Growth and regulatory metabolism of cellulose-utilizing thermophilic *Caldicellulosiruptor* species

by

Tina Suen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This thesis is my own work and all literature sources have been cited and acknowledged by means of complete references. This work has not previously been submitted elsewhere for the award of another degree. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

Hydrogen biofuels have been proposed as an environmentally-friendly and renewable energy substitute for petroleum-based fossil fuels. One species of bacteria capable of degrading cellulosic materials and producing vast amounts of hydrogen is *Caldicellulosiruptor saccharolyticus*. This species is one of the more well-studied thermophilic hydrogen producers as it can utilize a broad range of substrates and produce near theoretical maximum amounts of hydrogen by dark fermentation. Another species, *Caldicellulosiruptor kristjanssonii*, had its genome recently published but there is limited research done on this bacterium. It is studied here because of its ability to withstand even higher temperatures than *C. saccharolyticus*. It was found that both organisms were able to utilize glucose, xylose, cellobiose, Avicel, CMC, switchgrass, and xylan as sole carbon sources, with the exception of *C. kristjanssonii* which could only grow on CMC with yeast extract supplementation. Yeast extract was also determined to decrease, if not eliminate, the lag phase of growth for both organisms though growth was possible without yeast extract with the supplementation of a vitamin solution. *C. saccharolyticus* grew the best on xylose with cell densities reaching 2.5×10^8 cells/mL while *C. kristjanssonii* grew the best on cellobiose, reaching cell densities of about 3.2×10^8 cells/mL. Growth was not hindered on modified media containing 2 g/L soluble sugars or 4 g/L insoluble polymers, but it did have a detrimental effect on the growth rate and hydrogen yields, although carbon balances were near 100%. Enzyme assays were performed to study the native cellulase activities of these organisms while only recombinant enzyme assays have been done previously. Greater exoglucanase, endoglucanase, and xylanase activity was observed in the supernatant portion of the cultures compared to the cell-free extracts, suggesting that

cellulases are secreted into the extracellular environment. Generally, the cellulases had the highest specific activity with the corresponding growth substrate, but it is noteworthy that growth on Avicel induced the synthesis of xylanases in *C. saccharolyticus*. Proteomics and RNA-sequencing were done to confirm the expression of enzymes previously annotated in the genomes as well as to examine metabolic regulation when grown in different conditions. Remarkably, growth on glucose and xylan stimulated flagella production even though both organisms are reportedly non-motile. These flagellar proteins are thought to be associated with substrate attachment rather than motility. Moreover, many glycoside hydrolase proteins were upregulated according to the growth substrates the organisms were subjected to and several genes encoding ABC transporters (Csac_2504, Csac_2506, Csac_0681 and Calkr_2435) were found to have broad substrate specificities. Proteomic analyses revealed that xylose isomerases, an enzyme of the pentose phosphate pathway, were upregulated not only in growth conditions with pentose sugars, but also in cultures grown on hexose sugars. This finding indicates that the hexose sugars are being converted into fructose, facilitating a carbohydrate preference. Enzymes in the EMP pathway including phosphofructokinase (Csac_2441 and Calkr_1980) were upregulated when grown on glucose compared to other conditions, as expected. With these findings, the aim is to gain a better understanding of the metabolism of *C. saccharolyticus* and *C. kristjanssonii* in hopes of optimizing downstream processes in hydrogen production.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Kesen Ma, for making me a part of his lab, involving me in many opportunities and advising me throughout this research project. I would also like to thank my committee members, Dr. Owen P. Ward and Dr. Brendan J. McConkey, for their help and guidance.

Special thanks goes to Dr. Andrew C. Doxey for his help with bioinformatics, Ricardo Nordeste and Dr. Jiujun Cheng for their advice on RNA extraction, and to Vic Spicer, Justin Zhang, and Dr. Richard Sparling from the University of Manitoba, for their help with collecting proteomics and RNA-sequencing datasets. This research would not be possible without funding and support from the University of Waterloo, Genome Canada, Genome Prairie, and the Microbial Genomics for Biofuels and Co-Products from Biorefining Processes (MGCB^2) project.

I would also like to express my deep appreciation to all of the lab members who supported me throughout my years in this lab including Emma Chen, Rashid Shah, Emily Tse, Sheila Chan, Steffie Mendaglio, and my volunteers Drew Neish, Kashuf Ijaz, Paul Laffin, Nathan Ing, Ryan Tran, Vivian Nguyen, Kevin Woo, and Afzaal Hessan. You have truly made this a greater experience and I would not have been able to do it without you.

Thank you to the members of the Biology Department: Susan Lolle, Cheryl Duxbury, Jim Chambers, Karen Miinch, Terence Tang, Nathan Vo, Maya D'Alessio, Jessica Mendoza, Laura Beecraft, Laura Bell, and the other members of the BGSA, as well as all the others who I have worked with throughout my masters program.

Last but not least, I would like to thank my family, Kent, Man Wah, and Felix Suen and friends Sheldon Chan, Elisabeth Huang, Gina Kim, Matthew Leung, and Michael Ng for their love and support throughout my life and this academic endeavor. I dedicate this work to them.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABC	ATP-binding cassette transporter
ADH	Alcohol dehydrogenase
CAZy	Carbohydrate-Active enzymes database
CBM	Carbohydrate binding motif
CBP	Consolidated bioprocessing
CEC	Cell-free extracellular cellulase
CFE	Cell free extract
CMC	Carboxymethyl cellulose
CODH	Carbon monoxide dehydrogenase
DNA	Deoxyribonucleic acid
DNS	3,5-dinitrosalicylic acid
DOE	U.S. Department of Energy
ED	Entner-Doudoroff pathway
EDTA	Ethylenediaminetetraacetic acid
EM/EMP	Embden–Meyerhof–Parnas pathway
EtOH	Ethanol
GC	Gas chromatography
GH	Glycoside hydrolase
HPLC	High performance liquid chromatography
JGI	Joint Genome Institute
KDPG	2-keto-3-deoxy-6-phosphogluconate
LC	Liquid chromatography
LDH	Lactate dehydrogenase
MS/MS	Tandem mass spectrometry
NCBI	National Center for Biotechnology Information
PFK	Phosphofructokinase
PP/PPP	Pentose phosphate pathway
PPDK	Pyruvate phosphate dikinase
PTS	Phosphotransferase system
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SG	Switchgrass
YE	Yeast extract

1.0 Introduction

With an ever increasing demand for energy to power urbanization and development, biofuels are becoming a highly desired and essential commodity. Fossil fuels are approaching depletion and they contribute to the release of greenhouse gases which lead to climate change (Chandel, Giese, Singh, & Silva, 2013). However, biofuels such as biohydrogen and bioethanol are renewable and environmentally-friendly. These compounds can be produced by some microorganisms using abundant biomass such as cellulosic materials (Chandra, Takeuchi, & Hasegawa, 2012; Gowen & Fong, 2010).

1.1 Biofuels as an alternative energy source

Biofuels are renewable sources of energy that are converted from organic matter instead of petroleum or fossil-based materials into energy dense compounds such as bioethanol or biodiesel (Zhang, Rodriguez, & Keasling, 2011). These biofuels are commercially available in some places around the world such as the United States and Brazil, while biomethanol and biobutanol are in their pilot plant stages, and biohydrogen is even earlier in its development at the laboratory research stage. Although bioethanol and biodiesel can be mass produced at this time, they spark controversy as food crops such as corn, sugarcane, and soybeans are being used to produce these fuels (Antoni, Zverlov, & Schwarz, 2007). There is great potential for biohydrogen because it is still in the research stage of development and hydrogen fuel cells have a greater combustion efficiency than other fuels. Additionally, since carbon is converted within the culture media during microbial fermentation rather than released during combustion, carbon emissions can be controlled (Hallenbeck, 2009). Hence, biohydrogen is being proposed as a promising alternative fuel.

1.1.1 Hydrogen Biofuels

Biohydrogen is an optimal biofuel because it combusts cleanly to water (see Equation 1 below) (Brynjarsdottir, Scully, & Orlygsson, 2013) as opposed to the carbon dioxide and carbon monoxide that is released from other fuels (Kapdan & Kargi, 2006).

$$\mathbf{2H_{2(g)} + O_{2(g)} \rightarrow 2H_2O_{(g)}}$$

Equation 1. Hydrogen combustion to water.

A large portion of hydrogen is currently being produced from fossil fuels including natural gas and coal, which is still an issue as these are non-renewable resources and carbon dioxide is emitted as a byproduct (Rothstein, 1993). More recently, it has been proposed that hydrogen could be produced from types of waste materials such as wastewater, sludge or agricultural wastes containing starch or cellulose (Kapdan & Kargi, 2006).

1.1.2 Cellulosic Substrates

Of the types of waste materials mentioned above, cellulosic materials are one of the most abundant as they can come from agricultural, forest, and industrial residues, and it is estimated that about 220 billion tons of cellulosic biomass is produced each year (Ren, Guo, Liu, Cao, & Ding, 2011). Cellulose is a structural component in plants which contains other components such as hemicellulose and lignin (Rubin, 2008), Although cellulosic materials generally require mechanical or chemical pretreatment prior to fermentation (Kapdan & Kargi, 2006), several microorganisms have been found to be able to degrade cellulose as well as produce hydrogen. The ability to hydrolyze cellulosic materials and ferment the sugars into biofuels is known as consolidated bioprocessing (CBP). Organisms capable of CBP are wanted because this is an efficient process going from substrate to product in one step (Lynd, Weimer, Zyl, & Isak, 2002).

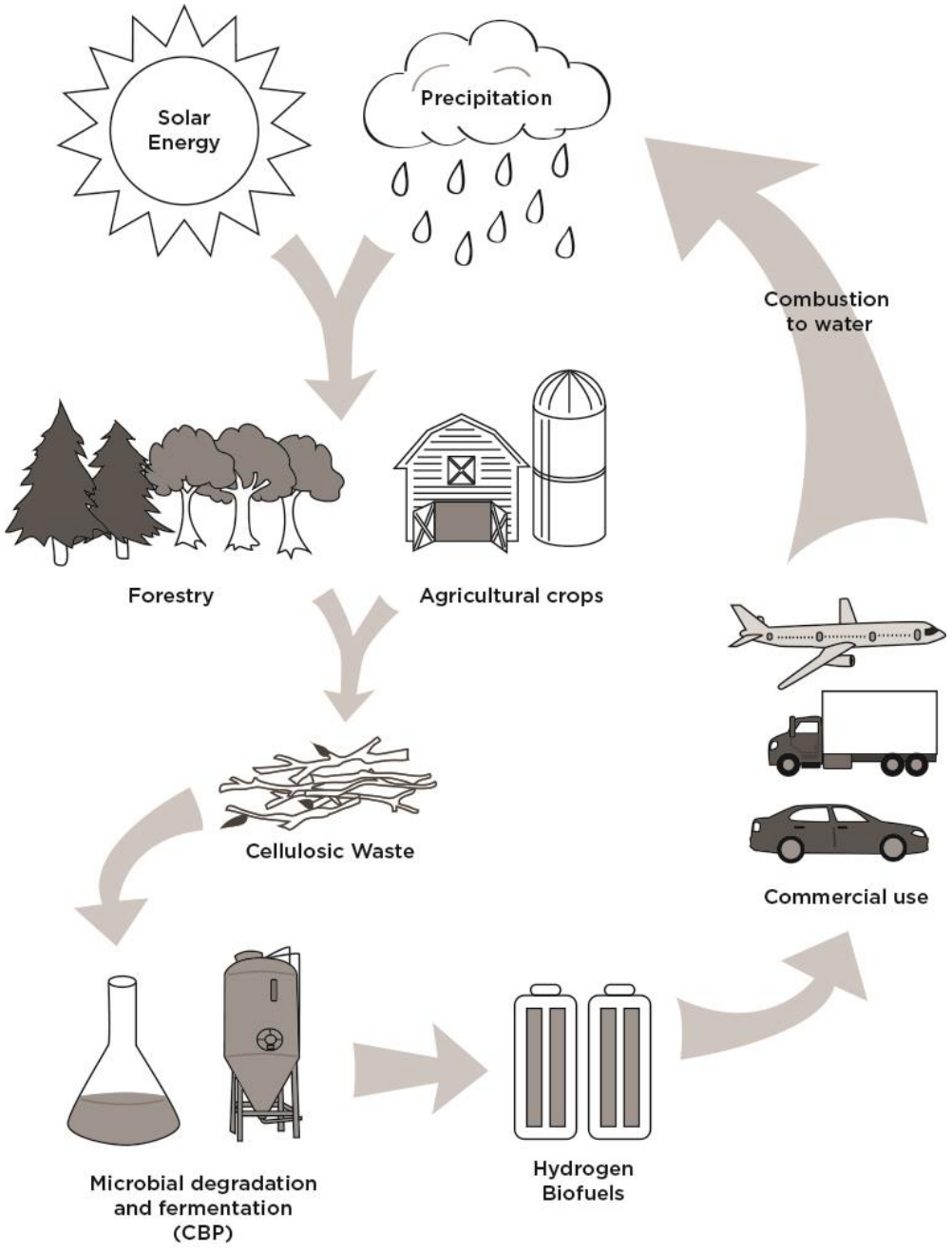


Figure 1. Concept for hydrogen biofuel production from cellulosic materials
Solar energy and CO_2 is converted into sugars within plants via photosynthesis. After forests and farms have been harvested for wood, paper, and food crops, the cellulosic waste materials can be collected for biofuel production. Microorganisms can break down cellulose and ferment the resulting sugars into hydrogen in a single step called consolidated bioprocessing (CBP). Further engineering is needed before fuel cells can effectively store hydrogen but when hydrogen fuel cells become commercially available, they can be used to power many forms of transportation and machinery. When combusted, the hydrogen turns into water or steam and is released and recycled into the environment.

1.2 Cellulolytic and Hydrogen-Producing Microorganisms

Microorganisms have the ability to degrade cellulose as well as produce hydrogen as natural processes in their metabolism. Both eukaryotic and prokaryotic microorganisms can be used for biofuel production. Currently, several types of fungi such as *Trichoderma reesei* (Durand, Soucaille, & Tiraby, 1984), *Penicillium* sp., *Humicola insolens* (Gowen & Fong, 2010), and *Phanerochaete chrysosporium* (Martinez *et al*., 2004) are exploited for their capacity to produce cellulases such as endoglucanases, exoglucanases, β-glucosidases, and hemicellulases (Durand *et al*., 1984; Gowen & Fong, 2010; Schwarz, 2001). *T. reesei* produces the most common cellulase enzymes used today for plant biomass saccharification (Kanafusa-Shinkai *et al*., 2013). Its enzyme profile contains two exoglucanases, four endoglucanases and one β-glucosidase which work together synergistically to hydrolyze cellulose (Kumar, Singh, & Singh, 2008). However, since *T. reesei* is a mesophilic organism, the enzymes it produces are heat labile with half-lives of just a few hours at 60°C (Durand *et al*., 1984).

The yeast *Saccharomyces cerevisiae* and bacterium *Klebsiella oxytoca* are also studied for their ability to degrade α-linked substrates such as starch and β-linked substrates such as Avicel and cellobiose (Lynd *et al*., 2002). Several species of *Clostridium* are also able to degrade cellulose and produce hydrogen including: *Clostridium cellulolyticum* (Levin, Carere, Cicek, & Sparling, 2009), *C. phytofermentans* (Warnick, Methé, & Leschine, 2002), and *C. cellovorans* (Sleat, Mah, & Robinson, 1984) but these organisms are mesophilic and do not produce thermostable enzymes. Since heat and acid pretreatment is used to breakdown cellulose, enzymes produced by thermoacidophiles are desired (Rubin, 2008).

