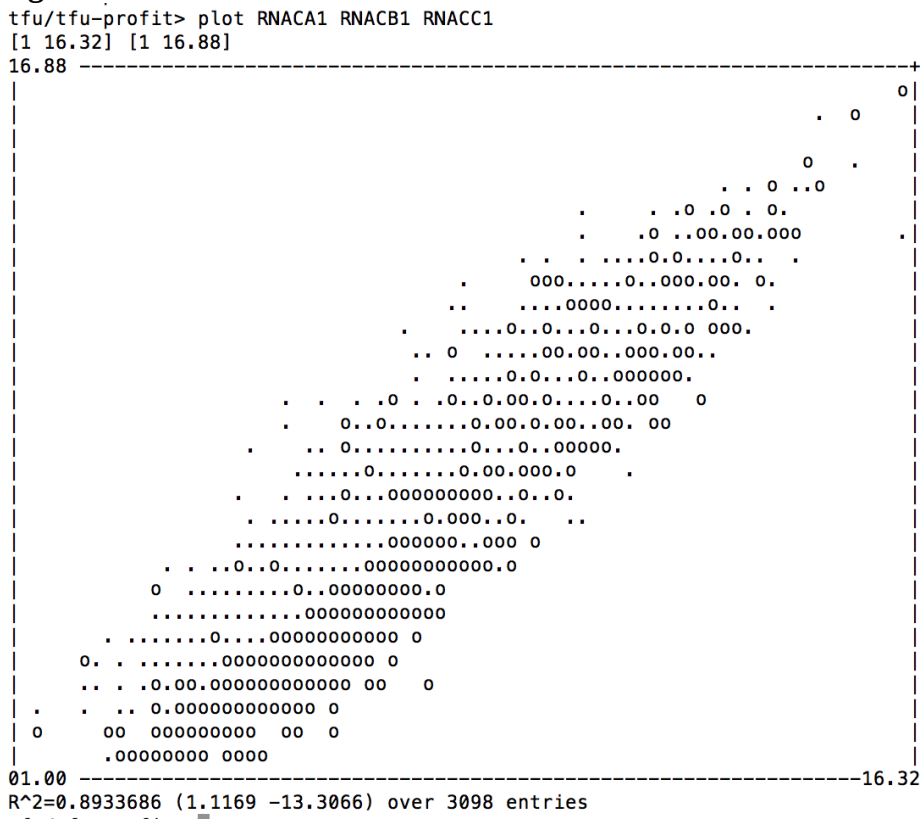
The RNASeq sequencing was performed at the Genome Prairie subsidiary of Genome Canada in Alberta, Canada. The equipment used for sequencing was the Illumina HiSeq 2000 SR100. This sequencing technology could perform an analysis on mRNA as low as 100 ng/µL, but anything below that would not be sequenced.

Analyzing RNASeq Results

The data analysis platform provided by Genome Prairie allows this to be done through porting into a server where the data is stored and analyzing the data through a proprietary

software platform. This service is continually being developed, and they are building into this platform a broad bioinformatics toolset. Figure 4-2 shows one of the most basic comparisons between samples which indicate how well the three samples correlate with each other based on expression levels, and this is also supported by a calculated R^2 value.

Figure 4-2 Correlation among all transcripts when comparing the three different nitrogen sources (CA1, CB1, and CC1). The R^2 value is displayed at the bottom of the chart indicating the overall correlation.



The linear relationship among the samples as shown in figure 4-2, indicates a high correlation among the samples that were plotted. Low correlation samples are not linearly related.

RNASeq Results

Sequenced Samples

Of the 40 samples that were sent to Genome Prairie, 26 samples had sufficient concentration while the others had too low of a concentration. The charts below summarize which samples were analyzed. Table 4-2 is a comparison of different fermentation parameters, while Table 4-4 analyzes the response to butyric acid.

Table 4-3 Summary of sequenced/non-sequenced samples from fermentation parameter study. Green indicates sufficient mRNA was present (>100 ng/µL) for mRNA sequencing. Red indicates the sample did not meet criteria to be sequenced.

T. fusca strain	Carbon Source and Amount	Nitrogen Source		
		Both yeast extract and NH_4	Yeast extract	NH_4
muS	5 g/L cellobiose	SA1	SB1	SC1
	5 g/L cellulose	SA2	SB2	SC2
	50 g/L cellobiose	SA3	SB3	SC3
	50 g/L cellulose	SA4	SB4	SC4
muC	5 g/L cellobiose	CA1	CB1	CC1
	5 g/L cellulose	CA2	CB2	CC2
	50 g/L cellobiose	CA3	CB3	CC3
	50 g/L cellulose	CA4	CB4	CC4

Table 4-4 Summary of sequenced/non-sequenced samples from organic acid accumulation study. Green indicates sufficient mRNA was present (>100 ng/µL) and Genome Prairie was able to sequence the sample. Red indicates the sample did not meet criteria to be sequenced.

T. fusca strain	Time Relative to Acid Addition	Butyric Acid Addition	
		Low Level Butyric Acid	High Level Butyric Acid

		(10 g/L)	(15 g/L)
muS	1 h before	SL-1	SH-1
	0.5 h after	SL.5	SH.5
	1 h after	SL1	SH1
	2 h after	SL2	SH2
muC	1 h before	CL-1	CH-1
	0.5 h after	CL.5	CH.5
	1 h after	CL1	CH1
	2 h after	CL2	CB2

Comparison and Correlation Analysis of Samples Sequenced

The data in this study is intended to be used in one of two ways. The first way is as an as an exploratory study where the data generated in one condition is compared to a different condition. In this exploratory study, a basic correlation could be performed between different conditions to analyze which condition caused the largest variability and deviation in fermentation. This would provide useful information to determine which fermentation conditions would be worth optimizing (large deviation among samples), and which had minimal impact on the cell (small deviation). The genes that are transcribed could also be compared in a differential manner to understand the fundamental differences in expression between the two scenarios.

The table below provides a list of potential studies that would be of interest. This list is a starting point for future comparisons, but the options in the list were determined because only one variable was different between the samples under comparison.

Table 4-5 A listing of possible comparison options using the mRNA sequence results. All samples compared to each other only have one variable difference. These samples could be compared on an exploratory bases to determine correlation between the two, or in a differential manner to determine actual transcription differences.

Sample Comparison	Rationale/Comment about Comparison
SA1 to CA1	Same fermentation parameters but with different species of *T. fusca*
SA3 to CA3	
SB4 to CB4	
SC3 to CC3	
SC4 to CC4	
SA4 to SB4 to SC4	Varying nitrogen source through full range of three options
CA1 to CB1 to CC1	
CA3 to CB3 to CC3	
CA1 to CA3	Different cellobiose to nitrogen ratio
CB1 to CB3	
CC1 to CC3	
SA1 to SA3	
CB2 to CB4	Different cellulose to nitrogen ratio
CC2 to CC4	
SA2 to SA4	
CB1 to CB2	Same nitrogen source but different carbon source (at 5 g/L carbon)
CC1 to CC2	
SA1 to SA2	
CB3 to CB4	Same nitrogen source but different carbon

CC3 to CC4	source (at 50 g/L carbon)
SA3 to SA4	
SC3 to SC4	
SB4 to SC4	Yeast Extract (organic nitrogen source) versus NH_4 (inorganic nitrogen source)
CB1 to CC1	
CB2 to CC2	
CB3 to CC3	
CB4 to CC4	
CL-1 to CL.5 to CL1	Progression of adaption to butyric acid
SL-1 to SL.5 to SL1	
SH-1 to SH.5 to SH1	
SL.5 to SH.5	High and how levels in the same species
SL1 to SH1	
CL.5 to SL.5	Low levels in different species
CL1 to SL1	

As previously mentioned, a number of items could be evaluated at a high level to determine the overall impact on transcription levels the fermentatiopn conditions had on the cells. This can be performed by comparing correlation among the conditions shown in table 4-5 and compare the correlation of the samples. By doing this, it could be determined which fermentation parameters have the largest impact on the expression level of the *T. fusca.* The samples that have the least correlation (low R^2) indicate the largest deviation in gene transcription levels; the opposite is also

true where the smallest deviation (high R^2) indicates the smallest deviation in gene transcription levels.

Table 4-6 List of determined exploratory options based on the RNASeq results. This table has been populated with the data from correlation among samples from the RNASeq database.