1.2.1 Thermophilic microorganisms capable of utilizing cellulose

By definition, thermophilic bacteria are microorganisms that can grow in temperatures ranging from 45°C to 80°C (Bergey, 1919). Thermophiles are desired for their ability to tolerate high temperatures, especially in industrial applications which may require heat treatment or in processes that discharge heat as a byproduct. Furthermore, it has been found that thermophilic strains have higher substrate conversion efficiency than their mesophilic counterparts (Rittmann & Herwig, 2012), generating hydrogen yields near the theoretical maximum of 4 moles of hydrogen per 1 mole of hexose sugar (Bielen, Verhaart, van der Oost, & Kengen, 2013). Some of the thermophilic archaea currently being studied for their ability to degrade cellulose and produce hydrogen include: *Pyrococcus furious, P. abyssi*, *P. horikoshii*, *Thermococcus kodakarensis*, and *Sulfolobus solfataricus* (Blumer-Schuette, Kataeva, Westpheling, Adams, & Kelly, 2008). *Myceliophthora thermophila* and *Thielavia terrestris* are thermophilic fungi that are able to degrade cellulosic biomass (Berka *et al*., 2011). The cellulose enzymes collected from both *T. terrestris* and another thermophilic fungi, *Sporotrichum cellulophilum,* were compared to those of *T. reesei* but were found to be either less efficient at cellulose hydrolysis in the case of *T. terrestris* or just as heat labile from *S. cellulophilum* (Durand *et al*., 1984). Thermophilic bacteria that are also capable of utilizing cellulosic substrates include: *Thermotoga maritima*, *T. neapolitana*, *T. lettingae*, *T. naphthophila*, *T. petrophila*, *T. elfii*, *Anaerocellum thermophilum* (Blumer-Schuette *et al*., 2008), *Acidothermus cellulolyticus* (Rubin, 2008)*, Clostridium stercorarium*, *C. thermocellum*, and *C. straminisolvens* (Schwarz, 2001; Sizova, Izquierdo, Panikov, & Lynd, 2011).

C. thermocellum is one of the most studied organisms when it comes to thermophilic bacteria capable of performing CBP (Svetlitchnyi *et al*., 2013). It is able to produce both hydrogen and ethanol as end products. When grown on glucose, it produces a hydrogen yield of 1.64 moles and an ethanol yield of 1.36 moles per mole glucose (Islam, Özmihçi, Cicek, Sparling, & Levin, 2013). Interestingly, depending on the growth media composition provided, *C. thermocellum* would shift its metabolism to produce either more hydrogen or more ethanol (Islam, Sparling, Cicek, & Levin, 2015). Though *C. thermocellum* is a promising candidate for biofuel production, the hydrogen and ethanol yields are relatively low and since it is unable to ferment C5 sugars, co-cultures may be needed to ferment complex cellulosic and hemicellulosic materials (Svetlitchnyi *et al*., 2013).

While most of the other organisms mentioned above are able to utilize glucans and hemicellulose such as xylan, many of them are unable to utilize crystalline cellulose such as Avicel (Blumer-Schuette *et al*., 2008). Amongst the thermophilic bacteria which are adept at cellulose degradation, the *Caldicellulosiruptor* species are of particular interest since they able to utilize a broad range of sugars and cellulosic substrates including crystalline cellulose (Vanfossen, Verhaart, Kengen, & Kelly, 2009).

1.3 *Caldicellulosiruptor* spp.

Caldicellulosiruptor species are extremely thermophilic and cellulolytic bacteria that are capable of saccharification as well as hydrogen production (Blumer-Schuette *et al*., 2012). The genus *Caldicellulosiruptor* is classified under the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales* and family *Syntrophomonadaceae* (Wagner & Wiegel, 2008). At this time, there are nine isolated species of *Caldicellulosiruptor*, including *C. saccharolyticus* (Rainey *et al*., 1994), *C. lactoaceticus* (Mladenovska, Mathrani, & Ahring, 1995), *C. owensensis* (Huang, Patel, Mah, & Baresi, 1998), *C. kristjanssonii* (Bredholt, Sonne-Hansen, Nielsen, Mathrani, & Ahring, 1999), *C. acetigenus* (Onyenwoke, Lee, Dabrowski, Ahring, & Wiegel, 2006), *C. kronotskyensis*, *C. hydrothermalis* (Miroshnichenko *et al*., 2008), *C. bescii* (Yang *et al*., 2010), and *C. obsidiansis* (Hamilton-Brehm *et al*., 2010), with a tenth species *C. changbaiensis* being proposed this year (Bing *et al*., 2015).

Eight of the *Caldicellulosiruptor* species, excluding *C. acetigenus*, have their genomes sequenced. These species have a genome size ranging from 2.4 to 2.97 Mb, a 16S rRNA gene sequence identity range of 94.8 to 99.4% (Blumer-Schuette, Lewis, & Kelly, 2010), and they have a G+C content of 35 to 36% (Blumer-Schuette *et al*., 2011). It was found that of the eight species, *C. saccharolyticus* was the most phylogenetically divergent, possibly due to the fact that it is the only species found in New Zealand, as shown in Figure 2 (Blumer-Schuette *et al*., 2012). The genomic analysis also showed that endoglucanases and exoglucanase classified in the Carbohydrate-Active enzymes database (CAZy) as glycoside hydrolases (GH) belonging to family 5, 9, 44, and 48, were present in most but not all species of *Caldicellulosiruptor*.

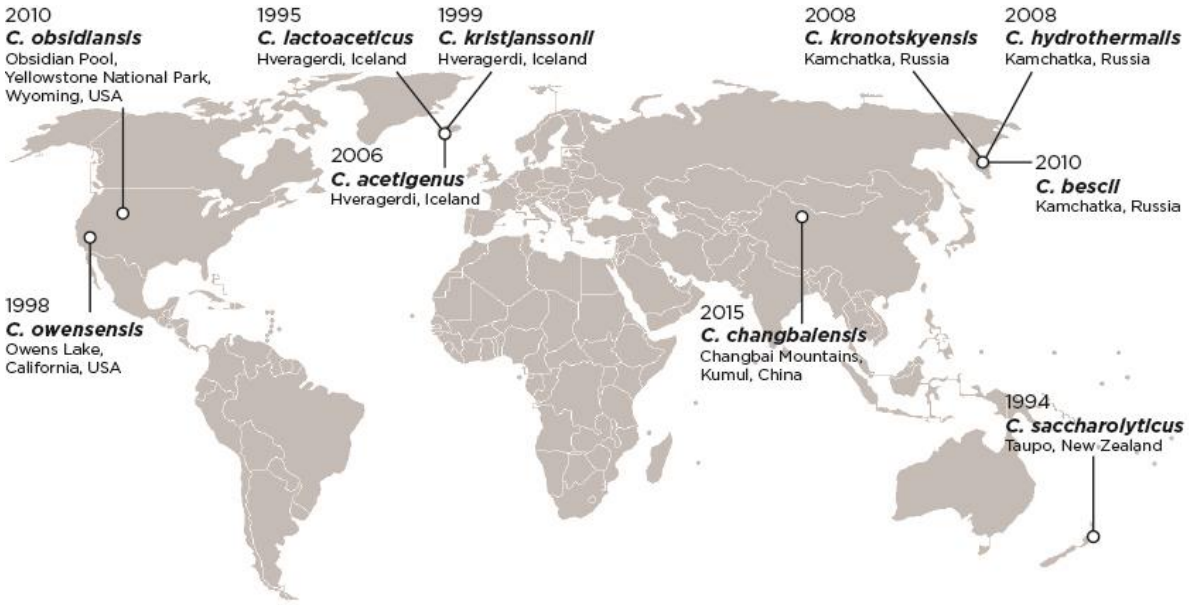


Figure 2. Geographical distribution of *Caldicellulosiruptor* species
Caldicellulosiruptor species have been discovered globally with several species being clustered in Iceland and Russia. *C. saccharolyticus* may be divergent from the other species since it is the only one found in New Zealand.

Furthermore, it was determined that all species of *Caldicellulosiruptor* are able to hydrolyze the β-1,4-xyloside linkages in the backbone of xylan and other types of hemicellulose (Blumer-Schuette, Lewis, & Kelly, 2010). This research group was also able to find new genomic loci that encode for GH9 and GH48 which differentiates the strongly cellulolytic species from the weaker ones. The loci were also associated with carbohydrate binding motifs (CBM) and a novel type IV pili, which are thought to aid in substrate adhesion. In addition, it is suggested that the ability to degrade crystalline cellulose was lost during evolution while the ability to hydrolyze amorphous and hemicellulose remains intact in *Caldicellulosiruptor* spp. (Blumer-Schuette *et al*., 2012).

All of these species are able to grow on glucose, xylose, xylan, and pretreated switchgrass with the exception that *C. lactoaceticus* was unable to grow on glucose. Growth on

crystalline cellulose was variable (Blumer-Schuette *et al*., 2012). *C. saccharolyticus* was able to reach the highest cell densities when grown on glucose, xylose, Avicel and switchgrass in comparison to the other *Caldicellulosiruptor* species. Moreover, since all of them have at least one endo-acting glycoside hydrolase (GH5) present in their genome, they are all able to hydrolyze carboxymethyl cellulose (CMC) (Blumer-Schuette *et al*., 2010).

Lately, one of the more commonly studied species is *C. bescii*. *C. bescii* was recently found to be the most thermophilic species of *Caldicellulosiruptor* with a maximum growth temperature of 90°C and an optimal growth temperature of 78°C (Yang *et al*., 2010). It has been co-cultured with *C. thermocellum* and shown to produce more ethanol than just with *C. thermocellum* alone (Kridelbaugh, Nelson, Engle, Tschaplinski, & Graham, 2013). The cell-free extracellular cellulase (CEC) system of *C. bescii* was also examined as opposed to cellulosomes typically found in *Clostridium* species. When comparing their CEC to the CEC derived from *T. reesei*, it was found that *C. bescii* degraded more than 2 times the amount of cellulose than *T. reesei* (Kanafusa-Shinkai *et al*., 2013). Studies have also been done to compare the secretomes of *C. bescii* with *C. obsidiansis*. Both species produced more glycoside hydrolases (GH) from families 5, 9, 10, 44, and 48 as well as family 3 carbohydrate binding modules (CBM3) when grown on Avicel. However, enzymes from *C. obsidiansis* had higher cellulase specific activity and higher thermostability than *C. bescii* (Lochner *et al*., 2011). Many other studies have been completed but nonetheless, *C.saccharolyticus* was the first of the genus to be identified and even more studies have been conducted on this organism than *C. bescii*.

1.3.1 *Caldicellulosiruptor saccharolyticus*

C. saccharolyticus is the most extensively studied species of the *Caldicellulosiruptor* genus due to its hydrogen producing capabilities and its capacity to utilize a broad range of cellulosic substrates (Vanfossen *et al*., 2009). *C. saccharolyticus* was isolated in 1994, from a geothermal springs located at Taupo, New Zealand. It is a thermophilic anaerobe that is non-motile and a non-spore-former. It is also a Gram-positive rod measuring about 3.0-4.0 μm by 0.4-0.6 μm, existing both singly and in pairs. It tolerates a pH range from 5.5 to 8.0 with optimal growth at pH 7.0, and its growth temperature ranges from 45°C to 80°C, while its optimal growth temperature is at 70°C (Rainey *et al*., 1994).

Since *C. saccharolyticus* has great potential for hydrogen production, many studies have been done with regards to the wide range of substrates it can utilize as well as its hydrogen yields. Growth on simple sugars such as glucose, xylose, galactose, arabinose, mannose, fructose, and on mixtures of these sugars, resulted in a doubling time of about 95 minutes and final cell densities of about 1×10^8 to 3×10^8 cells/mL (Vanfossen *et al*., 2009). It was also found that even though the metabolism of *C. saccharolyticus* was not affected by carbon catabolite repression (CCR) and it is able to utilize different substrates simultaneously, there was a substrate preference for fructose over xylose or arabinose and over glucose, mannose, or galactose (Vanfossen *et al*., 2009). The major end products observed include acetate, lactate, hydrogen, carbon dioxide and ethanol (Bielen *et al*., 2013), but a recent study using NMR spectroscopy found that acetoin, 2,3-butanediol, hydroxyacetone and ethylene glycol were substantial end products as well (Isern, Xue, Rao, Cort, & Ahring, 2013).

A study examining the hydrogen yield of *C. saccharolyticus* grown on glucose revealed that it is capable of producing 3.6 moles of hydrogen per mole of glucose. This is fascinating as the reported hydrogen yield is 90% of the theoretical maximum of 4 mol H_2/mol hexose in dark fermentation (de Vrije *et al*., 2007). A wide range of energy crops have also been studied including: sweet sorghum, sugarcane bagasse, wheat straw, maize leaves and silphium. These energy crops are thought to possess great potential for biofuel conversion since they can be regrown and harvested annually without the need to replant the crops. Wheat straw produced the most hydrogen at 3.8 mol H_2/mol glucose, surpassing the previously documented maximum yield and translating to 95% of the theoretical maximum. Maize leaves also come close and exceeds the previous hydrogen yields with 3.67 mol H_2/mol glucose but the other substrates tested do not yield significant amounts of hydrogen (Ivanova, Rákhely, & Kovács, 2009). As one way to make use of cellulosic waste, a study was done on paper sludge. Paper sludge offers a cheap and renewable resource but it was found that the hydrogen yield and production rates were lower than with simple sugars such as glucose and xylose (Kádár *et al*., 2004). Sugar beets were also investigated because the pulp of this crop is readily available as it is left over from sugar refineries in Europe. The main components of sugar beets comprise of sucrose, cellulose, hemicellulose, and pectin and when fermented by *C. saccharolyticus*, the hydrogen yields are about 10% higher than sucrose alone (Panagiotopoulos *et al*., 2010). The same group also investigated barley straw and sweet sorghum bagasse and found that *C. saccharolyticus* was able to grow on sugar concentrations of up to 20 g/L but concentrations of 30 g/L inhibited fermentation. The hydrogen yields on sorghum bagasse was lower at 2.6 mol H_2/mol hexose (Panagiotopoulos, Bakker, de Vrije, Claassen, & Koukios, 2012; Panagiotopoulos,

Bakker, de Vrije, Koukios, & Claassen, 2010) which is comparable to the sugarcane bagasse in the other study which produced 2.3 mol H_2/mol glucose (Ivanova *et al*., 2009). Another cellulosic substrate studied was carrot pulp which results from commercial carrot juice production. *C. saccharolyticus* was found to produce 2.8 mol H_2/mol hexose on carrot pulp hydrolysate (de Vrije *et al*., 2010). From these metabolic studies, it was concluded that hydrogen yields were higher on simple sugars or mixtures of sugars than on most complex polymers or biomass hydrolysates. This difference may be due to growth-inhibiting compounds being released from the biomass substrates during its pretreatment steps (Bielen *et al*., 2013). It has been proposed that hydrogen production yield and efficiencies be improved by using increasing cell densities. This could possibly be done by using higher concentrations of substrates yet genetically engineering *C. saccharolyticus* to be more osmotolerant to media containing added solutes. Additionally, cell densities can be increased by inducing biofilm formation (Willquist, Zeidan, & van Niel, 2010). A new study found that in fact, biofilm formation in co-cultures of *C. saccharolyticus* and *C. owensensis* increased biomass retention in reactors. This resulted in improved growth and hydrogen productivities (Pawar, Vongkumpeang, Grey, & van Niel, 2015).

Although many different growth substrates have been tested on *C. saccharolyticus*, most studies have been done on media containing yeast extract, which can be considered another carbon or nutrient source for the bacteria to grow on rather than solely on the selected sugars or biomass of interest. One study has shown that yeast extract may be excluded but instead supplemented with peptone or a vitamin cocktail containing necessary growth factors (Willquist & van Niel, 2012). Previous studies done in this lab

have also demonstrated that both *C. saccharolyticus* and *C. kristjanssonii* are able to grow on media containing a vitamin solution and various substrates as a sole carbon source, including: glucose, xylose, cellobiose, xylan, switchgrass, and Avicel (Ling, 2012). This chemically defined medium considers the effect of yeast extract on carbon balances.

When considering central metabolism, genomic studies as well as carbon isotope labelling and nuclear magnetic resonance (NMR) spectroscopy experiments have pointed to the Embden-Meyerhof-Parnas (EMP) pathway being the main route taken for glycolysis, with no evidence of the Entner-Doudoroff (ED) pathway being present (de Vrije *et al*., 2007; van de Werken *et al*., 2008). Genes encoding enzymes in the nonoxidative pentose phosphate pathway (PPP) have also been found in the genome of *C. saccharolyticus* (van de Werken *et al*., 2008). Remarkably, it has been reported that inorganic pyrophosphate (PPi) is an energy carrier in central metabolism and that it affects glycolysis. It has been found that the activity of the glycolytic enzymes, phosphofructokinase (PFK) and phosphate dikinase (PPDK), are dependent on PPi and because of a high PPi to ATP ratio during exponential growth, PPi-dependent glycolysis could possibly be a way for *C. saccharolyticus* to deal with a lower ATP yield (Bielen *et al*., 2010). Regarding fermentation, it has been reported that hydrogen, as an end product, inhibits the growth of *C. saccharolyticus* and causes a shift from acetate production to lactate and ethanol production (Van Niel, Claassen, & Stams, 2003; Willquist, Pawar, & Van Niel, 2011). Lactate production was also found to affect hydrogen production although the cause for inconsistent lactate production during the exponential growth phase is unknown (Kádár *et al*., 2004). The production of lactate occurs when the growth of *C. saccharolyticus* reaches stationary phase as a method of reductant disposal and to recycle electron carriers such as NAD^+ (Willquist & Van Niel, 2010). Lactate

dehydrogenase becomes active when ATP requirements are reduced, as it is no longer in the exponential growth phase (Willquist *et al*., 2010).

RNA-sequencing, also known as transcriptomics, looks at the RNA sequences that are transcribed given a certain growth condition. A transcriptome analysis of *C. saccharolyticus* investigating growth on glucose and xylose showed that growth on these substrates upregulated enzymes of the EM pathway including: fructose bisphosphate aldolase, glyceraldehyde-3-phosphate dehydrogenase, PPDK, and POR. Acetate kinase, belonging to a fermentative pathway, was also upregulated. Furthermore, several gene clusters, Csac0692 to Csac_0696, Csac0240 to Csac_0242, and Csac2416 to Csac_ 2419, were upregulated when grown on xylose. These genes are related to xylan and xylose degradation and include ABC transport systems and endoxylanases (van de Werken *et al*., 2008). Another study examined the transcriptomes of *C. saccharolyticus* grown on glucose, xylose, xyloglucan, switchgrass, and poplar. This study revealed that certain glycoside hydrolases and transporters were upregulated on complex biomass substrates. Carbohydrate uptake family 1 (CUT1) transporters including Csac_0679-Csac_0682, Csac_1557-Csac_1559, Csac_2412-Csac_2414, and Csac_2417-Csac_2419, were upregulated when grown on switchgrass or poplar biomass since these locus tags are annotated as xyloglucan or xylooligosaccharide transporters. It was also found that many GHs were constitutively expressed, enabling *C. saccharolyticus* to utilize a broad range of substrates at any time, and that CelB was highly upregulated when grown on plant biomass, suggesting that CelB is important in lignocellulosic hydrolysis (VanFossen, Ozdemir, Zelin, & Kelly, 2011). Another study examining the transcriptome response of *C. saccharolyticus* to high partial pressures

of hydrogen found that it shifts from acetate and hydrogen production to lactate and ethanol production. The enzymes upregulated include lactate dehydrogenase (Csac_1027), alcohol dehydrogenases (Csac_0407, Csac_0622) and hydrogenases (Csac_1860, Csac_1862-1864 and Csac_1534-1539) (Abraham A.M. Bielen *et al*., 2013).

Genomic and proteomic studies have revealed that *C. saccharolyticus* does not have cellulosomes for cellulosic degradation (van de Werken *et al*., 2008), but rather a wide array of glycoside hydrolases that are up-regulated when it is grown on Avicel (Blumer-Schuette *et al*., 2012). It has also been shown that *C. saccharolyticus* possesses genes that encode for ATP binding cassettes (ABC) and phosphotransferase system (PTS) transporters though carbohydrates are mainly transported through ABC transporters. A xylose-specific transport system was encoded by Csac_2504, Csac_2506 and Csac_2510. It is also notable that these genes were upregulated when grown on glucose or xylose suggesting that several sugars may be transported by the same system (Vanfossen *et al*., 2009). In fact, one type of sugar may be transported via several different transporters and each transport system may have a range of substrate specificities (Bielen *et al*., 2013). Previous proteomic studies done in this lab reveal that *C. saccharolyticus* possesses flagellar and attachment proteins, even though it is a non-motile bacteria. It is thought that the flagellar proteins enable the organisms to attach to insoluble substrates such as Avicel and switchgrass (Ling, 2012). *C. saccharolyticus* attachment to switchgrass was also documented in another study (VanFossen *et al*., 2011). Interestingly, a complete set of genes encoding flagella is found in *C. saccharolyticus* (Willquist *et al*., 2010) and is again suggested that flagella are used for cellulose substrate adhesion (Blumer-Schuette *et al*., 2012).

1.3.2 *Caldicellulosiruptor kristjanssonii*

C. kristjanssonii is one of the least studied species of *Caldicellulosiruptor*. *C. kristjanssonii* was isolated in 1999 from hot springs in Hveragerdi, Iceland (Blumer-Schuette *et al*., 2010). Similar to *C. saccharolyticus*, *C. kristjanssonii* is a cellulolytic, anaerobic thermophile that is a non-motile, and non-spore-forming bacterium. It is also a rod-shaped bacteria, measuring about 2.8-9.4 μm in length by 0.7-1.0 μm in width. It has a growth pH ranging from 5.8 to 8.0, and an optimal pH of 7.0, however, it has a higher growth temperature range from 45 to 82°C and an optimal growth temperature of 78°C (Bredholt *et al*., 1999). Previously known as the most thermophilic species of *Caldicellulosiruptor*, *C. kristjanssonii* is now second to *C. bescii* which can withstand temperatures of up to 90°C, although it shares a common optimal growth temperature of 78°C (Yang *et al*., 2010). In optimal conditions, it has a doubling time of approximately 2 hours (Bredholt *et al*., 1999). *C. kristjanssonii* differentiates itself from *C. saccharolyticus* in the fact that it is a Gram-negative bacteria (Bredholt *et al*., 1999).

C. kristjanssonii is able to grow on substrates including: arabinose, fructose, galactose, glucose, mannose, raffinose, sucrose, trehalose, Avicel (Onyenwoke *et al*., 2006), xylose, and yeast extract (Blumer-Schuette *et al*., 2010) and it is not inhibited by partial pressures of hydrogen of up to 50 kPa (Willquist *et al*., 2010). Moreover, a co-culture of *C. saccharolyticus* and *C. kristjanssonii* has been established to be able to co-exist stably and produce even more hydrogen than with either culture alone. The co-culture produced 3.7 mol H_2/mol hexose while each individual pure culture was able to produce 3.5 mol H_2/mol hexose alone (Zeidan, Rådström, & Van Niel, 2010).

In terms of genomic studies, the whole genome for *C. kristjanssonii* was recently sequenced by the Joint Genome Institute (JGI) website (Blumer-Schuette *et al*., 2011), and it is determined that *C. kristjanssonii* is most similar to *C. lactoaceticus* based on 16S rRNA (Onyenwoke *et al*., 2006). As for proteomic studies, only glycoside hydrolases and cellulosic attachment proteins or CBMs have been examined briefly from the *C. kristjanssonii* genome. *C. kristjanssonii* possesses the lowest number of GH enzymes in the genus and is classified as a weakly cellulolytic species, especially because it is missing GH48 which helps to hydrolyze crystalline cellulose (Blumer-Schuette *et al*., 2012). It was also found that in genome annotations from the NCBI database, *C. kristjanssonii* has 31 glycoside hydrolases compared to 49 glycoside hydrolases for *C. saccharolyticus*. This is much less than the over 400 glycoside hydrolase sequences *C. thermocellum* has but is still comparable to the 13 annotations found in the genome of *T. reesei*. Other than this, not many studies have been done on *C. kristjanssonii* since it is considerably harder to culture and its genome sequence was not readily available for many years after its discovery.

1.4 Advancing research on *Caldicellulosiruptor* spp.

Although numerous studies have already been done on *C. saccharolyticus* and some limited knowledge is obtained on *C. kristjanssonii*, there is still a lot more that could be explored in order to understand their metabolism and optimize biofuel production or other industrial applications. For instance, research has been done to show the broad substrate utilization of these organisms, but not to demonstrate growth on modified media without yeast extract or with limited amounts of carbon substrates available. How would these growth conditions affect the metabolites produced and the enzyme regulation in these organisms? Putative biochemical pathways have been proposed based on genomic analysis, but can these pathways be experimentally proven to exist? Proteomic analyses have been done on the secretome of *C. saccharolyticus* but not on the proteins that are present inside the cell. Transcriptomics analyses have also been completed for *C. saccharolyticus* but only on a few growth conditions. Likewise, enzyme activities have been examined for recombinant cellulases, but what is the activity like for native enzymes in these two species?
As *C. saccharolyticus* differs phylogenetically from *C. kristjanssonii*, will they show similar responses to varied growth conditions? Especially for *C. kristjanssonii*, barely any of these aspects have been studied so far, hence a lot can be gained from further research on these two organisms.

1.5 Purpose and Research Objectives

The purpose of this research is to gain a better understanding of the biochemical pathways and regulation involved in cellulose degradation, carbohydrate metabolism, and fermentation in *Caldicellulosiruptor* spp. To discover more about the growth and regulatory metabolism of *C. saccharolyticus* and *C. kristjanssonii*, the following objectives were set for this research project:

1. To compare the effect of different cellulosic substrates on growth. Substrates include: glucose (C6 monosaccharide), xylose (C5 monosaccharide), cellobiose (C6 disaccharide), Avicel (crystalline cellulose), CMC (amorphous cellulose), switchgrass (natural, complex polymer), and xylan (hemicellulose).
2. To quantify metabolic end products which give insights into the metabolic pathways being used as well as the organisms' potential for biofuel production. Major end products include: hydrogen, carbon dioxide, acetate, lactate, and ethanol.
3. To examine the cellulase activity of each organism from within the cell and from enzymes secreted extracellularly. Exoglucanase, endoglucanase and xylanase activity will be tested and compared between growth conditions.
4. To study the regulatory metabolism of the organisms through proteomic analysis where proteins of interest are expressed depending on growth substrates.
5. To study the regulatory metabolism of the organisms through transcriptomic analysis where certain RNA sequences are upregulated depending on growth substrate.

2.0 Materials and Methods

2.1 General Experimental Design

To generate a metabolic profile, the project was separated into several experiments including monitoring the growth of the organisms, examining the compounds they produce, conducting RNA-sequencing and proteomic analysis, and performing enzyme assays for some key enzymes involved in these metabolic pathways. Figure 3 is a flow diagram that summarizes the experiments and processes involved in this project.

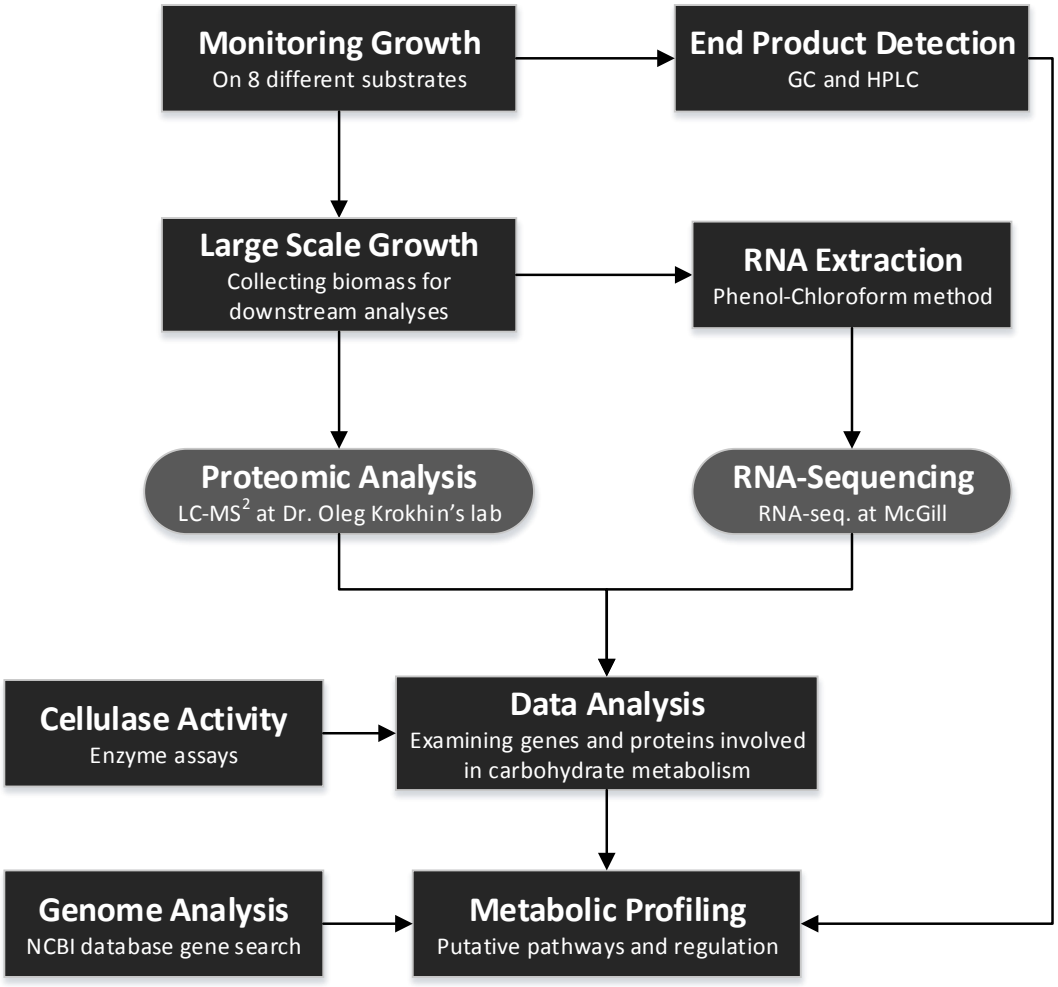


Figure 3. Experimental workflow to generate metabolic profiles for *C. saccharolyticus* and C*. kristjanssonii*. This diagram shows the general approaches taken to analyze the growth and metabolism of these organisms.

2.2 Microorganisms

The bacteria used in this project were *Caldicellulosiruptor saccharolyticus* strain Tp8T.6.3.3.1. (DSM 8903, ATCC 43494) and *Caldicellulosiruptor kristjanssonii* strain I77R1B (DSM 12137, ATCC 700853). Lyophilized cultures of these two organisms were obtained from the Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures.

The growth substrates used in this experiment include: glucose (Dextrose, Anhydrous, GR ACS, from EMD Millipore), xylose (D-(+)-Xylose, 98+%, from Alfa Aesar), cellobiose (D-(+)-Cellobiose, ≥98%, from Sigma-Aldrich), yeast extract (Yeast Extract, Granulated, from EMD Millipore), Avicel (Avicel PH 102, microcrystalline cellulose, NF, Ph. Eur., JP, BP, from FMC BioPolymer), CMC (Carboxymethylcellulose sodium salt, low viscosity, from Sigma-Aldrich), Switchgrass (Switchgrass, milled to 0.25 mm particle size using a UDY Cyclone sample milling machine (UDY 3010-080P, USA) and washed by stirring 20 g per 1L DI water incubated at 75°C for 24 hours and dried in a 50°C incubator overnight, produced and harvested from Nott Farms, Clinton, Ontario), and xylan (Xylan, from beechwood, ≥90% (HPLC), cell wall polysaccharide, from Sigma-Aldrich). Please see Appendix A for other chemicals used and exact media compositions.

2.3 Growth Experiments

Firstly, both *C. saccharolyticus* and *C. kristjanssonii* were rehydrated and cultivated in small volumes (approximately 3-5 mL) of the full DSMZ 640 and DSMZ 671 media at slightly lower than optimal temperatures of 65°C and 70°C respectively, as recommended by the DSMZ company website. The two microorganisms were then grown in their respective modified media and incubated at their optimal temperatures of 70°C and 78°C. To ensure that there was no contamination between these two closely related organisms, separate sets of instruments were used for each organism, and the cultures were stored on different benches in the lab. Growth was monitored through cell counts using a Petroff-Hausser counting chamber and growth curves were determined to ensure reproducible growth and consistent results for further downstream investigations.

2.3.1 Growth Media

Modified DSMZ Medium 640 – *Caldicellulosiruptor* Medium was used for cultivating *C. saccharolyticus* while modified DSMZ Medium 671 – Modified BA Medium was used for cultivating *C. kristjanssonii*. See Appendix A for growth media recipes and modifications. The 8 different carbon sources used in this experiment include: glucose, xylose, cellobiose, microcrystalline cellulose (Avicel), carboxymethyl cellulose (CMC), switchgrass (*Panicum virgatum*), and xylan. For more details about the substrates used, please refer to Appendix A.