Sample Comparison	Sample Correlation	Rationale/Comment about Comparison
SA1 to CA1	0.40	Same fermentation parameters but with different species of *T. fusca*
SA3 to CA3	0.80	
SB4 to CB4	0.70	
SC3 to CC3	0.76	
SC4 to CC4	0.74	
SA4 to SB4 to SC4	0.84	Varying nitrogen source through full range of three options
CA1 to CB1 to CC1	0.89	
CA3 to CB3 to CC3	0.71	
CA1 to CA3	0.41	Different cellobiose to nitrogen ratio
CB1 to CB3	0.63	
CC1 to CC3	0.83	
SA1 to SA3	0.39	
CB2 to CB4	0.72	Different cellulose to nitrogen ratio
CC2 to CC4	0.84	
SA2 to SA4	0.13	
CB1 to CB2	0.83	Same nitrogen source but different

CC1 to CC2	0.75	carbon source (at 5 g/L carbon)
SA1 to SA2	0.58	
CB3 to CB4	0.58	Same nitrogen source but different carbon source (at 50 g/L carbon)
CC3 to CC4	0.78	
SA3 to SA4	0.63	
SC3 to SC4	0.69	
SB4 to SC4	0.74	Yeast extract (organic nitrogen source) versus NH_4 (inorganic nitrogen source)
CB1 to CC1	0.80	
CB2 to CC2	0.63	
CB3 to CC3	0.78	
CB4 to CC4	0.81	
CL-1 to CL.5	0.92	Progression of adaption to butyric acid
CL.5 to CL1	0.99	
SL-1 to SL.5	0.70	
SL.5 to SL1	0.80	
SH-1 to SH.5	0.71	
SH.5 to SH1	0.87	
SL.5 to SH.5	0.85	High and how levels in the same species
SL1 to SH1	0.90	
CL.5 to SL.5	0.31	Low levels in different species
CL1 to SL1	0.71	

The high level exploratory analysis indicates which conditions have the largest impact on overall gene expression deviation. Because of limitations in not having duplicates of samples, a transcription differential analysis between various samples was not performed.

Discussion

An important point regarding the results is that each sample was not performed in duplicate, as should have been for good scientific practice. The decision was made that, based on the goals of the lab, it would be more useful to have a variety of information rather than results that would be scientifically structured enough for a major publication. It was expected that the results would be used as a tool in order to perform future systems biology and metabolic engineering work. Knowing how *T. fusca* responds to various fermentation conditions was deemed more useful to future experiments than knowing 100% certainty on the data obtained. There was a duplicate of each species (CL-1 to CA1 and SL-1 to SA1) to provide the most rudimentary of checks. This decision was made without prior experience in the RNASeq data analysis field. It is recommended that all future studies do a replicate of each experiment, because most of the commands built into the provided bioinformatics platform *require* a replicate in order to plot or process results. While processing could occur through additional statistical platforms, it would be much simpler in the future to perform replicates during the experiment.

While the samples weren't sequenced in duplicate according to good scientific protocol, the fermentation conditions were performed in triplicate to help mitigate risks of an anomaly in the fermentation conditions. These samples were then combined into one master sample of each condition, with the mRNA extraction occurring on the master sample.

When performing the correlation between samples, there are some immediate samples that stand out as a concern, because of extremely high or extremely low correlation between the two. These samples would likely be discarded or need further data filtering to ensure that these truly are real values.

One interesting observation is the reaction of both species of *T. fusca* to additional butyric acid, either the high or low levels. It appears that both species initially have a change in expression between the pre-butyric acid sample and the sample that was taken 0.5 h after addition. But the samples have a very high correlation between the 0.5 h after and the 1 hour after addition of butyric acid. This follows in line with the logic that butyric acid is extremely difficult for the cells, and causes high stress in a short period of time. Once the stress is induced on the cell, then it continues to try and produce the same cellular machinery to keep it alive, and the expression doesn't change after the initial shock. This information is important when considering the amount of butyric acid that can be produced by *T. fusca*. If the butyric acid is not removed from the fermentation broth, it will have a negative impact on the cell expression levels. Butyric acid, in this study, was added suddenly which is why the expression changed suddenly. By doing this, we can see the pathways most impacted by the butyric acid, and could possibly use this as the detailed expression analysis in the future. It would be useful to study the differential expression between both species when the butyric acid shock occurred and examine if any future genetic engineering efforts could help improve the cells tolerance to butyric acid.

When grown on the same fermentation parameters, it was obvious that the two different lab evolved species of *T. fusca* expressed different genes. The one sample of 40% correlation is likely an outlier and caused by an error, and that other samples correlation was likely accurate. This correlation falls in line with what would be expected, with the knowledge of how the two

were developed. Because they were laboratory evolved species over an extended period of time, they likely evolved to express different genes based on the growth conditions.

In terms of variability of the fermentation parameters, it appears that the carbon to nitrogen has a larger impact on fermentation conditions than does the specific type of nitrogen source. The varying nitrogen sources all correlated in the 70-80% range, which indicated a strong correlation. However, the carbon to nitrogen ration ranged from 13% correlation up to 84% correlation. Even if those two values are outliers, there is still an extreme variability in correlation when comparing cellobiose to nitrogen ratio. In the context of the study discussed in the previous chapter, *T. fusca* stopped consuming carbon, but when nitrogen was added the fermentation resumed. The RNASeq result supports that observation, because the cells themselves express markedly different genes depending on the carbon to nitrogen ratio.

Another conclusion from this study is that there is an extremely large amount of data generated as a result of RNASeq analysis. The Genome Prairie lab presented the results in two ways. The first was by delivering a .fastq file, which was the raw output of the RNASeq process and would need further processing through a software package like Tophat. The second way that the lab presented the data was in a form where their lab did the processing. As part of this processing, they aligned the reads with their source gene and counted up the number of hits per gene. They then created a non-redundant library from this data. Genome Prairie offered the ability to port into their network using a secure shell connection and analyze the data using a command-based system called lobe, which contained a wide array of bioinformatics tools.

While this information can be useful for a correlation among different samples or by confirmation of gene expression, it was discovered that there are certain inherent limitations with RNASeq results. The most challenging issue is the variability in gene expression in a cell.

Because the gene expression varies greatly depending on the phase of growth, the timing of sample collection is critical. Gene expression is most abundant during the growth phase, but depending on the contents of the fermentation broth, the growth phase could occur at anywhere from 18 h to 72 h. Before a future RNASeq analysis is performed, a detailed growth curve on each broth should be determined, which was not the case in this study.

A second issue with an RNASeq study is that the fermentation parameters, including the nutrients, shaking speed, temperature, and numerous other things will largely impact the gene expression and subsequent metabolite production. This study specifically evaluated the difference between cellulose and cellobiose, organic and inorganic nitrogen sources, as well as inhibition to butyric acid. While this is useful information, even more useful information could have been obtained had the ideal carbon to nitrogen ratio been obtained prior, or determined if the ideal organic nitrogen source was yeast extract or peptone or some other organic source. All of these factors could have changed the results of gene expression considerably. Therefore, it makes sense that other studies be performed at later stages after the fermentation conditions are more fully evaluated for *T. fusca*.

Finally, a limitation based on RNA degradation is the most fundamental challenge in RNASeq technology. Isolation of RNA is very challenging and RNA degrades very quickly. The area must be kept complete clean of RNase. It must remain cold during isolation, and be frozen immediately at -80^{o}C. During transport, it must not thaw out or all RNA will be degraded. During this study, the lab performing the RNASeq analysis first performed a quality control check to evaluate the amount of mRNA present in the sample to ensure it was of sufficient quantity. Several of the samples did not pass this check, and Genome Prairie allowed one additional submission of those samples to see if the expression could be increased. Even

after the second round of samples, there were still samples that did not have enough mRNA for analysis. Because of this degradation, it is difficult to make quantitative gene expression evaluations based on quantity. The RNASeq data can prove a gene is expressed, but not how much of that gene is expressed, since it could have degraded. Additionally, it can't prove that a gene is not expressed either, because the gene could be initially expressed in a very small amount, but that small amount degraded during transport.

While improving the plan for RNA isolation can mitigate the first two limitations discussed, the degradation of mRNA will not be improved through better planning. Because mRNA is so sensitive, unless a fully automated system can perform the isolation as well as RNASeq immediately upon taking a sample, then degradation of a certain amount of mRNA will likely occur.

Chapter 5 - Conclusion and Future Work

In Chapter 1, the need and desires to utilize biomass were introduced. The recalcitrant structure and difficulties in utilizing this structure for fermentation systems were discussed. Numerous methods of biomass pretreatment were evaluated based on their current state. Finally, the chapter introduced the use of enzymes to break down the pretreated biomass into useful sugars for fermentation. This was discussed in the context of either enzymatic hydrolysis or consolidated bioprocessing.

The background information from Chapter 1 was expanded upon further in the specific context of utilizing biomass as a feedstock for *Thermobifida fusca.* Chapter 2 discussed previous work using *T. fusca* and provided further context for the work performed. The laboratory evolved *T. fusca* that was discussed in Chapter 2 was the same bacteria used for this study, so a background on how it was developed is useful for context in future chapters.

For the original work of this study, Chapter 3 discussed the fermentation conditions that were optimized to produce butyric acid using *T. fusca.* The best stir speed and aeration combination was found to be 400 rpm and 2 vvm and a yield of 0.52 g/g-C. A batch fermentation starting with a media containing 9.66 g/L of cellulose produced a maximum titer of 2.1 g/L butyric acid as well as 2.37 g/L of butyric acid using a fed-batch fermentation with corn stover as a substrate. Another important result was that it was found that *T. fusca* could produce butyric acid from lignocellulosic biomass (corn stover). These results of 2.37 g/L of butyric acid produced are relatively low compared to other studies; however, no other studies produced butyric acid direct from cellulose in a CBP scheme. Therefore, this result is extremely promising in the context of producing biochemicals directly from biomass.

Optimization of the fermentation parameters provides a basis for production of butyric acid from *T. fusca*. The optimized parameters can be expanded upon in future work by further defining necessary nutrients in the fermentation broth, by testing fermentation schemes such as fed-batch or continuous fermentation, by evaluating impacts of pH on fermentation, or any other number of fermentation and process engineering work. Also, other organism development could be tested in the previously discussed optimized conditions to understand if changes made to the organism are beneficial to the cell or not.

It was shown that *T. fusca* prefers complex carbon sources for fermentation, since glucose was the last carbon source to be used by the cell. Also growth on pure glucose yielded very little butyric acid compared to complex carbon sources. Because of these two observations, a conclusion could be drawn that *T. fusca* could be combined in a co-culture or continuous fermentation system where *T. fusca* breaks down complex carbon sources while producing both glucose and a value added chemical such as butyric acid. Meanwhile, a second organism (such as *S. cerevisiae)* could use the glucose produced by *T. fusca* as its primary feedstock. Because *T. fusca* does not selectively use glucose, while most organisms do, this would be a symbiotic relationship. Inherent limitations to this proposed co-culture scheme do exist. The limitations include finding a second organism that favors similar fermentation conditions such as temperature and pH, maintaining a proper carbon to nitrogen ratio for each organism, as well as preventing either organism from being overtaking the other in the culture. However, it would be worth investigating as a possible future study.

The cellulase production in *T. fusca* was found to have a negative impact on butyric acid production. This negative impact is expected because when *T. fusca* is producing cellulases and breaking down complex carbons, a lot of the energy and carbon is being re-directed to producing

the necessary machinery to do so. This redirects it away from producing the enzymes necessary to produce butyric acid. If using a lignocellulosic biomass source, it is recommended that some pretreatment occur to liberate monomeric sugars. It was proven that *T. fusca* could both grow and produce butyric acid on a fermentation of milled corn stover, the fermentation took a very long time compared to a tradition fermentation of cellulose or cellobiose (>20 h longer). Future work in this field could continue to use corn stover as a feedstock, but to change pretreatment methods to evaluate the impacts of pretreatment on fermentation with *T. fusca*.

The metabolic pathway was determined for butyric acid production in *T. fusca* and was also evaluated experimentally by determining the mRNA expression of the genes whose proteins catalyze the necessary reactions. The final step did not have a specific gene determined from the *T. fusca* genome, but a washed cell culture was performed to determine if butanoyl-CoA was first converted to butanoyl-P and then to butyric acid or if butanoyl-CoA was converted directly to butyric acid through a butyryl-CoA transferase. The culture indicated that no butyrate kinase was present (no butyric acid was produced from butanoyl-P as carbon source), whereas butyric acid was produced from butanoyl-CoA as a carbon source. An assay was then performed to determine the activity of the butyryl-CoA transferase. This work is important to understand as a baseline for future metabolic engineering of *T. fusca* for butyric acid production. Genes in this pathway could be overexpressed to see the impacts on butyric acid production. Additionally, this gives information to use when evaluating RNASeq data from various conditions.

An exploratory, correlation based study was performed on the samples that were analyzed. This initial evaluation determined that the carbon/nitrogen ratio has the largest overall impact on changes in gene transcription, which is consistent with the observation of *T. fusca* halting carbon consumption until additional nitrogen was added. A future analysis could

determine specifics of the cause of this, and fermentation could occur to determine this optimum. Once the optimum is determined, a fed-batch system could maintain the optimum ratio during fermentation to maximize the titer.

Another conclusion from the RNASeq exploratory analysis was that the addition of butyric acid had a negative impact on the cell. This negative impact was indicated by a low correlation between the genes expressed before butyric acid was added, and the expression 30 minutes after the butyric acid was added. Cells did not adapt to butyric acid, as seen by the high correlation of all samples after the addition of butyric acid. The correlation results also showed that the two different levels of butyric acid used in this study both caused detrimental effects to *T. fusca*. Finally, the correlation also showed the both the muC and muS strain of *T. fusca* had similar detrimental reactions to the butyric acid addition. These results show the extreme impact that butyric acid has on *T. fusca* cells. Examining specific transcription patterns in these scenarios could provide context on why *T. fusca* is so greatly impacted by butyric acid.

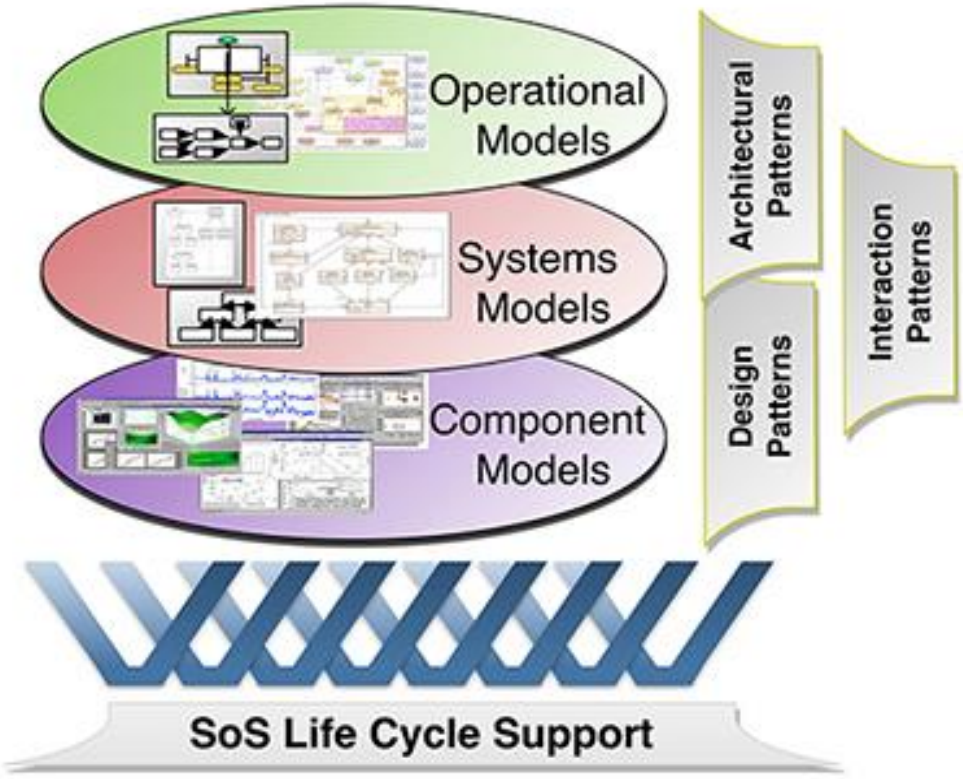
The data obtained during the RNASeq results is expected to be a tool for future studies, and not as a sole platform for scientific discoveries. The reasoning is that the samples were not performed in duplicate for RNASeq like the scientific standards require. It is impossible to truly perform a good solid comparison if the data point is skewed, and the chances are higher of that occurring when there is only one sample. However, this data is not useless, and should be used as described above – as potential ideas for future studies. It also can be used to determine if a gene from the *T. fusca* transcript is expressed or not during fermentation.

A conclusion, and a realization, drawn from the RNASeq results is that there is much future work needed in the "omics" field of study. The amount of data generated as a result of these studies is enormous (excel file of >2MB), and is very difficult to analyze. It takes a lot of

statistical techniques to evaluate. Mining this data and applying for practical genetic engineering efforts will also be best analyzed in a "systems biology" role where the genome, transcripts, proteomics, and metabolomics are all analyzed as a system rather than individual components.

This "systems biology" platform will be similar to the "Systems of Systems" (a.k.a. "SoS") lifestyle methodology used in product development and big data analysis. The "Designing for Adaptability and evolution in System of Systems Engineering" (a.k.a. "DANSE") project consortium is a consortium that aims to improve design and management of these complex systems and provides methodologies for these complex systems [129]. The overall goal of "SoS" engineering is to interatively improve and model how complex systems interact with each other, with the ultimately goal of optimizing the system of systems [129]. This "SoS" methodology is exhibited in the image below.

Figure 5-1 High level representation of "Systems of Systems" methodology that could potentially be applied to "Systems Biology" and other approaches in the future to improve upon the work of this study. Used with permission from [129].



These "System of Systems" can be broken down into the areas explored in this study. The "Operational Models" are the fermentation parameters study that was explored in the butyric acid production. The "Component Models" could be described as the component level details of *T. fusca*, specifically transcriptomics. It can be seen that these two models interact in the "Systems Models" and this is where the field of "Systems Biology" is located. On the bottom of the chart, is a bar titled "SoS Life Cycle Support" which indicates the evolution of the system. Each study evolves the system a little further, and they all interconnect. With the ever increasing complexity of data generated, this DANSE model or similar systems biology models will be pivotal for design of fermentation platforms and organism development.

Much work still needs to be performed to improve the fermentation conditions and to use the RNASeq results in this forward looking "System of Systems" manner. However, the results of this study prove that production of butyric acid directly from lignocellulosic biomass with *T. fusca* is possible. This study was the first example of this system and it worth exploring through future studies to further improve the technology. This study provided fermentation conditions that would produce butyric acid, and provides a point for future iteration of this system. Production of value added chemicals from a renewable feedstock like biomass is an important future need, and *Thermobifida fusca* could help provide information to fill this need.