2.3.2 Equipment for Monitoring Growth

Serum bottles and vacuum flasks were used to grow *C. saccharolyticus* and *C. kristjanssonii*. Small-scale cultures (50 mL) were grown in serum bottles with butyl rubber stoppers, crimp-sealed with aluminum caps to make the bottles airtight. These small-scale cultures were used to keep a running culture of active cells readily available through weekly transfers. They were also used for monitoring growth and for detection of end products. Large-scale flasks were required for collection of biomass for downstream analyses including proteomics, RNA-seq., and enzyme assays. For soluble substrates, 1 L cultures grown to collect biomass, while for insoluble substrates, 2 L cultures were grown and the substrates filtered out through cheesecloth before centrifuging the cell pellet.

To create anaerobic conditions for growing these strict anaerobes, a manifold was used to degas the culture bottles and flush with nitrogen gas in the following manner:
For 50mL cultures: degas bottles for 30 minutes with occasional agitation; gas with nitrogen for 3 minutes and degas for 3 minutes (repeated 3 times); and finally, the bottles were flushed with nitrogen for 1 minute and then pressurized for 2 minutes.
For 1-2L cultures: degas bottles for 45 minutes with occasional agitation; gas with nitrogen for 5 minutes and degas for 10 minutes (repeated 3 times); and finally, the bottles were flushed with nitrogen for 5 minute and then pressurized for 5 minutes.

For monitoring the growth and observing cell morphologies, a Nikon Eclipse E600 phase contrast microscope (Nikon, Japan) and a Petroff-Hausser counting chamber (Fisher Scientific, USA) were used.

2.4 End Product Detection

To observe the utilization of substrates and the production of desired end products such as hydrogen and ethanol, as well as to examine the pathways involved in carbohydrate metabolism, high performance liquid chromatography (HPLC) and gas chromatography (GC) are used to examine the liquid media and gas phase respectively.

2.4.1 High Performance Liquid Chromatography

Soluble substrates, metabolic end products, and intermediates were measured using high performance liquid chromatography (HPLC) (Prominence HPLC, Shimadzu, Japan) equipped with a LC-20AT solvent delivery unit, SIL-20AC autosampler, CBM-20A system controller, RID-10A refractive index detector, SPD-20AV UV-Vis detector, CTO-20AC column oven and an Alltech IOA-1000, 7.8 x 300 mm organic acid column. The mobile phase consisted of Millipore water combined with 0.02125 M H_2SO_4 as the solvent to get a final concentration of 0.0085 M. The flow rate was 0.4 mL/min using the gradient flow setting, with a column temperature of 60°C and a column pressure between 680 to 720 psi. The samples were kept cool at 4°C in the autosampler and volumes of 20 µL were taken for each sample. The software program used to collect and display the HPLC data was LCsolutions provided by Shimadzu and the compounds set to be detected using HPLC include: glucose, xylose, cellobiose, arabinose, acetate, citric acid, formate, succinate, lactate, and ethanol. A master mix of all of these components was made to run a standard curve. The glucose, xylose, cellobiose, arabinose and acetate were added at 10 mM concentrations while the lactate, succinate, formate, citric acid, and ethanol were added at 5 mM concentrations. This master mix was run using the same method set up for the

unknown samples, and the peak data was fitted automatically by the LCsolution software to produce a standard curve. Please see Appendix B for the standard curves.

Media samples of approximately 1mL were collected at initial, mid-log, late-log, and stationary phases of growth on each condition. These samples were centrifuged at 8,000 RPM for 10 minutes at 4°C (Allegra 21R centrifuge, F2402H rotor, Beckman Coulter, USA) and filtered through a 0.2 μm pore sized filter (Nalgene 4 mm, ThermoFisher, USA) before being placed into sample vials (10 mm glass vials, ThermoFisher, USA). Samples were run in the HPLC for 60 minutes each, after which the LCsolution software automatically calculates the compound concentrations.

2.4.2 Gas Chromatography

Products in the gaseous phase, namely hydrogen and carbon dioxide, were measured using gas chromatography (GC) (GC-2014, Shimadzu, Japan) equipped with a Grace Porapak Q 80/100 – 6’ × 1/8” × 0.85” SS divinylbenzene/ ethylvinylbenzene column. A thermal conductivity detector (TCD) was used with helium and nitrogen gas and a flame-ionization detector (FID) was used with hydrogen and air, both at a flow rate of 25 mL/min and a running column temperature of 140°C. The software used to gather and display the GC data was GCsolution provided by Shimadzu. To produce a standard curve for hydrogen and carbon dioxide, 150 mL serum bottles were degased and flushed with 100% nitrogen gas to act as a control (0%). Five other bottles containing nitrogen gas were injected with 0.1, 0.5, 1, 2, and 4% of either hydrogen or carbon dioxide, where the same amount of nitrogen was removed previously to maintain the same pressure. To create the standard curves, 100 μL

of gas from each bottle was run using the same method as the unknown samples. Please see Appendix B for the standard curves.

For the test samples, the pressure of gas in the headspace of each culture bottle was measured using a gauge, after which 100 µL samples from the gas phase were taken and directly injected into the GC septum with a microsyringe. These samples were collected at several time points including: mid-log, late-log, and stationary phases of growth on each condition. Since the GC is not automatic, samples were run in the GC for about 10 minutes each, at which time the machine was stopped manually and the peak areas recorded.

2.5 Bradford Assay for Protein Determination

Prior to performing enzyme assays, a Bradford assay was performed to estimate the total amount of protein in each concentrated supernatant or CFE sample. After samples were thawed on ice, they were inverted to mix and 800 µL of each sample was added to 200 µL of Bradford reagent (Bio-Rad protein assay reagent concentrate, Bio-Rad, USA). The assay mixtures were incubated at room temperature for 30 minutes and then their absorbances were read at 595 nm (Genesys 10 Spectrophotometer, Thermo Scientific, USA).

2.6 Enzyme Assays

Enzyme assays were performed to examine cellulases including endoglucanases, exoglucanases, and xylanases. *C. saccharolyticus* and *C. kristjanssonii* cultures were grown on glucose, xylose, Avicel, CMC, and xylan to compare their enzyme activities when grown on different conditions.

2.6.1 Sample Preparation

Large scale cultures (1-2L) of *C. saccharolyticus* and *C. kristjanssonii* were grown to late-log phase on their respective media containing each of the aforementioned substrates. Cultures grown on Avicel were vacuum filtered through cheesecloth (Cheesecloth wipes, 9" × 9", VWR, USA) to remove the insoluble substrates. The cells were pelleted by centrifuging them at 8,000 RPM for 20 minutes at 4°C (Sorvall RC 6 Plus centrifuge, SLA-3000 Super-Lite rotor, Thermo Scientific, USA), and then further condensed into 50 mL Falcon tubes (50mL Centrifuge tube, Sarstedt, Germany) for storage, by centrifuging them at 8,000 RPM for 20 minutes at 4°C (Sigma 3-18K, 19776-H rotor, Sigma, UK). The pellets were stored at -80°C prior to lysing and creating cell free extracts (CFE). The 500 mL of supernatant or spent culture media was also kept and concentrated by using a stirred cell (Amicon Stirred Cell 8400, 400 mL protein concentrator, EMD Millipore, Germany) fitted with a 10 kDa membrane disc (Ultrafiltration Disc, PLGC, Ultracel regenerated cellulose, 10 kDa NMWL, 76mm, EMD Millipore, Germany). This spent media was concentrated using 50 psi of nitrogen gas to approximately 5 mL while keeping the stirred cell set-up on ice. After concentrating, approximately 1 mL aliquots were placed into 1.5 mL microcentrifuge tubes and stored at -20°C until enzyme assays were to be performed. To make CFE, a cell pellet

was thawed on ice and placed into a 5 mL glass serum bottle with a small stir bar and sealed with a butyl rubber stopper and an aluminum crimp seal. The bottle was degased for 1 minute and flushed with nitrogen gas. Using a syringe, 5 parts (v/w) of Tris-HCl buffer (containing 0.1 M Tris-base and 0.01 M $MgCl_2 \times 6\ H_2O$, pH 7.5) and 10% (v/v) lysozyme and DNase solution (1 mg/mL lysozyme, 0.1 mg/mL DNase) were added. The serum bottle was again degased for 1 minute and then flushed with nitrogen gas for 1 minute, and this cycle was repeated 3 times in total. The serum bottle was then incubated at room temperature (~22°C) for 1 hour and then centrifuged at 8,000 RPM for 15 minutes at 4°C (Allegra 21R centrifuge, F0850 rotor, Beckman Coulter, USA) to pellet the cell debris. The resulting supernatant, which was the CFE, was collected using a syringe and injected into degased and nitrogen flushed glass vials (2 mL, Screw top vials, VWR, USA) with red rubber stoppers (Sleeve-type septa, Ace Glass Incorporated, VWR, USA) to be stored at -20°C until enzyme assays were to be performed.

2.6.2 Endoglucanase, Exoglucanase, and Xylanase Assays

Enzyme assays for endoglucanases, exoglucanases, and xylanases were done by testing with CMC, Avicel, and Xylan (from beechwood) as substrates respectively. To each 5 mL glass serum bottle, 0.016 g of the corresponding substrate and 0.91 mL of 0.1 M MES buffer (0.1 M MES and 0.01 M $MgCl_2 \times 6\ H_2O$, pH 6.0) were added. The bottles were then sealed with red rubber sleeve stoppers and placed into an 80°C waterbath for exoglucanase or endoglucanase assays or a 70°C waterbath for xylanases assays to equilibrate their temperature. Once the concentrated supernatant or CFE sample is thawed on ice, a syringe was used to take 0.7 mL of the sample and inject it into the serum bottle. The bottle was

mixed immediately by inverting and a 0.5 mL sample was removed and placed on ice as the t=0 sample. Samples were also taken out at 60 and 150 minutes for endoglucanase and exoglucanase assays and samples were taken at 30 and 60 minutes for xylanase assays.

2.6.3 DNS Method for detecting Cellulase and Xylanase Activity

The enzyme activity was measured indirectly through the DNS method for assaying reducing sugars (Sumner, 1921). After all of the samples have been taken from the enzyme assays, 400 µL of the sample was added to 400 µL of 1% DNS solution (1% (w/v) 3,5-dinitrosalicylic acid dissolved in 0.4 M NaOH) and incubated at 95°C for 15 minutes to allow for colour development. Samples were cooled to room temperature and then their absorbance was read at 540 nm (Genesys 10 Spectrophotometer, Thermo Scientific, USA).

2.7 Genome Annotations

A thorough search of the currently available genome annotations for proteins involved in specific metabolic pathways of interest was done. This was done online on the NCBI (http://www.ncbi.nlm.nih.gov/protein/?term=caldicellulosiruptor %20saccharolyticus and http://www.ncbi.nlm.nih.gov/protein/?term= caldicellulosiruptor+kristjanssonii) and the BioCyc (http://flamingo.cs.umanitoba. ca:1555) databases and it provides some clues for a putative metabolic pathway that *C. saccharolyticus* and *C. kristjanssonii* may take to catabolize and utilize carbohydrates.

2.8 Proteomics

Proteomics is the study of the entire set of proteins produced by an organism. For this project, it provides a snapshot of the proteins expressed when *C. saccharolyticus* and *C. kristjanssonii* are grown on eight different substrates: glucose, xylose, cellobiose, Avicel, CMC, switchgrass, xylan and yeast extract only.

2.8.1 Sample Preparation

Cell pellets of both *C. saccharolyticus* and *C. kristjanssonii* were prepared and shipped out to the University of Manitoba for proteomic analysis. First, 1-2 L cultures were grown on their respective media containing one of the eight test substrates. The cultures were grown to mid-log phase and harvested immediately. If the media contained insoluble substrates, the substrates were removed by vacuum filtration through cheesecloth (Cheesecloth wipes, 9" × 9", VWR, USA). The cells were pelleted by centrifuging them at 8,000 RPM for 20 minutes at 4°C (Sorvall RC 6 Plus centrifuge, SLA-3000 Super-Lite rotor, Thermo Scientific, USA), and then further condensed into 50 mL Falcon tubes (50mL Centrifuge tube, Sarstedt, Germany) for storage, by centrifuging them at 8,000 RPM for 20 minutes at 4°C (Sigma 3-18K, 19776-H rotor, Sigma, UK). The pellets were stored at -80°C until they were ready to be shipped.

2.8.2 Proteomic Analysis

Frozen samples of *C. saccharolyticus* and *C. kristjanssonii* were shipped in duplicate to the Manitoba Centre for Proteomics and Systems Biology for proteomic analysis.

Dr. Oleg Krokhin and members of his lab used 2D LC-MS/MS to identify and quantify proteins. The general workflow after receiving the cell pellets involves extracting, enriching

and digesting the proteins, and then all of the proteins are sprayed and ionized in a mass spectrophotometer using matrix-assisted laser desorption and ionization (MALDI), which allows for chromatographic separation. The data collected was then processed and organized through several bioinformatics tools before being sent back to our lab.

2.8.3 Proteomic Data Analysis

The proteomic datasets for both organisms received from the proteomics team at University of Manitoba were first processed by the bioinformatics team at the University of Manitoba who put the expression signal intensity data into an online database called UNITY. This data was organized into a table format and could be accessed and retrieved here at the University of Waterloo. Further biostatistical analyses for these datasets were done with the help of Dr. Andrew Doxey from the Department of Biology at the University of Waterloo. Dr. Doxey helped to normalize the data and run pairwise comparisons of the C6 sugar (glucose) versus C5 sugar (xylose), C6 sugar (glucose) versus C6 polymer (Avicel), and the C5 sugar (xylose) versus the C5 polymer (xylan), identifying overexpression or underexpression of proteins using the statistical program "R". A 2-fold change in regulation was used as a threshold for considering potentially relevant changes in protein expression. This data was summarized into tables which were then manually analyzed to find proteins related to cellulosic degradation, transport and carbohydrate metabolism. These datasets were also compared to the genes for metabolic pathway enzymes already available from the genome annotations available on NCBI and the U.S. Department of Energy Joint Genome Institute (DOE JGI) (https://img.jgi.doe.gov/cgi-bin/er/main.cgi) as well as the RNA-seq. data for *C. saccharolyticus* to confirm the presence of these enzymes.

2.9 RNA Sequencing

RNA sequencing (RNA-seq) uses shotgun sequencing techniques to profile the whole transcriptome of an organism. For this project, RNA-seq was used to examine the levels of transcripts present in *C. saccharolyticus* when it was grown on several selected conditions including growth on: glucose, xylose, cellobiose, xylan and yeast extract only.

2.9.1 RNA Extraction and Clean-up

To extract the RNA, the phenol-chloroform method was employed followed by the use of commercially available kits to clean up the RNA. The phenol-chloroform technique was adapted from the article, *The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction* (Chomczynski and Sacchi, 1987), and it involved extracting, precipitating, washing, and solubilizing the RNA. In this experiment, large-scale (1-2 L) bacterial cultures grown to mid-log phase, were first treated with 10% (v/v) cell stop solution consisting of 5% phenol in 100% absolute ethanol. If the media contained insoluble substrates, the substrates were removed by vacuum filtration through cheesecloth (Cheesecloth wipes, FisherBrand, USA) at this point. After any insoluble substrates were removed, 500 mL aliquots of the cultures were centrifuged at 8,000 RPM for 20 minutes at 4°C (Sorvall RC 6 Plus centrifuge, SLA-3000 Super-Lite rotor, Thermo Scientific, USA), and then the pellet was resuspended and further condensed into 50 mL Falcon tubes (Ambion RNase-free conical tubes, Life Technologies, USA) for storage, by centrifuging them at 8,000 RPM for 20 minutes at 4°C (Sigma 3-18K, 19776-H rotor, Sigma, UK). The pellet was stored at -80°C until ready to extract RNA.