References

[1] Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels: a comprehensive review. *Renewable and Sustainable Energy Reviews*, *14*(2), 578-597.

[2] Perlack, R. D., Eaton, L. M., Turhollow Jr, A. F., Langholtz, M. H., Brandt, C. C., Downing, M. E., ... & Lightle, D. (2011). US billion-ton update: biomass supply for a bioenergy and bioproducts industry.

[3] McCann, M. C., Buckeridge, M. S., & Carpita, N. C. (2014). *Plants and bioenergy*. Springer..

[4] Chundawat, S. P., Beckham, G. T., Himmel, M. E., & Dale, B. E. (2011). Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annual review of chemical and biomolecular engineering*, *2*, 121-145.

[5] Pauly, M., & Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal*, *54*(4), 559-568.

[6] McKendry, P. (2002). Energy production from biomass (part 1): overview of biomass. *Bioresource technology*, *83*(1), 37-46.

[7] Biomass Program: Feedstock Composition and Property Database. (n.d.). Retrieved September 30, 2014, from http://www.afdc.energy.gov/biomass/progs/search3.cgi?13124

[8] Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M., & Ladisch, M. (2005). Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource technology*, *96*(6), 673-686..

[9] Alvira, P., Tomás-Pejó, E., Ballesteros, M., & Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresource technology*, *101*(13), 4851-4861.

[10] Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource technology*, *83*(1), 1-11.

[11] Taherzadeh, M. J., & Karimi, K. (2008). Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. *International journal of molecular sciences*, *9*(9), 1621-1651.

[12] Cadoche, L., & López, G. D. (1989). Assessment of size reduction as a preliminary step in the production of ethanol from lignocellulosic wastes.*Biological Wastes*, *30*(2), 153-157.

[13] Lamsal, B., Yoo, J., Brijwani, K., & Alavi, S. (2010). Extrusion as a thermo-mechanical pre-treatment for lignocellulosic ethanol. *biomass and bioenergy*,*34*(12), 1703-1710.

[14] Yoo, J., Alavi, S., Vadlani, P., & Amanor-Boadu, V. (2011). Thermo-mechanical extrusion pretreatment for conversion of soybean hulls to fermentable sugars. *Bioresource technology*, *102*(16), 7583-7590.

[15] Gáspár, M., Kálmán, G., & Réczey, K. (2007). Corn fiber as a raw material for hemicellulose and ethanol production. *Process Biochemistry*, *42*(7), 1135-1139.

[16] Kumar, P., Barrett, D. M., Delwiche, M. J., & Stroeve, P. (2009). Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial & Engineering Chemistry Research*, *48*(8), 3713-3729.

[17] Taherzadeh, M. J., & Karimi, K. (2007). Enzymatic-based hydrolysis processes for ethanol. *BioResources*, *2*(4), 707-738.

[18] Silverstein, R. A., Chen, Y., Sharma-Shivappa, R. R., Boyette, M. D., & Osborne, J. (2007). A comparison of chemical pretreatment methods for improving saccharification of cotton stalks. *Bioresource Technology*, *98*(16), 3000-3011.

[19] Kim, S., & Holtzapple, M. T. (2006). Delignification kinetics of corn stover in lime pretreatment. *Bioresource Technology*, *97*(5), 778-785.

[20] Chang, V. S., Nagwani, M., Kim, C. H., & Holtzapple, M. T. (2001). Oxidative lime pretreatment of high-lignin biomass. *Applied Biochemistry and Biotechnology*, *94*(1), 1-28.

[21] Wyman, C. E., Dale, B. E., Elander, R. T., Holtzapple, M., Ladisch, M. R., & Lee, Y. Y. (2005). Coordinated development of leading biomass pretreatment technologies. *Bioresource technology*, *96*(18), 1959-1966.

[22] Yang, B., & Wyman, C. E. (2004). Effect of xylan and lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose. *Biotechnology and bioengineering*, *86*(1), 88-98.

[23] Cara, C., Ruiz, E., Oliva, J. M., Sáez, F., & Castro, E. (2008). Conversion of olive tree biomass into fermentable sugars by dilute acid pretreatment and enzymatic saccharification. *Bioresource Technology*, *99*(6), 1869-1876.

[24] Taherzadeh, M. J. (1999). *Ethanol from lignocellulose: physiological effects of inhibitors and fermentation strategies*. Chalmers University of Technology.

[25] Kootstra, A. M. J., Beeftink, H. H., Scott, E. L., & Sanders, J. P. (2009). Comparison of dilute mineral and organic acid pretreatment for enzymatic hydrolysis of wheat straw. *Biochemical Engineering Journal*, *46*(2), 126-131.

[26] Zhao, X., Cheng, K., & Liu, D. (2009). Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Applied microbiology and biotechnology*, *82*(5), 815-827.

[27] Papatheofanous, M. G., Billa, E., Koullas, D. P., Monties, B., & Koukios, E. G. (1995). Two-stage acid-catalyzed fractionation of lignocellulosic biomass in aqueous ethanol systems at low temperatures. *Bioresource Technology*, *54*(3), 305-310.

[28] Samayam, I. P., & Schall, C. A. (2010). Saccharification of ionic liquid pretreated biomass with commercial enzyme mixtures. *Bioresource Technology*, *101*(10), 3561-3566.

[29] Zhu, S., Wu, Y., Chen, Q., Yu, Z., Wang, C., Jin, S., ... & Wu, G. (2006). Dissolution of cellulose with ionic liquids and its application: a mini-review.*Green Chem.*, *8*(4), 325-327.

[30] Brodeur, G., Yau, E., Badal, K., Collier, J., Ramachandran, K. B., & Ramakrishnan, S. (2011). Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. *Enzyme research*, *2011*.

[31] Pu, Y., Jiang, N., & Ragauskas, A. J. (2007). Ionic liquid as a green solvent for lignin. *Journal of Wood Chemistry and Technology*, *27*(1), 23-33.

[32] Binder, J. B., & Raines, R. T. (2009). Simple chemical transformation of lignocellulosic biomass into furans for fuels and chemicals. *Journal of the American Chemical Society*, *131*(5), 1979-1985.

[33] Lee, S. H., Doherty, T. V., Linhardt, R. J., & Dordick, J. S. (2009). Ionic liquid-mediated selective extraction of lignin from wood leading to enhanced enzymatic cellulose hydrolysis. *Biotechnology and Bioengineering*, *102*(5), 1368-1376.

[34] Swatloski, R. P., Spear, S. K., Holbrey, J. D., & Rogers, R. D. (2002). Dissolution of cellose with ionic liquids. *Journal of the American Chemical Society*, *124*(18), 4974-4975.

[35] Socha, A. M., Parthasarathi, R., Shi, J., Pattathil, S., Whyte, D., Bergeron, M., ... & Singh, S. (2014). Efficient biomass pretreatment using ionic liquids derived from lignin and hemicellulose. *Proceedings of the National Academy of Sciences*, *111*(35), E3587-E3595.

[36] Overend, R. P., Chornet, E., & Gascoigne, J. A. (1987). Fractionation of lignocellulosics by steam-aqueous pretreatments [and discussion].*Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *321*(1561), 523-536.

[37] Ballesteros, I., Negro, M. J., Oliva, J. M., Cabañas, A., Manzanares, P., & Ballesteros, M. (2006, January). Ethanol production from steam-explosion pretreated wheat straw. In *Twenty-Seventh Symposium on Biotechnology for Fuels and Chemicals* (pp. 496-508). Humana Press.

[38] Yu, Q., Zhuang, X., Yuan, Z., Wang, Q., Qi, W., Wang, W., ... & Xu, H. (2010). Two-step liquid hot water pretreatment of Eucalyptus grandis to enhance sugar recovery and enzymatic digestibility of cellulose. *Bioresource Technology*,*101*(13), 4895-4899.

[39] Mosier, N., Hendrickson, R., Ho, N., Sedlak, M., & Ladisch, M. R. (2005). Optimization of pH controlled liquid hot water pretreatment of corn stover.*Bioresource technology*, *96*(18), 1986-1993.

[40] Laser, M., Schulman, D., Allen, S. G., Lichwa, J., Antal, M. J., & Lynd, L. R. (2002). A comparison of liquid hot water and steam pretreatments of sugar cane bagasse for bioconversion to ethanol. *Bioresource technology*, *81*(1), 33-44.

[41] Alizadeh, H., Teymouri, F., Gilbert, T. I., & Dale, B. E. (2005). Pretreatment of switchgrass by ammonia fiber explosion (AFEX). *Applied biochemistry and biotechnology*, *124*(1-3), 1133-1141.

[42] Laureano-Perez, L., Teymouri, F., Alizadeh, H., & Dale, B. E. (2005). Understanding factors that limit enzymatic hydrolysis of biomass. *Applied Biochemistry and Biotechnology*, *124*(1-3), 1081-1099.