When extracting the RNA, the pellet was thawed on ice and resuspended in 960 µL DEPC-treated water (deionized water treated with 0.1% diethylpyrocarbonate, stirred for 24 hours and autoclaved). The resuspended sample was split into two 480 µL aliquots and added to 480 µL of hot phenol solution (20 mM Tris base, 400 mM NaCl, 40 mM EDTA, pH 7.5, SDS solution added to 1% (v/v), and 1% β-mercaptoethanol added just prior to use) in microcentrifuge tubes. The two tubes were vortexed for 10 seconds and placed into a 95°C waterbath for 1 minute. The hot phenol solution disrupts cell membranes and denatures proteins including RNases. The tubes were then centrifuged at 13,000 RPM for 10 minutes at 4°C (Allegra 21R centrifuge, F2402H rotor, Beckman Coulter, USA). The supernatant was then added to two tubes of 600 µL of phenol:chloroform (1:1, v/v) respectively and vortexted for 10 seconds. During this step, proteins are trapped in the lower organic layer while nucleic acids persist in the upper aqueous layer. Phenol is used effectively for separating nucleic acids due to its non-polar nature, where nucleic acids are highly polar and will not dissolve in it, and because it is denser than water, it forms two distinct layers. Chloroform is added to the phenol to make the organic phase even denser, to ensure there is a clear distinction between the two phases. These tubes were then centrifuged at 13,000 RPM for 5 minutes at 4°C (Allegra 21R centrifuge, F2402H rotor, Beckman Coulter, USA). The transfer of the aqueous layer to new tubes of phenol:chloroform (1:1, v/v) was repeated twice more to ensure the RNA samples were clean and the proteins and other cell components are removed. The aqueous layer was transferred each time being cautious not to remove the interphase between the two layers since DNA may exist in the interphase or it may be in the organic phase since the solution was acidic. After that, the aqueous layer was transferred into 0.1 volumes of 3 M sodium acetate and 2.5 volumes of 100% ethanol.

The tubes were inverted to mix and stored on ice for 30 minutes so that the RNA could precipitate. The tubes were then centrifuged at 13,000 RPM for 10 minutes at 4°C (Allegra 21R centrifuge, F2402H rotor, Beckman Coulter, USA) to pellet the RNA and the supernatant was discarded. To wash the pellet, it was resuspended in approximately 600 µL of -20°C, 80% ethanol and centrifuged at 13,000 RPM for 5 minutes at 4°C (Allegra 21R centrifuge, F2402H rotor, Beckman Coulter, USA). The supernatant was discarded and the RNA pellet was left to air dry at room temperature for about 10 minutes or until it turned translucent. The RNA was then resuspended in 85 µL of RNase-free water (Sigma, USA) and 5 µL of DNase I and 10 µL of 10x reaction buffer (RQ1 RNase-free DNase, Promega, USA) was added and incubated for 20 minutes at room temperature to denature any DNA present in the sample.

Finally, columns from the QIAGEN RNeasy Mini Kit or the Geneaid Presto Mini RNA Bacteria Kit were used to purify the RNA samples according to the manufacturers' clean-up protocols. The RNA samples were finally eluted in RNase-free water provided in the respective kits and checked for quality (see section 2.9.2). The tubes were then dipped into liquid nitrogen and stored at -80°C to be shipped to external labs for further analysis.

2.9.2 Quality Control

To check the quality of the prepared RNA samples, a NanoDrop spectrophotometer and gel electrophoresis were employed. The NanoDrop (NanoDrop 2000, Thermo Scientific, USA) is used to examine the amount of RNA obtained as well as its purity by ensuring its A_{260}/A_{280} ratio is equal to or greater than 2.0. Lower absorbance ratios could indicate contamination of proteins or other cellular components (Chomczynski & Sacchi, 2006). Gel

electrophoresis was also done to visualize the RNA bands present. Formaldehyde agarose (1.2%) gels were prepared by dissolving 0.6 g of agarose into 36 mL of DEPC-treated water and adding 5 mL of 10x MOPS buffer (containing 0.2 M MOPS, 50 mM sodium acetate, 5 mM EDTA, adjusted to pH 7.0 and autoclaved), 9 mL of 37% formaldehyde, and setting in a gel rig (FB-SB-710 MiniGel System, FisherBiotech, USA) with a 10-well, 25 µL comb. When set, the gels were submerged in 300mL of 1X MOPS buffer and loaded with approximately 1 µg of RNA in each well. Gels were run at 90-100V (0.06A, 5W) (PowerPac HC High Current Power Supply, Bio-Rad, USA) for about 1 hour. To visualize the RNA bands, the gel was observed under UV light using the AlphaImager (AlphaImager 3400 Gel Documentation System, Alpha Innotech, USA). Two sharp bands were expected at 1.5 and 2.9 kb showing the 16S and 23S rRNA respectively and a small smear may be present around 250 bp or less, indicating the presence of small RNA (sRNA). If the two rRNA bands were not sharp, it signifies that the RNA was degraded due to possible RNase contamination. If there were any bands larger than 3 kb or even stuck in the loading wells, it was likely genomic DNA contamination. Only the RNA samples that had a NanoDrop reading of greater than or equal to 2.0 and two sharp bands visualized on the gel were sent for further RNA-seq analysis.

2.9.3 RNA Sequencing

RNA sequencing was done in McGill University and Genome Quebec Innovation Centre using Illumina technology. The *C. saccharolyticus* samples were shipped in duplicates and each of the samples were required to have a minimum of 10 µg of total RNA and the RNA concentration ≥ 300 ng/µL. About 5 µL of the RNA sample was first used for RNA quality control at their lab, verifying the quality on a RNA chip using an Agilent 2100 Bioanalyzer

(Agilent) and quantifying the RNA using a NanoDrop ND-1000 UV-VIS spectrophotometer (ThermoFisher). Messenger RNA was then purified from 1 μg of total RNA by selectively polyadenylating mRNAs using an *E. coli* poly(A) polymerase enzyme (Sorek & Cossart, 2010) and capturing the poly(A) mRNA using oligo-dT beads. The mRNA was reversed transcribed to generate cDNA fragments that were then sheared using a Covaris instrument to yield ~200 bp fragments. Libraries were made using the TruSeq RNA kit (Illumina) and quantified using PCR. Finally, an Illumina HiSeq2000 instrument was used to sequence the cDNA through 50-300 cycles and the reads were aligned to a reference transcriptome. Details about the procedures and technical workflows can be found on this website: http://gqinnovationcenter.com.

2.9.4 RNA Sequence Data Analysis

The raw RNA-seq. datasets received from McGill University were first processed by the bioinformatics team at the University of Manitoba who put the expression signal intensity data into an online database called UNITY. This data was organized into a table format which was retrievable at our lab. Further biostatistical analyses and data normalization for these datasets were done through the statistical program "R", with the help of Dr. Andrew Doxey from the Department of Biology at the University of Waterloo. The whole transcriptomes from select conditions were analyzed through pairwise comparisons of the C6 sugar (glucose) versus C5 sugar (xylose), C6 monosaccharide (glucose) versus C6 disaccharide (cellobiose), and the C5 sugar (xylose) versus the C5 polymer (xylan) for upregulation or downregulation. A 2-fold change in regulation was used as a threshold for considering potentially relevant changes in RNA. This data was organized into table format

and these transcripts were then compared to the genes for metabolic pathway enzymes available from the genome annotations found on NCBI and the U.S. Department of Energy Joint Genome Institute (DOE JGI) (https://img.jgi.doe.gov/cgi-bin/er/main.cgi) as well as proteomic datasets to confirm the presence of these enzymes.

3.0 Results

In order to better understand carbohydrate metabolism and regulation in *C. saccharolyticus* and *C. kristjanssonii*, the growth of these organisms were monitored when provided with different cellulosic substrates. Eight substrates were selected to compare metabolism on glucose (C6 monosaccharide), xylose (C5 monosaccharide), cellobiose (C6 disaccharide), Avicel (crystalline cellulose), CMC (amorphous cellulose), switchgrass (natural, complex polymer), xylan (hemicellulose), and yeast extract. End products were measured to examine metabolites produced during utilization of each substrate. Enzyme assays also confirm that cellulases are present and are utilized for substrate degradation. Additionally, a genomic analysis was compared with findings from RNA-sequencing and proteomic analyses to generate a metabolic map of enzymes and pathways involved in carbohydrate consumption, utilization and fermentation.

3.1 Growth on eight different substrates

To begin studies on growth and metabolism, the growth behavior and characteristics must be established for the specific conditions provided. Growth was monitored when each culture was grown on the eight conditions mentioned in section 2.3. Figures 4 and 5 below display the growth curves for *C. saccharolyticus* while Figures 6 and 7 show the growth curves for *C. kristjanssonii* grown on soluble and insoluble substrates respectively.

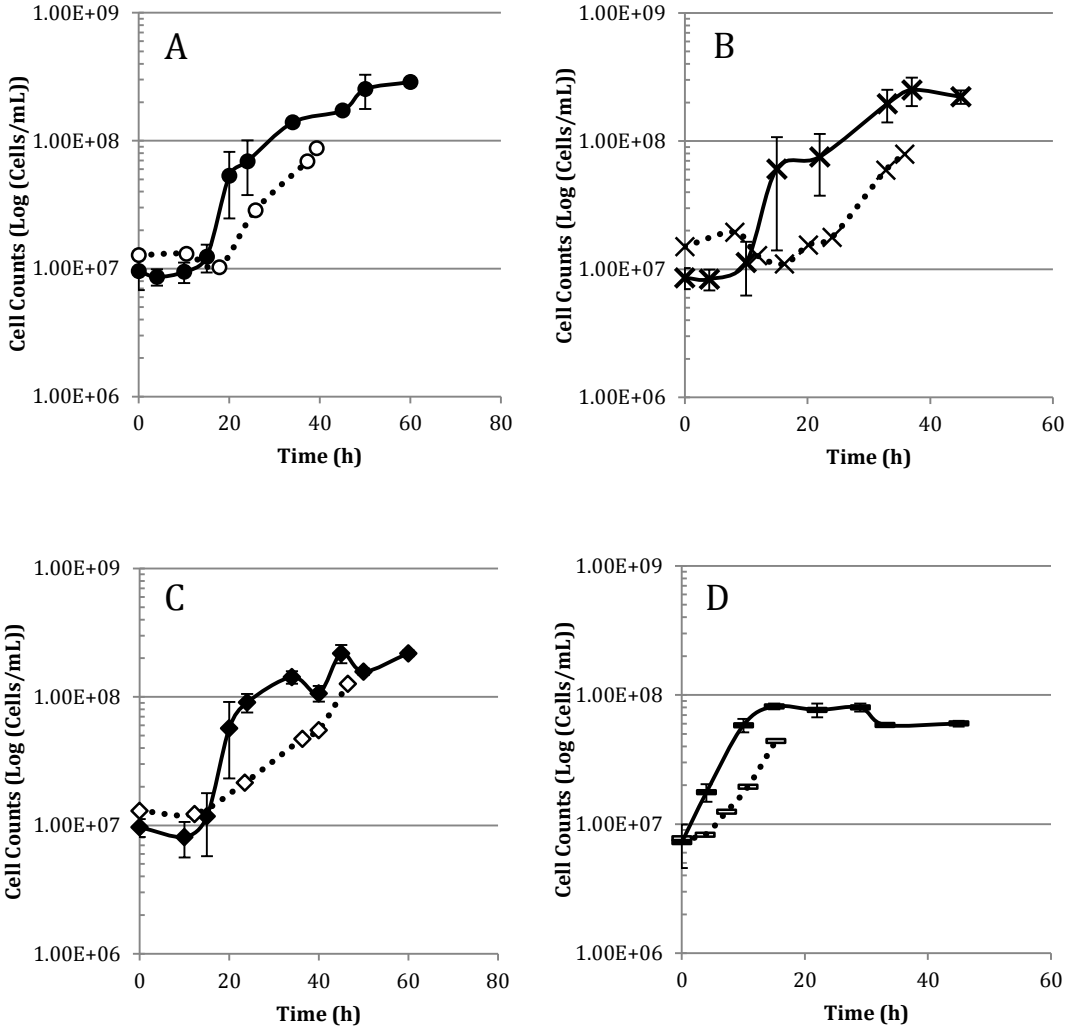


Figure 4. Small-scale and large-scale growth curves for *C. saccharolyticus* on four soluble substrates. Small-scale growth is shown in solid lines while large-scale growth is shown in dotted lines. (A) shows growth on glucose, small scale (—●—) and large scale (·○·). (B) shows growth on xylose, small scale (—✕—) and large scale (··✕··). (C) shows growth on cellobiose, small scale (—◆—) and large scale (·◇·). (D) shows growth on yeast extract only, small scale (——) and large scale (·▬·). The curves display an average of counts done in triplicates. Note that small-scale cultures were 50 mL while large-scale cultures were 1 L in volume.

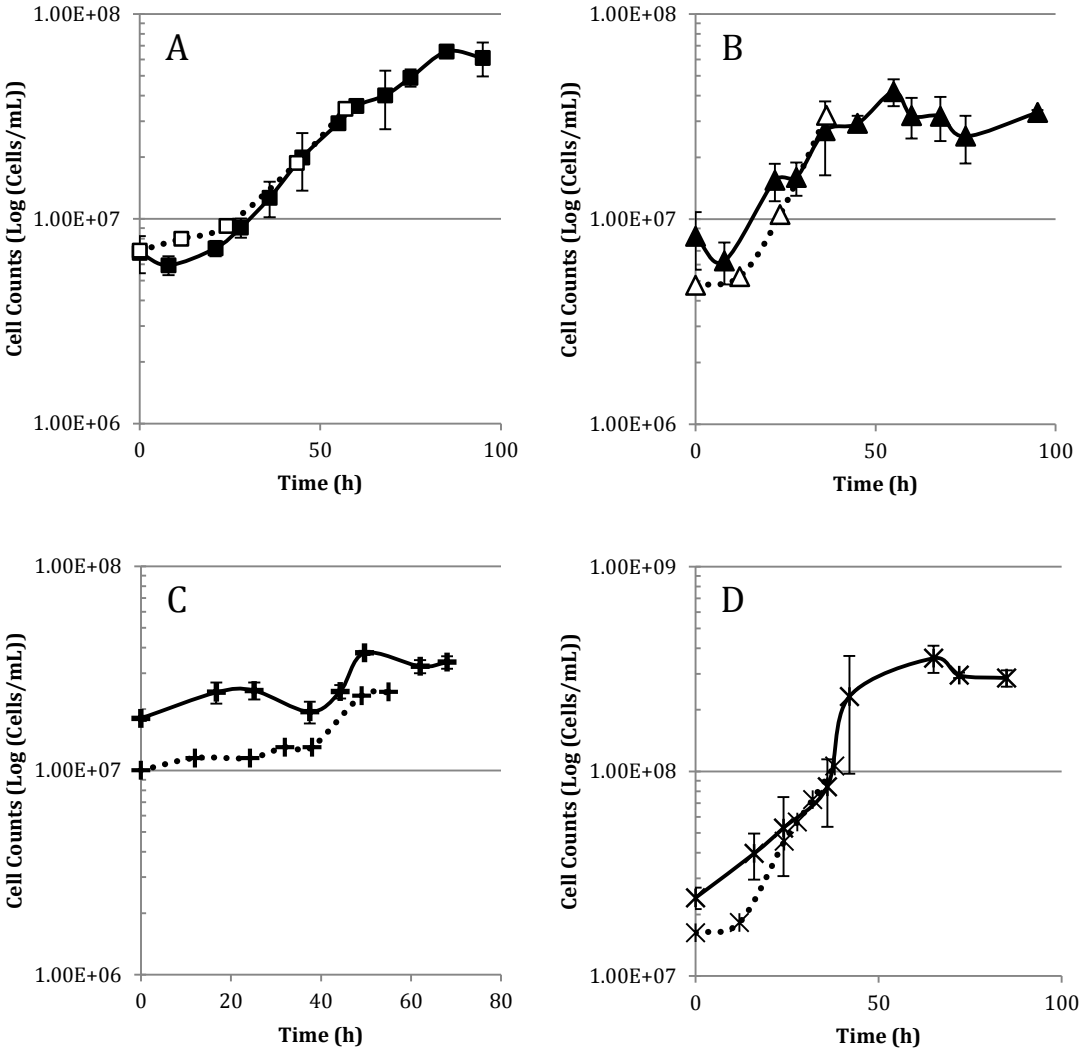


Figure 5. Small-scale and large-scale growth for *C. saccharolyticus* on four insoluble substrates. Small-scale growth is shown in solid lines while large-scale growth is shown in dotted lines. (A) shows growth on Avicel, small scale (—■—) and large scale (·□·). (B) shows growth on CMC, small scale (—▲—) and large scale (·△·). (C) shows growth on switchgrass, small scale (—+—) and large scale (·+·). (D) shows growth on xylan only, small scale (—✳—) and large scale (·✳·). The curves display an average of counts done in triplicates. Note that small-scale cultures were 50 mL while large-scale cultures were 1 L in volume.