[43] Kim, K. H., & Hong, J. (2001). Supercritical CO 2 pretreatment of lignocellulose enhances enzymatic cellulose hydrolysis. *Bioresource Technology*, *77*(2), 139-144.

[44] Taherzadeh, M. J., & Karimi, K. (2008). Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. *International journal of molecular sciences*, *9*(9), 1621-1651.

[45] Dumitriu, S. (Ed.). (2004). *Polysaccharides: structural diversity and functional versatility*. CRC press.

[46] Martín, C., Thomsen, M. H., Hauggaard-Nielsen, H., & BelindaThomsen, A. (2008). Wet oxidation pretreatment, enzymatic hydrolysis and simultaneous saccharification and fermentation of clover–ryegrass mixtures. *Bioresource technology*, *99*(18), 8777-8782.

[47] Keshwani, D. R. (2009). *Microwave pretreatment of switchgrass for bioethanol production*. ProQuest.

[48] Hasunuma, T., Okazaki, F., Okai, N., Hara, K. Y., Ishii, J., & Kondo, A. (2013). A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology.*Bioresource technology*, *135*, 513-522.

[49] Lynd, L. R., Weimer, P. J., Van Zyl, W. H., & Pretorius, I. S. (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and molecular biology reviews*, *66*(3), 506-577.

[50] Hyeon, J. E., Jeon, S. D., & Han, S. O. (2013). Cellulosome-based, Clostridium-derived multi-functional enzyme complexes for advanced biotechnology tool development: advances and applications. *Biotechnology advances*, *31*(6), 936-944..

[51] van den Brink, J., & de Vries, R. P. (2011). Fungal enzyme sets for plant polysaccharide degradation. *Applied microbiology and biotechnology*, *91*(6), 1477-1492.

[52] Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Plant Biology*, *61*(1), 263.

[53] Sweeney, M. D., & Xu, F. (2012). Biomass converting enzymes as industrial biocatalysts for fuels and chemicals: recent developments. *Catalysts*, *2*(2), 244-263.

[54] Kumar, R., Singh, S., & Singh, O. V. (2008). Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of industrial microbiology & biotechnology*, *35*(5), 377-391.

[55] Gao, D., Uppugundla, N., Chundawat, S. P., Yu, X., Hermanson, S., Gowda, K., ... & Dale, B. E. (2011). Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnol Biofuels*,*4*(5), 1-11.

[56] Couturier, M., Haon, M., Coutinho, P. M., Henrissat, B., Lesage-Meessen, L., & Berrin, J. G. (2011). Podospora anserina hemicellulases potentiate the Trichoderma reesei secretome for saccharification of lignocellulosic biomass.*Applied and environmental microbiology*, *77*(1), 237-246.

[57] Kumar, R., & Wyman, C. E. (2009). Effect of xylanase supplementation of cellulase on digestion of corn stover solids prepared by leading pretreatment technologies. *Bioresource Technology*, *100*(18), 4203-4213.

[58] Banerjee, G., Scott-Craig, J. S., & Walton, J. D. (2010). Improving enzymes for biomass conversion: a basic research perspective. *Bioenergy research*, *3*(1), 82-92.

[59] Banerjee, G., Car, S., Scott-Craig, J. S., Borrusch, M. S., & Walton, J. D. (2010). Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnology for biofuels*, *3*(1), 22.

[60] Griggs, A. J., Stickel, J. J., & Lischeske, J. J. (2012). A mechanistic model for enzymatic saccharification of cellulose using continuous distribution kinetics I: depolymerization by EGI and CBHI. *Biotechnology and bioengineering*, *109*(3), 665-675.

[61] Levine, S. E., Fox, J. M., Blanch, H. W., & Clark, D. S. (2010). A mechanistic model of the enzymatic hydrolysis of cellulose. *Biotechnology and bioengineering*, *107*(1), 37-51.

[62] Okazaki, M., & Moo-Young, M. (1978). Kinetics of enzymatic hydrolysis of cellulose: analytical description of a mechanistic model. *Biotechnology and bioengineering*, *20*(5), 637-663.

[63] Bansal, P., Hall, M., Realff, M. J., Lee, J. H., & Bommarius, A. S. (2009). Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnology advances*, *27*(6), 833-848.

[64] Belkacemi, K., & Hamoudi, S. (2003). Enzymatic hydrolysis of dissolved corn stalk hemicelluloses: reaction kinetics and modeling. *Journal of chemical technology and biotechnology*, *78*(7), 802-808.

[65] Feng, Y. Y., He, Z. M., Song, L. F., Ong, S. L., Hu, J. Y., Zhang, Z. G., & Ng, W. J. (2003). Kinetics of β-mannanase fermentation by Bacillus licheniformis.*Biotechnology letters*, *25*(14), 1143-1146.

[66] HarjunpÄÄ, V., Teleman, A., Siika-Aho, M., & Drakenberg, T. (1995). Kinetic and Stereochemical Studies of Manno-Oligosaccharide Hydrolysis Catalysed by β-Mannanases from Trichoderma Reesei. *European journal of biochemistry*,*234*(1), 278-283.

[67] Liu, X., Lu, M., Ai, N., Yu, F., & Ji, J. (2012). Kinetic model analysis of dilute sulfuric acid-catalyzed hemicellulose hydrolysis in sweet sorghum bagasse for xylose production. *Industrial Crops and Products*, *38*, 81-86.

[68] Zhang, Y., Xu, B., & Zhou, W. (2014). On a novel mechanistic model for simultaneous enzymatic hydrolysis of cellulose and hemicellulose considering morphology. *Biotechnology and bioengineering*, *111*(9), 1767-1781.

[69] Chiaramonti, D., Prussi, M., Ferrero, S., Oriani, L., Ottonello, P., Torre, P., & Cherchi, F. (2012). Review of pretreatment processes for lignocellulosic ethanol production, and development of an innovative method. *Biomass and Bioenergy*,*46*, 25-35.

[70] Olson, D. G., McBride, J. E., Shaw, A. J., & Lynd, L. R. (2012). Recent progress in consolidated bioprocessing. *Current opinion in biotechnology*, *23*(3), 396-405.

[71] Lynd, L. R., Van Zyl, W. H., McBride, J. E., & Laser, M. (2005). Consolidated bioprocessing of cellulosic biomass: an update. *Current opinion in biotechnology*, *16*(5), 577-583.

[72] Koeck, D. E., Pechtl, A., Zverlov, V. V., & Schwarz, W. H. (2014). Genomics of cellulolytic bacteria. *Current opinion in biotechnology*, *29*, 171-183.

[73] Deng, Y., & Fong, S. S. (2010). Development and application of a PCR-targeted gene disruption method for studying CelR function in Thermobifida fusca. *Applied and environmental microbiology*, *76*(7), 2098-2106.

[74] Deng, Y., & Fong, S. S. (2011). Metabolic engineering of Thermobifida fusca for direct aerobic bioconversion of untreated lignocellulosic biomass to 1-propanol. *Metabolic engineering*, *13*(5), 570-577.

[75] Diaz-Ricci, J. C., Regan, L., & Bailey, J. E. (1991). Effect of alteration of the acetic acid synthesis pathway on the fermentation pattern of Escherichia coli.*Biotechnology and bioengineering*, *38*(11), 1318-1324.

[76] Ingram, L. O., & Clark, D. P. (1991). *U.S. Patent No. 5,028,539*. Washington, DC: U.S. Patent and Trademark Office.

[77] Hols, P., Ramos, A., Hugenholtz, J., Delcour, J., de Vos, W. M., Santos, H., & Kleerebezem, M. (1999). Acetate utilization in Lactococcus lactis deficient in lactate dehydrogenase: a rescue pathway for maintaining redox balance.*Journal of bacteriology*, *181*(17), 5521-5526.

[78] Desai, R. P., & Papoutsakis, E. T. (1999). Antisense RNA strategies for metabolic engineering of Clostridium acetobutylicum. *Applied and Environmental Microbiology*, *65*(3), 936-945.

[79] Burgard, A. P., & Maranas, C. D. (2003). Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnology and bioengineering*, *82*(6), 670-677.

[80] Herrgård, M. J., Fong, S. S., & Palsson, B. Ø. (2006). Identification of genome-scale metabolic network models using experimentally measured flux profiles.

[81] Milne, C. B., Eddy, J. A., Raju, R., Ardekani, S., Kim, P. J., Senger, R. S., ... & Price, N. D. (2011). Metabolic network reconstruction and genome-scale model of butanol-producing strain Clostridium beijerinckii NCIMB 8052. *BMC systems biology*, *5*(1), 130.

[82] Feist, A. M., Zielinski, D. C., Orth, J. D., Schellenberger, J., Herrgard, M. J., & Palsson, B. Ø. (2010). Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. *Metabolic engineering*, *12*(3), 173-186.

[83] Tyurin, M. V., Desai, S. G., & Lynd, L. R. (2004). Electrotransformation of Clostridium thermocellum. *Applied and environmental microbiology*, *70*(2), 883-890.

[84] Tripathi, S. A., Olson, D. G., Argyros, D. A., Miller, B. B., Barrett, T. F., Murphy, D. M., ... & Caiazza, N. C. (2010). Development of pyrF-based genetic system for targeted gene deletion in Clostridium thermocellum and creation of a pta mutant. *Applied and environmental microbiology*, *76*(19), 6591-6599.