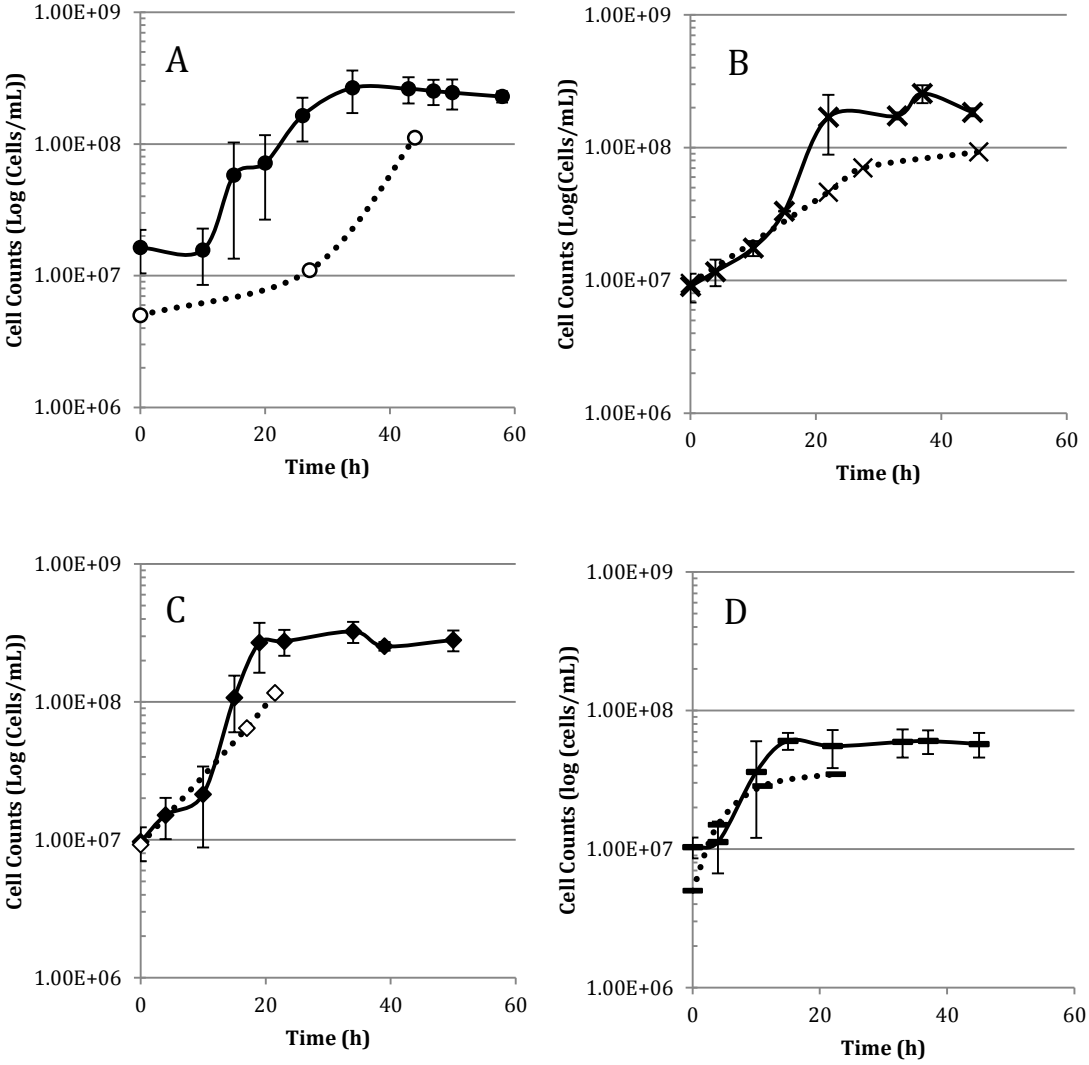


Figure 6. Small-scale and large-scale growth for *C. kristjanssonii* on four soluble substrates. Small-scale growth is shown in solid lines while large-scale growth is shown in dotted lines. (A) shows growth on glucose, small scale (—●—) and large scale (·○·). (B) shows growth on xylose, small scale (—✕—) and large scale (·✕·). (C) shows growth on cellobiose, small scale (—◆—) and large scale (·◇·). (D) shows growth on yeast extract only, small scale (——) and large scale (·▬·). The curves display an average of counts done in triplicates. Note that small-scale cultures were 50 mL while large-scale cultures were 1 L in volume.

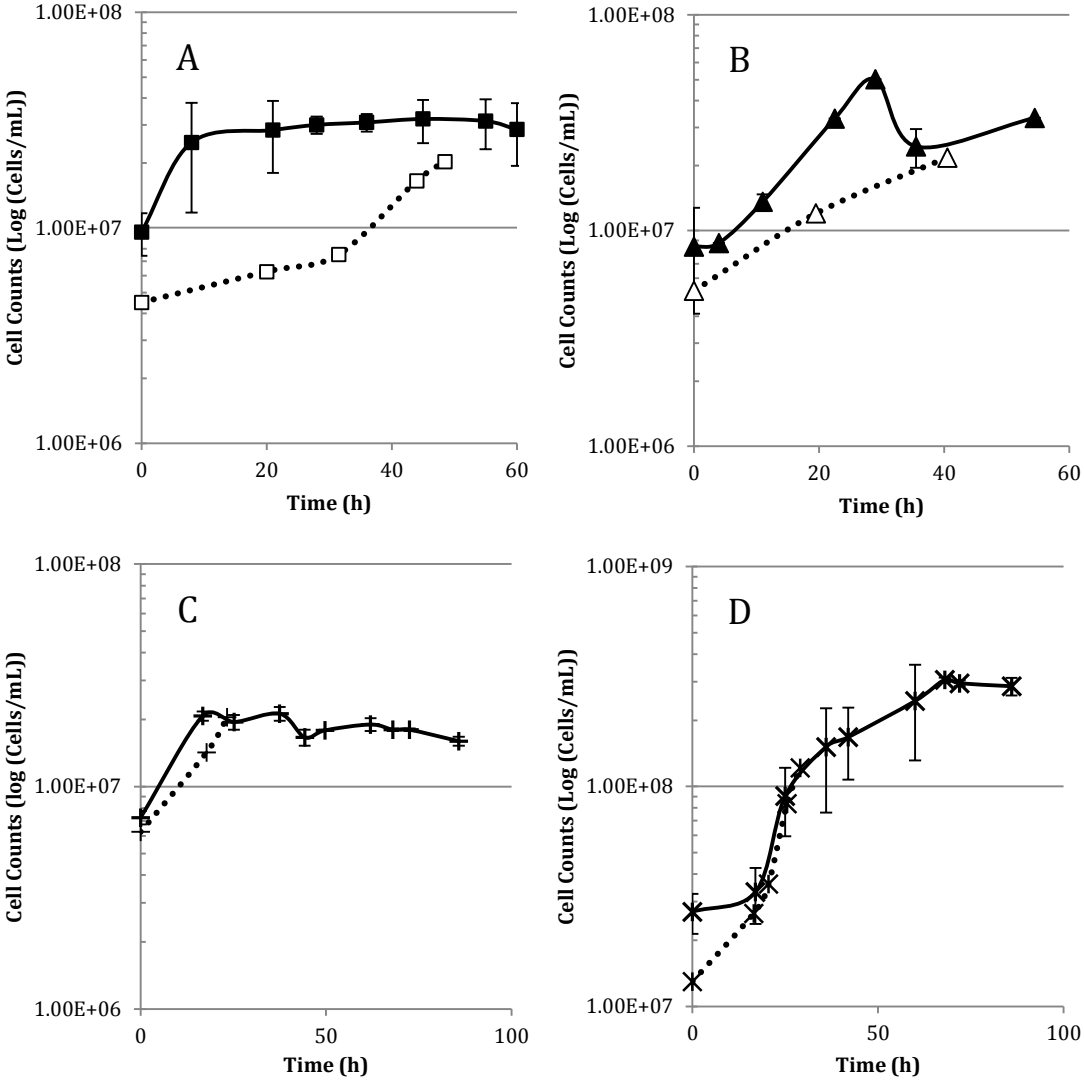


Figure 7. Small-scale and large-scale growth for *C. kristjanssonii* on four insoluble substrates. Small-scale growth is shown in solid lines while large-scale growth is shown in dotted lines. (A) shows growth on Avicel, small scale (▬■▬) and large scale (▪□▪). (B) shows growth on CMC, small scale (▬▲▬) and large scale (▪△▪). (C) shows growth on switchgrass, small scale (▬╋▬) and large scale (▪╋▪). (D) shows growth on xylan, small scale (▬✳▬) and large scale (▪✳▪). The curves display an average of counts done in triplicates. Note that small-scale cultures were 50 mL while large-scale cultures were 1 L in volume.

Both *C. saccharolyticus* and *C. kristjanssonii* were observed to be capable of growing on all eight substrates: glucose, xylose, cellobiose, Avicel, switchgrass, CMC, xylan, and yeast extract alone, with the exception of *C. kristjanssonii* which is not able to grow on CMC without supplementing with yeast extract. Details of the growth media composition can be found in Appendix A. A second exponential growth phase could be observed when *C. saccharolyticus* was grown on glucose, xylose, cellobiose, Avicel, CMC, and switchgrass and when *C. kristjanssonii* was grown on xylose, Avicel, and xylan. *C. saccharolyticus* reached a highest cell density of about 3.8×10^8 cells/mL when grown on xylan but the best growth was on xylose, giving a cell density of about 2.5×10^8 cells/mL and a generation time of 4.21 hours (see Table 1). *C. kristjanssonii* grew the best on cellobiose with cell densities reaching about 3.2×10^8 cells/mL with a generation time of 2.5 hours (see Table 1). The poorest growth for both organisms was observed on washed switchgrass *as C. saccharolyticus* only reached cell densities of about 3.8×10^7 cells/mL and *C. kristjanssonii* reached about 2.1×10^8 cells/mL. Interestingly, for both organisms, growth on yeast extract alone gave the fastest growth rates but resulted in relatively low final cell densities and conversely, growth on xylan produced the slowest growth rates but one of the highest cell densities compared to other growth conditions (see Figure 8).

After confirming that growth was possible on each of these substrates as well as determining the characteristic growth rates and phases, further experiments could be performed including growth in large-scale batch cultures for RNA extraction and proteomic sample preparation as well as end product detection using GC and HPLC at predetermined time intervals,

Table 1. Growth rate constants, generation times, and final cell densities of *C. saccharolyticus* and *C. kristjanssonii* grown on eight substrates

Growth Substrates	Small Scale 1st phase		Small Scale 2nd phase		Small Scale Final Cell Density	Large Scale		End Product Detection		EPD Final Cell Density
	k	***g***	***k***	***g***	**cells/mL**	***k***	***g***	***k***	***g***	**cells/mL**
C. saccharolyticus										
Glucose	0.15	4.51	0.05	13.76	2.88×10^{8}	0.10	7.22	0.09	7.91	1.91×10^{8}
Xylose	0.16	4.21	0.07	9.37	2.50×10^{8}	0.10	6.92	0.12	5.71	1.78×10^{8}
Cellobiose	0.16	4.25	0.04	17.22	2.18×10^{8}	0.07	10.53	0.08	8.61	1.87×10^{8}
Yeast extract	0.23	3.01	-	-	8.16×10^{7}	0.15	4.51	0.26	2.66	6.25×10^{7}
Avicel + YE	0.05	14.45	0.02	37.82	6.56×10^{7}	0.05	14.45	0.03	22.50	5.29×10^{7}
CMC + YE	0.06	10.70	0.01	112.69	4.18×10^{7}	0.07	9.65	0.06	12.23	6.99×10^{7}
Switchgrass + YE	0.01	51.76	0.05	13.33	3.76×10^{7}	0.04	17.03	0.03	21.02	4.70×10^{7}
Xylan	0.03	22.07	-	-	3.57×10^{8}	0.06	12.44	0.08	9.07	2.54×10^{8}
C. kristjanssonii										
Glucose	0.13	5.44	-	-	2.67×10^{8}	0.14	4.80	0.12	5.89	2.09×10^{8}
Xylose	0.13	5.41	0.09	7.82	2.55×10^{8}	0.09	7.75	0.11	6.46	1.47×10^{8}
Cellobiose	0.28	2.50	-	-	3.24×10^{8}	0.12	5.61	0.14	5.04	2.50×10^{8}
Yeast extract	0.19	3.70	-	-	6.04×10^{7}	0.19	3.65	0.18	3.87	5.69×10^{7}
Avicel + YE	0.09	7.49	0.01	136.17	3.20×10^{7}	0.06	10.79	0.03	21.02	3.93×10^{7}
CMC + YE	0.08	9.03	-	-	5.03×10^{7}	0.03	20.00	0.07	10.59	5.61×10^{7}
Switchgrass + YE	0.05	15.06	-	-	2.13×10^{7}	0.05	12.76	0.05	13.57	4.56×10^{7}
Xylan	0.13	5.53	0.02	32.37	3.06×10^{8}	0.13	5.53	0.13	5.53	1.98×10^{8}

This table displays the growth rate constants (*k*) in number of generations per hour, generation times (*g*) in hours and final cell densities in cells/mL, for *C. saccharolyticus* and *C. kristjanssonii* grown on eight substrates: glucose, xylose, cellobiose, yeast extract, Avicel, CMC, switchgrass, and xylan. Two growth phases were observed in several conditions as either the culture has a long lag or a slow transition into stationary phase. Final cell densities were not recorded here for large scale cultures because these were quenched during the mid-log phase of growth for further sample processing, whereas small scale cultures were observed into stationary phase. Note: EPD stands for "end product detection".

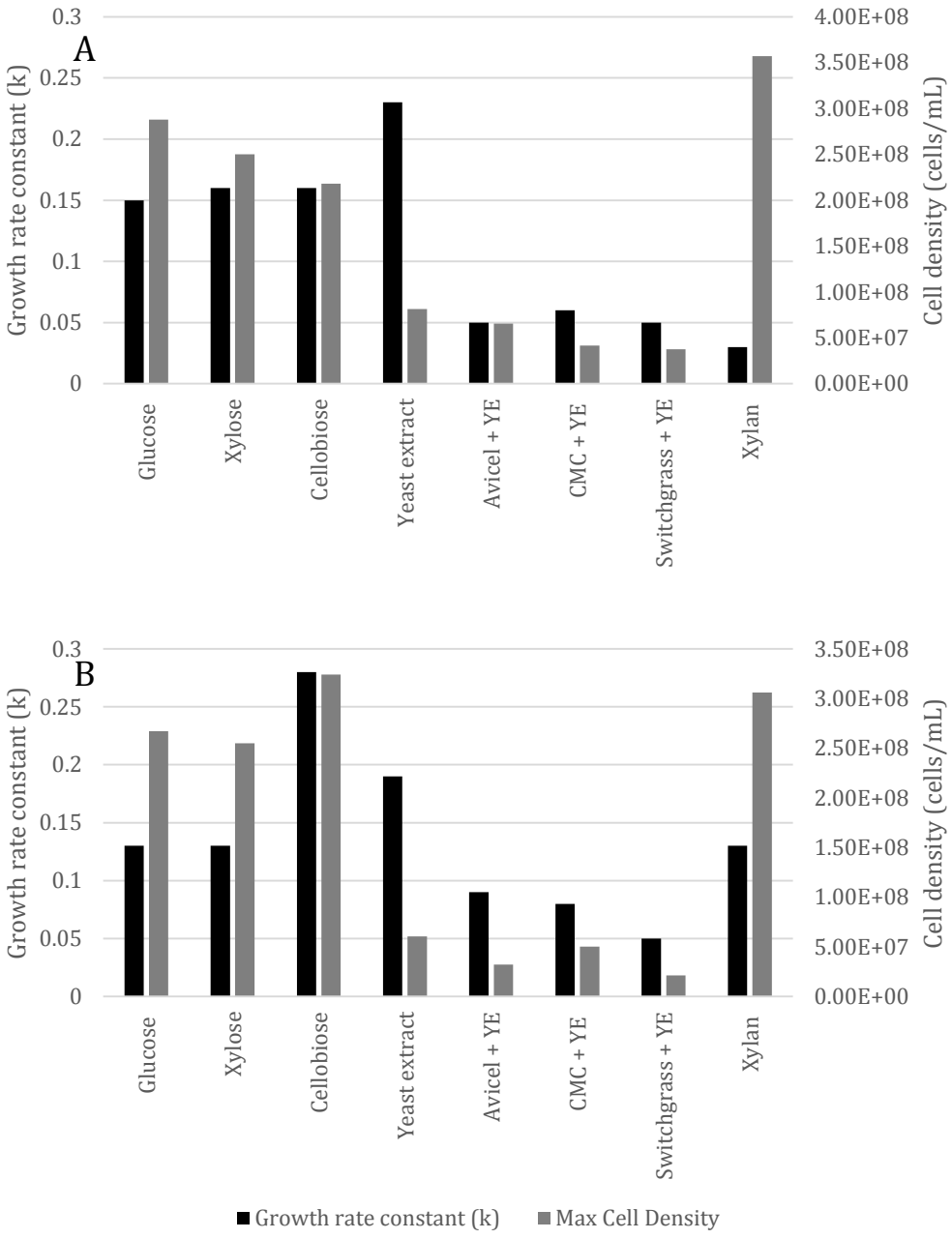


Figure 8. Growth rate constants and maximum cell densities for *Caldicellulosiruptor* spp. grown on cellulosic substrates. These graphs help to visualize the growth characteristics of (A) *C. saccharolyticus* and (B) *C. kristjanssonii* when they are grown in glucose, xylose, cellobiose, yeast extract, Avicel, CMC, switchgrass and xylan.