[85] Ingram, L. O. (1989). Ethanol tolerance in bacteria. *Critical reviews in biotechnology*, *9*(4), 305-319..

[86] Jung, S., Zeikus, J. G., & Hollingsworth, R. I. (1994). A new family of very long chain alpha, omega-dicarboxylic acids is a major structural fatty acyl component of the membrane lipids of Thermoanaerobacter ethanolicus 39E.*Journal of lipid research*, *35*(6), 1057-1065.

[87] Herrero, A. A., Gomez, R. F., & Roberts, M. F. (1982). Ethanol-induced changes in the membrane lipid composition of Clostridium thermocellum.*Biochimica et Biophysica Acta (BBA)-Biomembranes*, *693*(1), 195-204.

[88] Herrero, A. A., Gomez, R. F., & Roberts, M. F. (1985). 31P NMR studies of Clostridium thermocellum. Mechanism of end product inhibition by ethanol.*Journal of Biological Chemistry*, *260*(12), 7442-7451.

[89] Herrero, A. A., Gomez, R. F., Snedecor, B., Tolman, C. J., & Roberts, M. F. (1985). Growth inhibition of Clostridium thermocellum by carboxylic acids: a mechanism based on uncoupling by weak acids. *Applied microbiology and biotechnology*, *22*(1), 53-62.

[90] Herrero, A. A., & Gomez, R. F. (1980). Development of ethanol tolerance in Clostridium thermocellum: effect of growth temperature. *Applied and environmental microbiology*, *40*(3), 571-577.

[91] Kundu, S., Ghose, T. K., & Mukhopadhyay, S. N. (1983). Bioconversion of cellulose into ethanol by Clostridium thermocellum—product inhibition.*Biotechnology and bioengineering*, *25*(4), 1109-1126.

[92] Brown, S. D., Guss, A. M., Karpinets, T. V., Parks, J. M., Smolin, N., Yang, S., ... & Lynd, L. R. (2011). Mutant alcohol dehydrogenase leads to improved ethanol tolerance in Clostridium thermocellum. *Proceedings of the National Academy of Sciences*, *108*(33), 13752-13757.

[93] Williams, T. I., Combs, J. C., Lynn, B. C., & Strobel, H. J. (2007). Proteomic profile changes in membranes of ethanol-tolerant Clostridium thermocellum.*Applied microbiology and biotechnology*, *74*(2), 422-432.

[94] Demain, A. L., Newcomb, M., & Wu, J. D. (2005). Cellulase, clostridia, and ethanol. *Microbiology and molecular biology reviews*, *69*(1), 124-154.

[95] Lynd, L. R., Grethlein, H. E., & Wolkin, R. H. (1989). Fermentation of cellulosic substrates in batch and continuous culture by Clostridium thermocellum.*Applied and Environmental Microbiology*, *55*(12), 3131-3139.

[96] Saddler, J. N., & Chan, M. H. (1984). Conversion of pretreated lignocellulosic substrates to ethanol by Clostridium thermocellum in mono-and co-culture with Clostridium thermosaccharolyticum and Clostridium thermohydrosulphuricum.*Canadian journal of microbiology*, *30*(2), 212-220.

[97] Hörmeyer, H. F., Tailliez, P., Millet, J., Girard, H., Bonn, G., Bobleter, O., & Aubert, J. P. (1988). Ethanol production by Clostridium thermocellum grown on hydrothermally and organosolv-pretreated lignocellulosic materials. *Applied microbiology and biotechnology*, *29*(6), 528-535.

[98] Avgerinos, G. C., & Wang, D. C. (1983). Selective solvent delignification for fermentation enhancement. *Biotechnology and bioengineering*, *25*(1), 67-83.

[99] Avgerinos, G. C., & Wang, D. C. (1983). Selective solvent delignification for fermentation enhancement. *Biotechnology and bioengineering*, *25*(1), 67-83.

[100] Herrero, A. A., Gomez, R. F., Snedecor, B., Tolman, C. J., & Roberts, M. F. (1985). Growth inhibition of Clostridium thermocellum by carboxylic acids: a mechanism based on uncoupling by weak acids. *Applied microbiology and biotechnology*, *22*(1), 53-62.

[101] Russell, J. B. (1992). Another explanation for the toxicity of fermentation acids at low pH: anion accumulation versus uncoupling. *Journal of Applied Bacteriology*, *73*(5), 363-370.

[102] Wang, X., Miller, E. N., Yomano, L. P., Zhang, X., Shanmugam, K. T., & Ingram, L. O. (2011). Increased furfural tolerance due to overexpression of NADH-dependent oxidoreductase FucO in Escherichia coli strains engineered for the production of ethanol and lactate. *Applied and environmental microbiology*, *77*(15), 5132-5140.

[103] Wang, X., Yomano, L. P., Lee, J. Y., York, S. W., Zheng, H., Mullinnix, M. T., ... & Ingram, L. O. (2013). Engineering furfural tolerance in Escherichia coli improves the fermentation of lignocellulosic sugars into renewable chemicals.*Proceedings of the National Academy of Sciences*, *110*(10), 4021-4026..

[104] van Zyl, W. H., Lynd, L. R., den Haan, R., & McBride, J. E. (2007). Consolidated bioprocessing for bioethanol production using Saccharomyces cerevisiae. In *Biofuels* (pp. 205-235). Springer Berlin Heidelberg.

[105] Slininger, P. J., Bolen, P. L., & Kurtzman, C. P. (1987). Pachysolen tannophilus: properties and process considerations for ethanol production from D-xylose. *Enzyme and microbial technology*, *9*(1), 5-15.

[106] Prior, B. A., Kilian, S. G., & Du Preez, J. C. (1989). Fermentation of D-xylose by the yeasts Candida shehatae and Pichia stipitis. *Process Biochemistry*,*24*(1), 21-32.

[107] Xu, Q., Singh, A., & Himmel, M. E. (2009). Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose.*Current Opinion in Biotechnology*, *20*(3), 364-371.

[108] Lawford, H. G., & Rousseau, J. D. (2002). Performance testing of Zymomonas mobilis metabolically engineered for cofermentation of glucose, xylose, and arabinose. In *Biotechnology for Fuels and Chemicals* (pp. 429-448). Humana Press.

[109] Mohagheghi, A., Evans, K., Chou, Y. C., & Zhang, M. (2002). Cofermentation of Glucose, Xylose, and Arabinose by Genomic DNA-lntegrated Xylose/Arabinose Fermenting

Strain of Zymomonas mobilis AX101. In*Biotechnology for Fuels and Chemicals* (pp. 885-898). Humana Press.

[110] Ingram, L. O., Aldrich, H. C., Borges, A. C. C., Causey, T. B., Martinez, A., Morales, F., ... & Zhou, S. (1999). Enteric bacterial catalysts for fuel ethanol production. *Biotechnology progress*, *15*(5), 855-866.

[111] Tao, H., Gonzalez, R., Martinez, A., Rodriguez, M., Ingram, L. O., Preston, J. F., & Shanmugam, K. T. (2001). Engineering a homo-ethanol pathway inEscherichia coli: Increased glycolytic flux and levels of expression of glycolytic genes during xylose fermentation. *Journal of bacteriology*, *183*(10), 2979-2988.

[112] Dien, B. S., Cotta, M. A., & Jeffries, T. W. (2003). Bacteria engineered for fuel ethanol production: current status. *Applied microbiology and biotechnology*,*63*(3), 258-266.

[113] Wood, B. E., Yomano, L. P., York, S. W., & Ingram, L. O. (2005). Development of Industrial-Medium-Required Elimination of the 2, 3-Butanediol Fermentation Pathway To Maintain Ethanol Yield in an Ethanologenic Strain of Klebsiellaoxytoca. *Biotechnology progress*, *21*(5), 1366-1372.

[114] Lynd, L. R., Van Zyl, W. H., McBride, J. E., & Laser, M. (2005). Consolidated bioprocessing of cellulosic biomass: an update. *Current opinion in biotechnology*, *16*(5), 577-583.

[115] Herpoël-Gimbert, I., Margeot, A., Dolla, A., Jan, G., Mollé, D., Lignon, S., ... & Asther, M. (2008). Comparative secretome analyses of two Trichoderma reesei RUT-C30 and CL847 hypersecretory strains. *Biotechnol Biofuels*, *1*(1), 18.

[116] Nagendran, S., Hallen-Adams, H. E., Paper, J. M., Aslam, N., & Walton, J. D. (2009). Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus Amanita bisporigera, based on the secretome of Trichoderma reesei. *Fungal Genetics and Biology*, *46*(5), 427-435.

[117] Yamada, R., Yamakawa, S. I., Tanaka, T., Ogino, C., Fukuda, H., & Kondo, A. (2011). Direct and efficient ethanol production from high-yielding rice using a Saccharomyces cerevisiae strain that express amylases. *Enzyme and microbial technology*, *48*(4), 393-396.