3.2 End product detection using GC and HPLC

To measure the amounts of end products produced in low-nutrient media and get an idea of which pathways are being used for carbohydrate metabolism, end products were examined using GC and HPLC. The techniques and parameters used for measuring components in the liquid media were as described in section 2.4.1 whereas for gaseous samples, it is stated in section 2.4.2. The standard equation calculated for hydrogen was $y=11956x$ with an R^2 value of 0.997 while the standard equation for carbon dioxide was $y=(3\times10^6)x$ with an R^2 value of 0.998. Please see Figures A and B in Appendix B for the hydrogen and carbon dioxide standard curves and Figure C for the HPLC standard curves.

The following graphs compare the growth corresponding to the substrate utilization and the detectable end products generated over time. Figures 9 and 10 demonstrate the metabolic flux for *C. saccharolyticus* while Figures 11 and 12 show this for *C. kristjanssonii* grown on soluble and insoluble substrates respectively. The metabolites begin accumulating as growth proceeds into exponential phase and greatest amounts of end products including acetate, lactate, hydrogen and carbon dioxide are accumulated as the cultures reach stationary phase or a few hours after that.

In terms of potential biofuels, the highest hydrogen yield from *C. saccharolyticus* was only about 0.75 mol/mol xylose and for *C. kristjanssonii*, it was about 0.52 mol/mol xylose. Ethanol was only detected when both organisms were grown on xylan with *C. saccharolyticus* producing up to 0.847 mM and *C. kristjanssonii* producing up to 0.789 mM. However, it should be noted that the detection limit for ethanol on the HPLC was about 0.5 mM, so the other conditions may have produced ethanol but at undetectable levels.

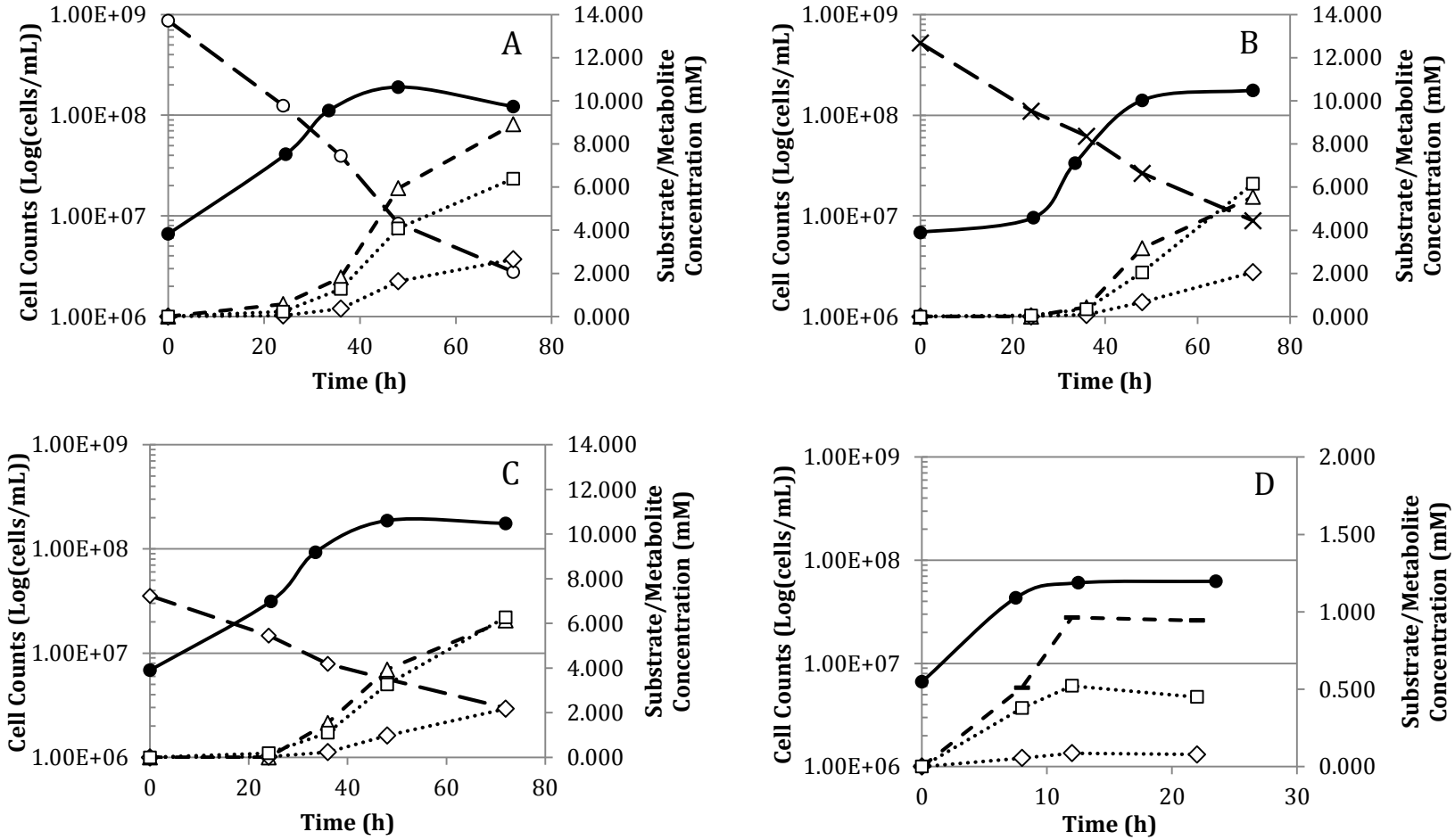


Figure 9. *C. saccharolyticus* growth, soluble substrate consumption, and end product formation. Associated growth curves are shown in solid lines (—●—). The graphs show growth and end product detection on (A) glucose, (B) xylose, (C) cellobiose, and (D) yeast extract. Dashed lines show the substrate consumption of glucose (– ○ –), xylose (– ✕ –), and cellobiose (– ◇ –), as well as soluble products acetate (– △ –), and lactate (– ▬ –). End products detected in the gas phase are shown in dotted lines including hydrogen (··□··) and carbon dioxide (··◇··). The curves represent an average of counts and chromatographic runs done in duplicates.

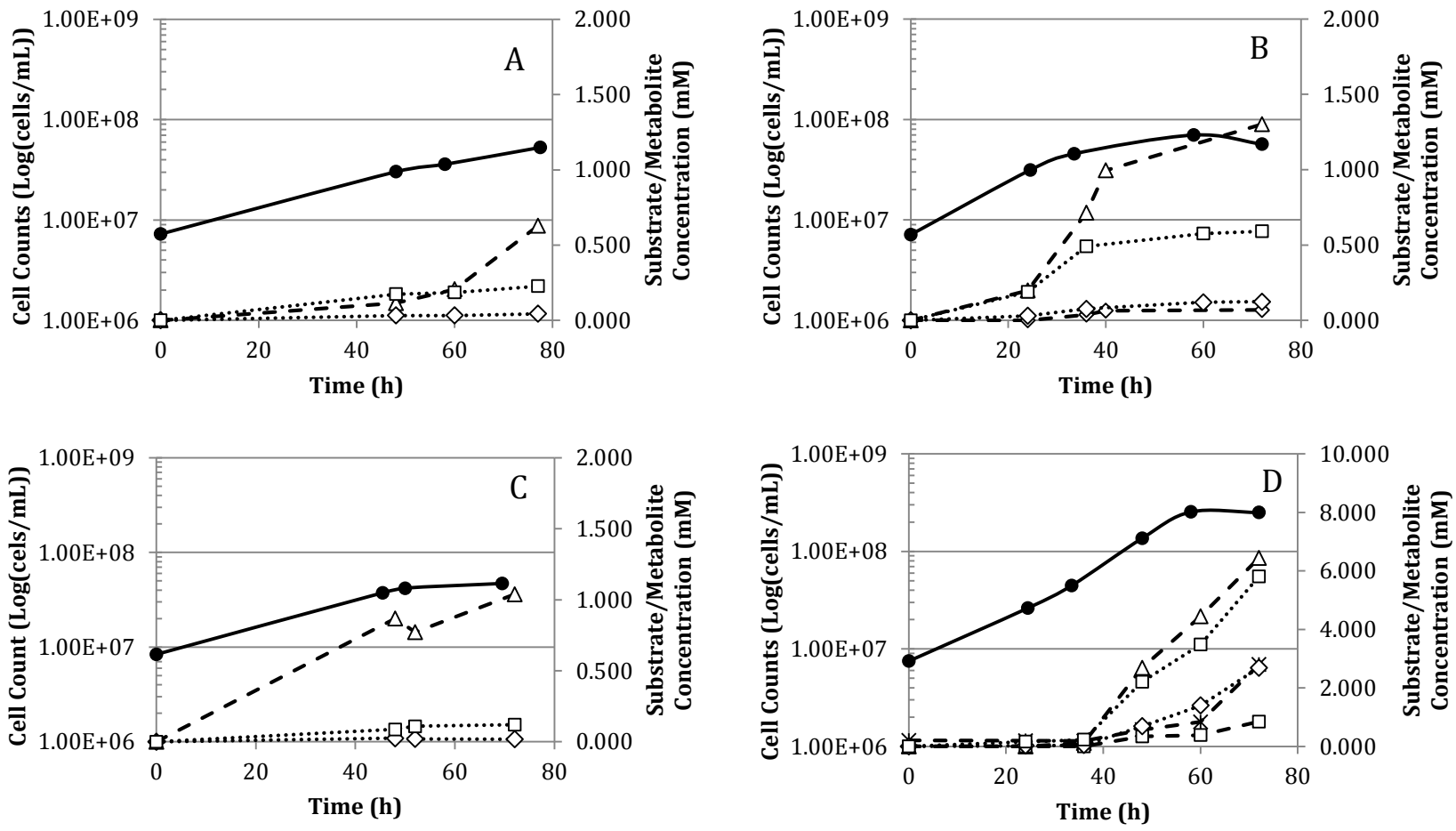


Figure 10. *C. saccharolyticus* growth, insoluble substrate utilization, and end product formation. Associated growth curves are shown in solid lines (—●—). The graphs show growth and end product detection on (A) Avicel, (B) CMC, (C) switchgrass, and (D) xylan. Dashed lines show compounds found in the liquid phase including cellobiose (— ◇ —), acetate (— △ —), citric acid (— ∗ —), and ethanol (— □ —). End products detected in the gas phase are shown in dotted lines including hydrogen (··□··) and carbon dioxide (··◇··). The curves represent an average of counts and chromatographic runs done in duplicates.

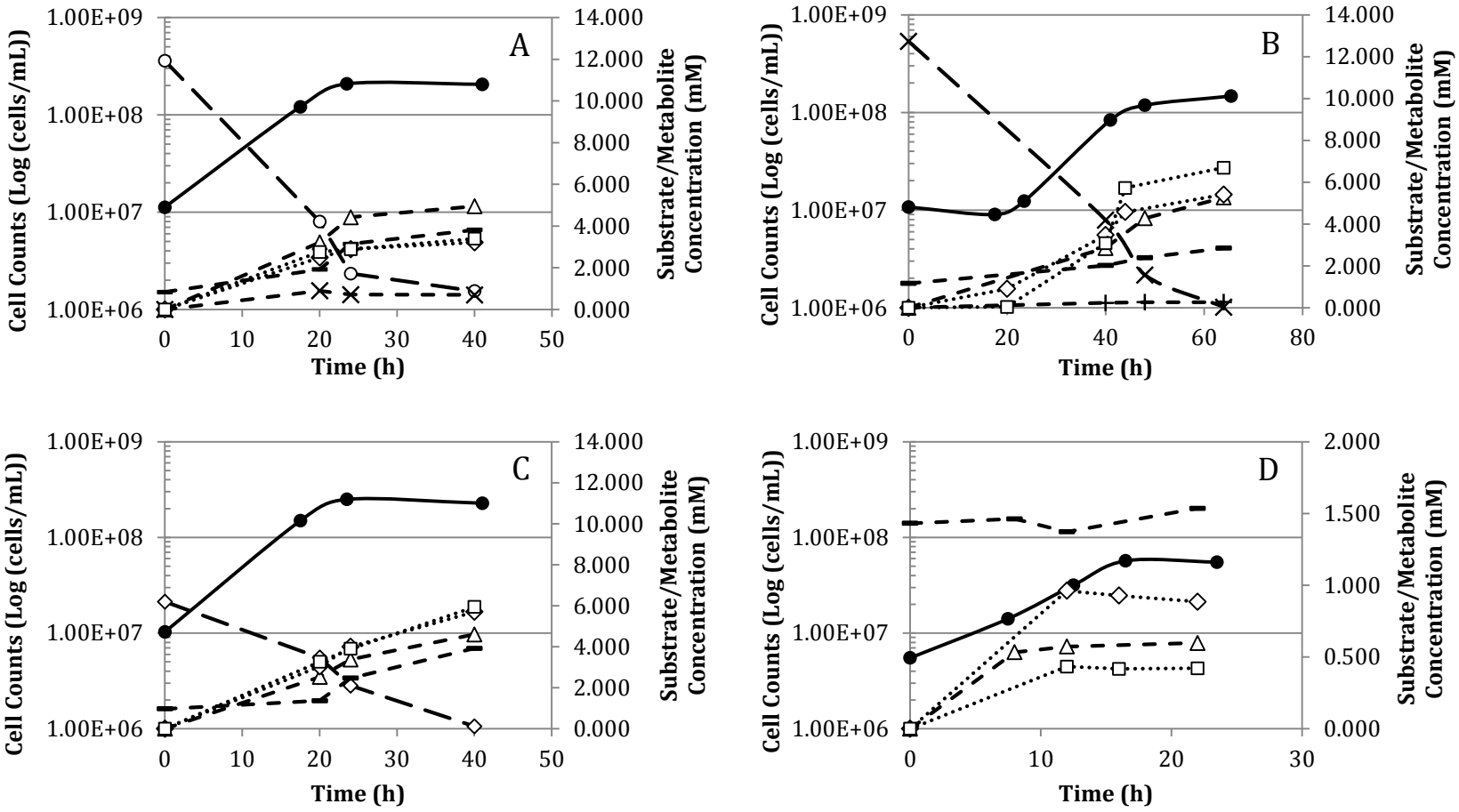


Figure 11. *C. kristjanssonii* growth, soluble substrate consumption, and end product formation. Associated growth curves are shown in solid lines (—●—). The graphs show growth and end product detection on (A) glucose, (B) xylose, (C) cellobiose, and (D) yeast extract. Dashed lines show the substrate consumption and production of glucose (— ○ —), xylose (— ✕ —), cellobiose (— ◇ —), arabinose (— + —), acetate (— △ —), and lactate (— ■ —). End products detected in the gas phase are shown in dotted lines including hydrogen (··□··) and carbon dioxide (··◇··). The curves represent an average of counts and chromatographic runs done in duplicates.