[118] Den Haan, R., Rose, S. H., Lynd, L. R., & van Zyl, W. H. (2007). Hydrolysis and fermentation of amorphous cellulose by recombinant Saccharomyces cerevisiae. *Metabolic engineering*, *9*(1), 87-94.

[119] Lykidis, A., Mavromatis, K., Ivanova, N., Anderson, I., Land, M., DiBartolo, G., ... & Kyrpides, N. (2007). Genome sequence and analysis of the soil cellulolytic actinomycete Thermobifida fusca YX. *Journal of bacteriology*, *189*(6), 2477-2486.

[120] Bachmann, S. L., & McCarthy, A. J. (1991). Purification and cooperative activity of enzymes constituting the xylan-degrading system of Thermomonospora fusca. *Applied and Environmental Microbiology*, *57*(8), 2121-2130.

[121] Deng, Y., & Fong, S. S. (2010). Influence of culture aeration on the cellulase activity of Thermobifida fusca. *Applied microbiology and biotechnology*, *85*(4), 965-974.

[122] Deng, Y., & Fong, S. S. (2011). Laboratory evolution and multi-platform genome re-sequencing of the cellulolytic actinobacterium Thermobifida fusca.*Journal of Biological Chemistry*, *286*(46), 39958-39966.

[123] Zhang, C., Yang, H., Yang, F., & Ma, Y. (2009). Current progress on butyric acid production by fermentation. *Current microbiology*, *59*(6), 656-663.

[124] Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.*Analytical biochemistry*, *72*(1), 248-254.

[125] Wilson, D. B. (2004). Studies of Thermobifida fusca plant cell wall degrading enzymes. *The Chemical Record*, *4*(2), 72-82.

[126] Chen, X., Shekiro, J., Franden, M. A., Wang, W., Zhang, M., Kuhn, E., ... & Tucker, M. P. (2012). The impacts of deacetylation prior to dilute acid pretreatment on the bioethanol process. *Biotechnol Biofuels*, *5*(8)..

[127] Jarvis, E., Adney, W. S., Decker, S. R., Sakon, J., Ding, S. Y., & Himmel, M. E. (2004, March). Increased thermal tolerance of Thermobifida fusca beta-glucosidase by protein engineering. In *ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY* (Vol. 227, pp. U291-U291). 1155 16TH ST, NW, WASHINGTON, DC 20036 USA: AMER CHEMICAL SOC.

[128] Tribe, L. A., Briens, C. L., & Margaritis, A. (1995). Determination of the volumetric mass transfer coefficient (kLa) using the dynamic “gas out–gas in” method: Analysis of errors caused by dissolved oxygen probes. *Biotechnology and bioengineering*, *46*(4), 388-392.

[129] Advanced VR Research Centre. (n.d.). Retrieved June 11, 2015, from http://www.lboro.ac.uk/research/avrrc/research/currentprojects/virtualengineering/designing-for-adaptability-and-evolution-in-system-of-systems-engineering-danse.html

[130] Zhang, C., Yang, H., Yang, F., & Ma, Y. (2009). Current progress on butyric acid production by fermentation. *Current microbiology*, *59*(6), 656-663.

[131] Dwidar, M., Park, J. Y., Mitchell, R. J., & Sang, B. I. (2012). The future of butyric acid in industry. *The Scientific World Journal*, *2012*.

Appendix A - Non-inclusive list of cellulolytic microorganisms

Phylogeny	Species	Source	Genome accession number
Class Clostridia Order Thermoanaerobacterales	Caldicellulosiruptor bescii		NC_012034
Family Incertae Sedis	Caldicellulosiruptor hydrothermalis	Hot spring	NC_014652
	Caldicellulosiruptor kristjanssonii	Hot spring	NC_014721
	Caldicellulosiruptor kronotskyensis	Hot spring	NC_014720
	Caldicellulosiruptor lactoaceticus		NC_015949
	Caldicellulosiruptor obsidiansis	Hot spring	NC_014392
	Caldicellulosiruptor saccharolyticus	Hot spring	NC_009437
	Thermoanaerobacter cellulolyticus	Hot spring	N
	Thermoanaerobacter thermocopriae	Hot spring	PRJNA224116
Class Clostridia Order Clostridiales	Butyrivibrio fibrisolvens	Rumen	NC_021031
Family Lachnospiraceae	Cellulosilyticum lentocellum	Rumen	NC_015275
	Cellulosilyticum	Rumen	N

	ruminicola		
Class Clostridia Order Clostridiales Family Eubacteriaceae	Eubacterium cellulosolvens	Rumen	NZ_CM001487
Class Clostridia	Clostridium aldrichii	Wood fermenter	N
Order Clostridiales	Clostridium alkalicellulosi	Soda lake	PRJNA65293
Family Clostridiaceae	Clostridium caenicola	Sludge	N
	Clostridium celerescens	Manure	N
	Clostridium cellobioparum	Rumen	N
	Clostridium cellulofermentans	Manure	N
	Clostridium cellulolyticum	Compost	NC_011898
	Clostridium cellulosi	Manure	PRJNA212730
	Clostridium cellulovorans	Wood fermenter	NC_014393
	Clostridium chartatabidum	Rumen	N
	Clostridium clariflavum	Sludge	NC_016627
	Clostridium herbivorans	Pig intestine	N
	Clostridium hungatei	Soil	N
	Clostridium josui	Compost	PRJNA195880
	Clostridium longisporum	Rumen	N

	Clostridium papyrosolvens	Paper mill	PRJNA55815
	Clostridum phytofermentans	Soil	NC_010001
	Clostridium populeti	Wood fermenter	N
	Clostridium sp. C7	Mud	N
	Clostridium stercorarium	Compost	NC_020134
	Clostridium straminisolvens	Rice straw	PRJDB821
	Clostridium sufflavum	Cattle waste	N
	Clostridium termitidis	Termite	PRJNA196409
	Clostridium thermocellum	Sewage + soil	NC_009012
	Clostridium thermopapyrolyticum	Mud	N
Class Clostridia	Acetivibrio cellulolyticus	Sewage	PRJNA51533
Order Clostridiales	Acetivibrio cellulosolvens	Sewage	N
Family Ruminococcaceae	Ruminococcus albus	Rumen	NC_014833
	Ruminococcus flavefaciens	Rumen	PRJNA224116
Class Clostridia Order Halanaerobiales Family Halanaerobiaceae	Hallocella cellulosilytica	Saline lake	N
Class Bacilli Order Bacillales	Thermoactinomyces sp. YX		N

Family Thermoactinomycetaceae			
Class Bacilli Order Bacillales Family Alicyclobacillaceae	Caldibacillus cellulovorans		N
Class Bacilli Order Bacillales Family Bacillaceae	Bacillus circulans Bacillus pumilis		PRJDB1314 NC_009848
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Frankineae Family Acidothermaceae	Acidothermus cellulolyticus	Acidic hot spring	NC_008578
Class Actinobacteria	Cellulomonas biazotea		N
Subclass Actinobacteridae	Cellulomonas cartae		N
Order Actinomycetales	Cellulomonas cellasea		N
Suborder Micrococcineae	Cellulomonas fimi	Soil	NC_015514
Family Cellulomonadaceae	Cellulomonas flavigena	Soil	NC_014151
	Cellulomonas gelida		N
	Cellulomonas iranensis	Forest soil	PRJDB435
	Cellulomonas persica	Forest soil	PRJDB980
	Cellulomonas terrae	Soil	N
	Cellulomonas uda	Sewage	N
Class Actinobacteria Subclass Actinobacteridae	Cellulosimicrobium cellulans	Soil	PRJEB571

Order Actinomycetales Suborder Micrococcineae Family Promicromonosporaceae			
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Micrococcineae Family Microbacteriaceae	Curtobacterium flaccumfaciens	Soil	PRJNA199964
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Micrococcineae Family Promicromonosporaceae	Xylanimonas cellulosilytica	Decayed tree	NC_013530
Class Actinobacteria Subclass Actinobacteridae	Actinoplanes aurantiaca	Soil	N
Order Actinomycetales	Micromonospora melonosporea	Compost	N
Suborder Micromonosporineae	Micromonospora chalcae	Soil	N
Family Micromonosporaceae	Micromonospora propionici	Termite	N
	Micromonospora ruminantium	Rumen	N
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Pseudonocardineae	Actinosynnema mirum	Grass blade	NC_013093

Family Pseudonocardiaceae			
Class Actinobacteria	Streptomyces albogriseolus		PRJNA73805
Subclass Actinobacteridae	Streptomyces aureofaciens	Compost	N
Order Actinomycetales	Streptomyces cellulolyticus		N
Suborder Streptomycineae	Streptomyces flavogriseus	Soil	N
Family Streptomycetaceae	Streptomyces lividans		NZ_CM001889
	Streptomyces nitrosporeus		N
	Streptomyces olivochromogenes		N
	Streptomyces reticuli	Soil	N
	Streptomyces rochei	Termite gut	N
	Streptomyces thermovulgaris		N
	Streptomyces viridosporus		PRJNA200053
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Streptosporangineae	Thermobifida alba		N
Family Nocardiopsaceae	Thermobifida cellulolytica	Compost	N
	Thermobifida fusca	Soil	NC_007333

Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Streptosporangineae Family Thermomonosporaceae	Thermomonospora curvata		NC_013510
Class Actinobacteria Subclass Actinobacteridae Order Actinomycetales Suborder Streptosporangineae Family Streptosporangiaceae	Thermobispora bispora Streptosporangium subroseum	Soil Soil	NC_014165 N
Class Fibrobacteria Order Fibrobacterales Family Fibrobacteraceae	Fibrobacter succinogenes	Rumen	NC_013410
Class Cytophagia	Cytophaga aurantiaca	Soil	PRJNA199184
Order Cytophagales	Cytophaga haloflava	Soil	N
Family Cytophagaceae	Cytophaga hutchinsonii	Soil	NC_008255
	Cytophaga krzemieniewskae	Soil	N
	Cytophaga rosea	Soil	N
	Sporocytophaga myxococcoides	Soil	PRJNA224116
Class Flavobacteriia Order Flavobacteriales Family Flavobacteriaceae	Flavobacterium johnsoniae	Soil	NC_009441
Class Bacteroidia	Bacteroides	Human faecal	PRJNA55279

	cellulosilyticus	sample	
Order Bacteroidales	Bacteroides cellulosolvens	Sewage	N
Family Bacteroidaceae	Bacteroides sp. P-1	Rotting biomass	N
Class Bacteroidia Order Bacteroidetes Order II. Incertae sedis Family Rhodothermaceae	Rhodothermus marinus	Hot spring	NC_013501
Class Betaproteobacteria Order Burkholderiales Family Alcaligenaceae	Achromobacter piechaudii	Soil	PRJNA47029
Class Gammaproteobacteria Order Enterobacteriales Family Enterobacteriaceae	Dickeya dadantii	Plant biomass	NC_012880
Class Gammaproteobacteria Order Xanthomonadales Family Xanthomonadaceae	Xanthomonas sp.	Brack water	N
	Rudaea cellulosilytica	Soil	PRJNA199387
Class Gammaproteobacteria	Cellvibrio gilvus	Soil	NC_015671
Order Pseudomonadales	Cellvibrio mixtus	Soil	N
Family Pseudomonadaceae	Cellvibrio vulgaris, fulvus	Soil	N
	Cellvibrio japonicus	Plant biomass	NC_010995

Subphylum delta/epsilon subdivisions Class Deltaproteobacteria Order Myxococcales Family Incertae Sedis	Myxobacter sp. AL-1	Soil	N
Class Thermotogae	Fervidobacterium islandicum	Hot spring	N
Order Thermotogales	Thermotoga maritima	Hot spring	NC_000853
Family Thermotogaceae	Thermotoga neopolitana	Hot spring	N

Appendix B - Protocol for RNA Isolation from *T. fusca*

Appendix A is the exact protocol used to isolate mRNA. Dr. Stephen Fong from Virginia Commonwealth University provided this protocol based on previous work that his lab did with isolation of mRNA from *T. fusca.* All comments and tips in the protocol are from the Fong Lab.

Provided Protocol:

Tips/Requirements:

- Changes gloves often!
- Avoid any DEPC-treated water or Detergents!
- Elute/Resuspend RNA sample in RNase-free water (no buffers)!
- Final samples should be ≥100ng/uL and ≥12 µL for Prokaryotic Gene Expression (they would like 2.5ug RNA total)
    - It is better to dilute concentrated samples and provide the minimal volume rather than submitting samples with higher concentrations but smaller volumes
    - Extra volume is always appreciated

Protocol for RNA Isolation from *T. fusca* using ZR Fungal/Bacterial RNA MiniPrep™

Equipment:

- Microcentrifuge
- Vortexer or Bead Basher/Cell Disruptor (Bead Basher is preferred but not necessary)

Supplies:

- ZR Fungal/Bacterial RNA MiniPrep (#R2014)

- DNase I (NEB #M0303 or Zymo #E1009) (used for optional In-Column DNase Treatment)
- RNase Decontamination Reagent (RNase Zap or RNase Away) (Optional but **highly encouraged**)
- RNAprotect Bacteria Reagent (Qaigen) (Optional but **highly encouraged**)
    - RNA/DNA Shield (Zymo) can be used, but lyses cells
- RNase/DNase free pipette tips and tubes (remember: RNases are not inactivated by autoclaving)
- Plenty of gloves

Prepare RNA area:

- Wipe down all surfaces with RNase Decontamination Reagent
- Wipe down the pipettes with RNase Decontamination Reagent
    - From here on, pipettes should be handled with fresh gloves whenever reentering RNA area

Prepare DNase I Mastermix and store on ice:

For each column to be treated:

10X DNase Reaction Buffer – 10uL

RNase-free DNase I enzyme – 10uL (1U/uL)

RNA Wash Buffer - 80uL

From a culture of *T. fusca:*

1. To 6mL of culture add 2 volumes (12mL) of RNAprotect Bacteria Reagent, vortex briefly, and incubate at room temperature for at least 5 minutes (no longer than 2 h). This step is optional, but encouraged **especially if the cell pellets are going to be stored (frozen) before RNA prep.**

- I used 15mL conical tubes and a benchtop centrifuge here, but this step can be done in 1.5/2mL tubes in series, recombining the pellets during the resuspension (step 3).

2. Centrifuge the sample and decant the supernatant. Either proceed directly to RNA prep or freeze the pellets at -70^{o}C. According to RNAprotect reagent protocols, the pellet can be stored at -70 for up to 4 weeks.

3. With fresh pellet or thawed frozen pellet, resuspend pellet in 800-1000uL of **RNA lysis buffer** and transfer the mixture into **ZR BashingBead Lysis Tube**

4. Use BeadBasher/Cell Disruptor if available and disrupt for 1-2 minutes at max speed. If using a vortexer, try laying the tubes flat on the vortexer and taping them down securely. Run the vortexer at full speed for no less than 5 minutes. Ensuring the mixtures are **fully lysed is essential to the downstream success of RNA isolation**.

- *T. fusca* is relatively tough to lyse and without a proper BeadBasher will need vigorous vortexing, since the motion of the vortexer is not the same as a BeadBasher and is much less efficient at lysis.

5. Centrifuge the **ZR BashingBead™ Lysis Tube** at ≥12,000 x *g* for 1 minute.

6. Transfer 400-500 µl supernatant to a **Zymo-Spin™ IIIC Column** in a **Collection Tube** and centrifuge at 8,000 x *g* for 30 seconds. **Save the flow-through!!!**

Note: Maximum load volumes of IIIC and IIC columns is 800uL

7. Add 0.8 volume ethanol (95-100%) to the flow-through in the **Collection Tube** and mix well *(e.g., 320 µl ethanol added to 400 µl flow-through).*

8. Transfer the mixture to a **Zymo-Spin™ IIC Column** in a **Collection Tube** and centrifuge at ≥12,000 x *g* for 30 seconds. Discard the flow-through.

9. Add 400uL of **RNA Wash Buffer** to the **Zymo-Spin™ IIC Column** in a **Collection Tube** and centrifuge at ≥12,000 x *g* for 1 minute. Discard the flow through.

10. Add 100uL of the **DNase I Mastermix** prepared earlier, directly to the matrix of the **Zymo-Spin™ IIC Column**. Keep the **Zymo-Spin™ IIC Column** in the **Collection tube.**

11. Incubate the column at 25-37^{o}C for ≥15 minutes (less time is needed at higher temperatures), then centrifuge at ≥12,000 x *g* for 30 seconds.

12. Add 400 µl **RNA Prep Buffer** to the column. Centrifuge at ≥12,000 x *g* for 1 minute. Discard the flow-through and replace the **Zymo-Spin™ IIC Column** back into the **Collection Tube**.

13. Add **800 µl RNA Wash Buffer** to the column. Centrifuge at ≥12,000 x *g* for 30 seconds. Discard the flow-through and replace the **Zymo-Spin™ IIC Column** back into the **Collection Tube**.

14. Add **400 µl RNA Wash Buffer** to the column. Centrifuge at ≥12,000 x *g* for 30 seconds. Discard the flow-through and replace the **Zymo-Spin™ IIC Column** back into the **Collection Tube**.

15. Centrifuge the **Zymo-Spin™ IIC Column** at ≥12,000 x *g* for 2 minutes in the emptied **Collection Tube** to ensure complete removal of the wash buffer.

16. Carefully remove the **Zymo-Spin™ IIC Column** from the **Collection Tube** and place into a **DNase/RNase-Free 1.5 mL Tube**. Add ≥25 µl **DNase/RNase-Free Water** directly to the column matrix and let stand for 1 minute.

17. Centrifuge at 10,000 × *g* for 30 seconds to elute the RNA from the column. RNA can be used immediately or stored at ≤-70 ^{o}C

- Once eluted is **crucial** to make sure the samples do not encounter any contamination from RNases. This means always using RNase/DNase-free pipette tips, using fresh gloves and minimizing handling. Taking out a small aliquot of the eluent for subsequent analysis is a good idea.

- **Once frozen, the RNA sample should stay frozen until it arrives at the Sequencing center. Even one freeze-thaw cycle can significantly affect the quality of the RNA.**