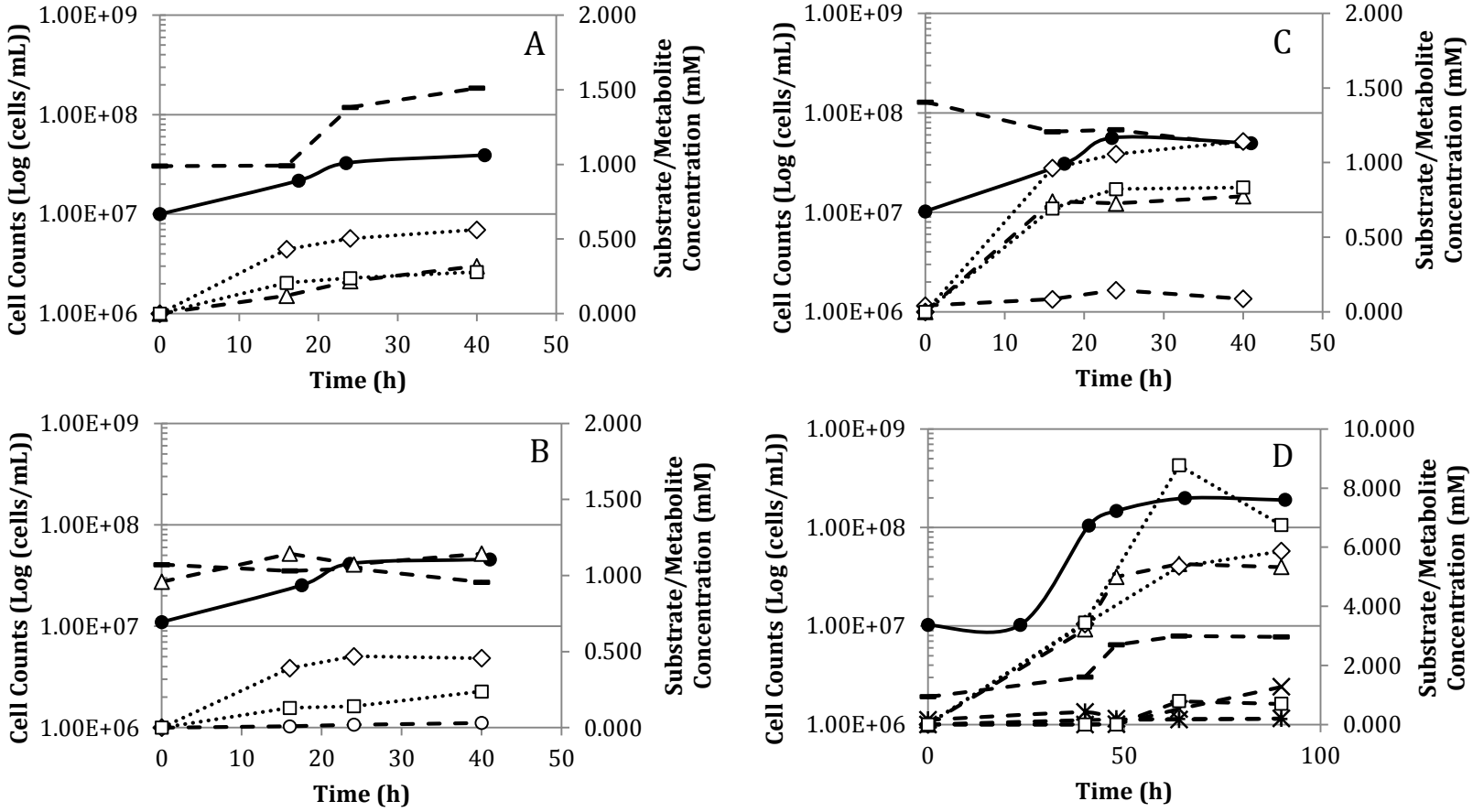


Figure 12. ***C. kristjanssonii*** **growth, insoluble substrate utilization, and end product formation.** Associated growth curves are shown in solid lines (—●—). The graphs show growth and end product detection on (A) Avicel, (B) CMC, (C) switchgrass, and (D) xylan. Dashed lines show the substrate consumption and production of glucose (– ○ –), xylose (– ✱ –), cellobiose (– ◇ –), arabinose (– + –), acetate (– △ –), lactate (– ■ –), citric acid (– ∗ –) and ethanol (– □ –). End products detected in the gas phase are shown in dotted lines including hydrogen (··□··) and carbon dioxide (··◇··). The curves represent an average of counts and chromatographic runs done in duplicates.

Table 1 summarizes the growth rates, generation times and final cell densities recorded for *C. saccharolyticus* and *C. kristjanssonii* when grown in small scale, large scale, and during end product detection. As seen in Figure 8, the highest cell density for *C. saccharolyticus* was reached on xylan and for *C. kristjanssonii*, it was when it was on cellobiose. Growth on cellobiose also gave *C. kristjanssonii* the fastest growth rate, but for *C. saccharolyticus*, growth on xylan gave the slowest growth rate, whereas yeast extract was the fastest.

As for the metabolites produced, Table 2 below profiles the major end products observed when the cultures reach stationary phase. The most acetate and hydrogen and carbon dioxide was observed when both organisms were grown on glucose, xylose, cellobiose, and xylan (see Table 2 above). *C. saccharolyticus* produced the most acetate (about 0.45 mmol) when grown on glucose and *C. kristjanssonii* produced about 0.26 mmol acetate when grown on xylose. Lactate was not detected with *C. saccharolyticus* except for growth on yeast extract, where about 0.05 mmol of lactate was detectable, since this culture was monitored for several hours after it reached stationary phase. Interestingly, arabinose was detected in small amounts when *C. kristjanssonii* was grown on xylose and xylan, suggesting that it produces enzymes to isomerize these C5 sugars. Citric acid was also detected when *C. saccharolyticus* and *C. kristjanssonii* were grown on xylan, but it is suspected that this is a compound called ferulic acid rather than citric acid.

Table 2. End product summary for *C. saccharolyticus* and *C. kristjanssonii* grown on eight substrates

	Amount of end products detected (mM)					
Substrate	**Acetate**	**Lactate**	**Arabinose**	**Ethanol**	**H_2**	**CO_2**
C. saccharolyticus						
Control	-	-	-	-	-	0.0004
Glucose	8.913	-	-	-	6.392	2.654
Xylose	5.553	-	-	-	6.163	2.059
Cellobiose	6.136	-	-	-	6.274	2.177
Avicel + YE	0.628	-	-	-	0.227	0.044
CMC + YE	1.302	-	-	-	0.591	0.124
Switchgrass + YE	1.039	-	-	-	0.120	0.018
Xylan	6.438	-	-	0.847	5.805	2.685
Yeast Extract	-	0.945	-	-	0.450	0.078
C. kristjanssonii						
Control	-	0.732	-	-	-	0.0073
Glucose	4.951	3.822	-	-	3.396	3.216
Xylose	5.271	2.851	0.262	-	6.691	5.412
Cellobiose	4.603	3.921	-	-	5.938	5.700
Avicel + YE	0.319	1.512	-	-	0.279	0.562
CMC + YE	0.774	1.115	-	-	0.833	1.142
Switchgrass + YE	1.144	0.957	-	-	0.238	0.456
Xylan	5.326	2.965	0.186	0.789	6.748	5.863
Yeast Extract	0.598	1.536	-	-	0.422	0.886

Note: the control contains the respective medium without any carbon sources and was inoculated with the same amount of inoculum as all of the other conditions. Remaining substrates represent the amounts of the corresponding carbon substrate that was left over after the cultures reached stationary phase. The detection limit for ethanol is about 0.5 mM, thus other conditions that produced ethanol in small amounts may not be detectable.

Table 3. Carbon balance for *C. saccharolyticus* and *C. kristjanssonii* grown on three soluble sugars

	Compound Concentration (mM)								
Substrate	**Substrate Used**	**Acetate**	**Lactate**	**Carbon dioxide**	**Dissolved CO_2**	**Dissolved HCO_3**	**Cell Biomass**	**Total Output**	**Yield**
C. saccharolyticus									
Glucose	11.66	8.92	0	2.651			0.20 g		
Carbon Content	**69.96**	**17.84**	**0**	**2.651**	**0.810**	**0.034**	**50 mM**	**71.335**	**101.97%**
Xylose	8.26	5.56	0	2.055			0.13 g		
Carbon Content	**41.3**	**11.12**	**0**	**2.055**	**0.690**	**0.029**	**32 mM**	**45.894**	**111.12%**
Cellobiose	5.04	6.14	0	2.178			0.17 g		
Carbon Content	**60.48**	**12.28**	**0**	**2.178**	**0.719**	**0.030**	**42 mM**	**57.207**	**94.59%**
C. kristjanssonii									
Glucose	11.06	4.96	3.82	3.219			0.19 g		
Carbon content	**66.36**	**9.92**	**11.47**	**3.219**	**0.397**	**0.167**	**47 mM**	**72.173**	**108.76%**
Xylose	12.76	5.28	2.85	5.416			0.18 g		
Carbon content	**63.80**	**10.56**	**8.55**	**5.416**	**0.632**	**0.265**	**45 mM**	**70.423**	**110.38%**
Cellobiose	6.08	4.60	3.92	5.700			0.20 g		
Carbon content	**72.96**	**9.20**	**11.76**	**5.700**	**0.671**	**0.282**	**50 mM**	**77.613**	**106.38%**

The carbon content listed in bold corresponds to the substrate in each row above. The carbon content was calculated by multiplying the substrate or product concentration by the moles of carbon in its chemical composition. The substrate used is the difference between the amount of substrates added and the substrates remaining after fermentation. The CO_2 dissolved in the liquid phase was calculated using Henry's law, and the HCO_3 was calculated using the Henderson-Hasselbach equation. The carbon concentration in the cell biomass was estimated assuming that 30% of the pellet was dry weight (Bakken & Olsen, 1983) and about 50% of this pellet was composed of carbon (Bratbak & Dundas, 1984).

From the end products detected, a carbon balance was also calculated for the soluble sugars, as summarized in Table 3. Carbon balances were shown for the simple sugars because these conditions had a defined carbon content. All of the carbon balances were near 100% with some that are over 100% because the biomass weights may be overestimated since they are wet pellets. The CO_2 dissolved in the liquid phase was calculated using Henry's law, and the HCO_3 was calculated using the Henderson-Hasselbach equation.

Considering growth on complex substrates is made possible by cellulase enzymes first degrading the cellulosic matter, enzyme assays were done to examine cellulase activity.

3.3 Enzyme activities of native cellulase and xylanase

Enzyme assays were performed to confirm whether or not functional cellulases are produced by *C. saccharolyticus* and *C. kristjanssonii*. More specifically, endoglucanase, exoglucanase, and xylanase activity was targeted and measured through the DNS reducing sugar detection method (see sections 2.6.3 and 2.6.4) and these enzymes were tested to see if they predominantly act intracellularly or if they are secreted outside of the cell.

Table 4 summarizes the percentage of total activity observed from the extracellular (supernatant) and intracellular (CFE) environments. It should be noted that the total activity stated here does not include any cellulases that may be bound to the cell wall or cell surface. Surprisingly, when looking at the total activity, the majority of endoglucanase activity was observed from the cell-free extract in both organisms, with the exception of growth on Avicel. Growth on glucose resulted in 85% and 98% of the total endoglucanase activity occurring intracellularly in *C. saccharolyticus* and *C. kristjanssonii* respectively. *C. saccharolyticus* also had 98% and 97% of the total endoglucanase activity occurring inside the cells when grown on xylose and xylan respectively. Exoglucanase activity differed as *C. saccharolyticus* generally showed more total activity extracellularly while *C. kristjanssonii* generally had more intracellularly. As for xylanase activity, *C. saccharolyticus* generally showed more total activity intracellularly, especially with glucose (99%) and xylose (96%), while *C. kristjanssonii* had much variation and it was almost split between the supernatant and the CFE samples. Figures 13 and 14 have stacked column graphs to help visualize the amounts of total activity occurring inside and outside the cell for *C. saccharolyticus* and *C. kristjanssonii*, respectively.

Table 4. Enzyme activities for exoglucanase, endoglucanase, and xylanase observed in *C. saccharolyticus* and *C. kristjanssonii*

		C. saccharolyticus					
		Supernatant			**Cell-free Extract**		
	Growth Condition	Specific Activity (U/mg)	Total Activity (U)	Percentage of Total Activity in SN+CFE	Specific Activity (U/mg)	Total Activity (U)	Percentage of Total Activity in SN+CFE
Exoglucanase	Glucose	0.169	0.056	45%	0.072	0.069	55%
	Avicel	1.109	0.060	56%	0.035	0.046	44%
	CMC	1.625	0.524	87%	0.050	0.081	13%
	Xylose	0.209	0.067	92%	0.009	0.006	8%
	Xylan	4.911	1.391	93%	0.028	0.109	7%
Endoglucanase	Glucose	0.587	0.115	15%	0.278	0.651	85%
	Avicel	0.409	0.021	59%	0.020	0.015	41%
	CMC	1.663	0.540	23%	1.181	1.801	77%
	Xylose	0.012	0.004	2%	0.278	0.195	98%
	Xylan	0.017	0.026	3%	0.245	0.927	97%
Xylanase	Glucose	0.082	0.016	1%	0.640	1.135	99%
	Avicel	28.391	1.496	70%	0.942	0.649	30%
	CMC	4.648	1.771	58%	0.848	1.266	42%
	Xylose	0.075	0.024	4%	0.812	0.599	96%
	Xylan	0.592	0.168	18%	0.228	0.767	82%
		C. kristjanssonii					
Exoglucanase	Glucose	0.424	0.142	20%	0.262	0.564	80%
	Avicel	5.470	1.692	69%	0.763	0.777	31%
	CMC	0.902	0.532	40%	0.565	0.789	60%
	Xylose	0.039	0.075	11%	0.091	0.639	89%
	Xylan	0.270	0.189	15%	0.286	1.075	85%
Endoglucanase	Glucose	0.018	0.011	2%	0.150	0.470	98%
	Avicel	1.034	0.312	92%	0.011	0.026	8%
	CMC	1.465	0.848	29%	1.778	2.119	71%
	Xylose	0.069	0.132	10%	0.173	1.219	90%
	Xylan	0.287	0.201	41%	0.085	0.289	59%
Xylanase	Glucose	2.576	0.860	60%	0.241	0.564	40%
	Avicel	5.299	1.536	80%	0.600	0.395	20%
	CMC	0.024	0.032	7%	0.340	0.406	93%
	Xylose	1.713	3.233	44%	0.585	4.109	56%
	Xylan	6.454	8.046	59%	2.382	5.517	41%

This table summarizes all of the total enzyme activities as well as the specific activities calculated with respect to the amounts of protein present in each sample. Activities from SN and CFE are shown, but do not include cellulases bound to the cell wall or cell surface. Note that the enzyme unit (U) represents μmol/min.

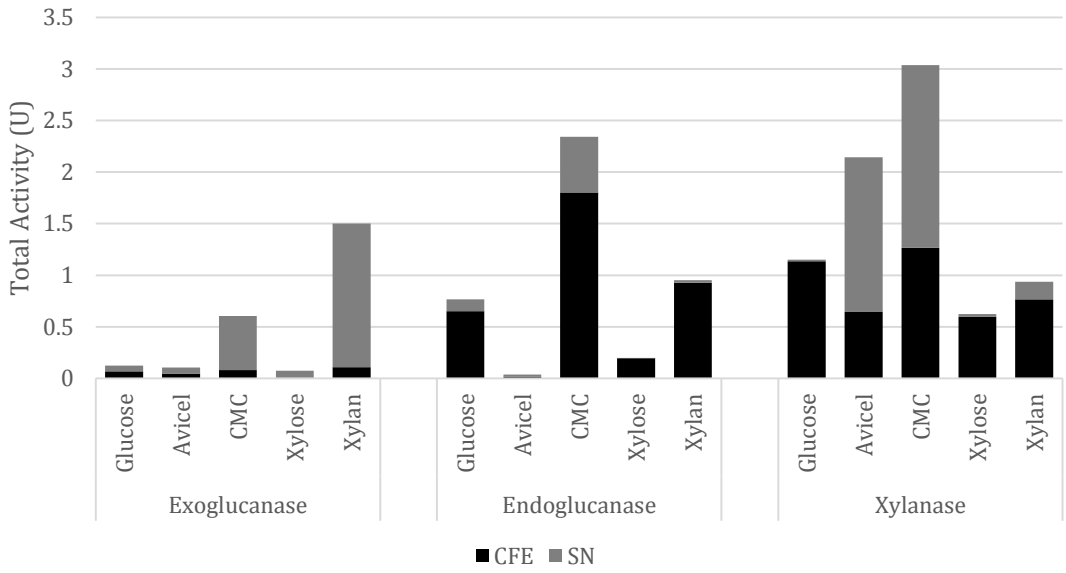


Figure 13. Total activity of cellulases in *C. saccharolyticus* cell-free extracts and supernatant. This stacked column graph shows the total activity of exoglucanases, endoglucanases, and xylanases measured in *C. saccharolyticus* samples from both intracellular (CFE) and extracellular (SN) environments. Note that CFE stands for cell-free extract and SN stands for supernatant.

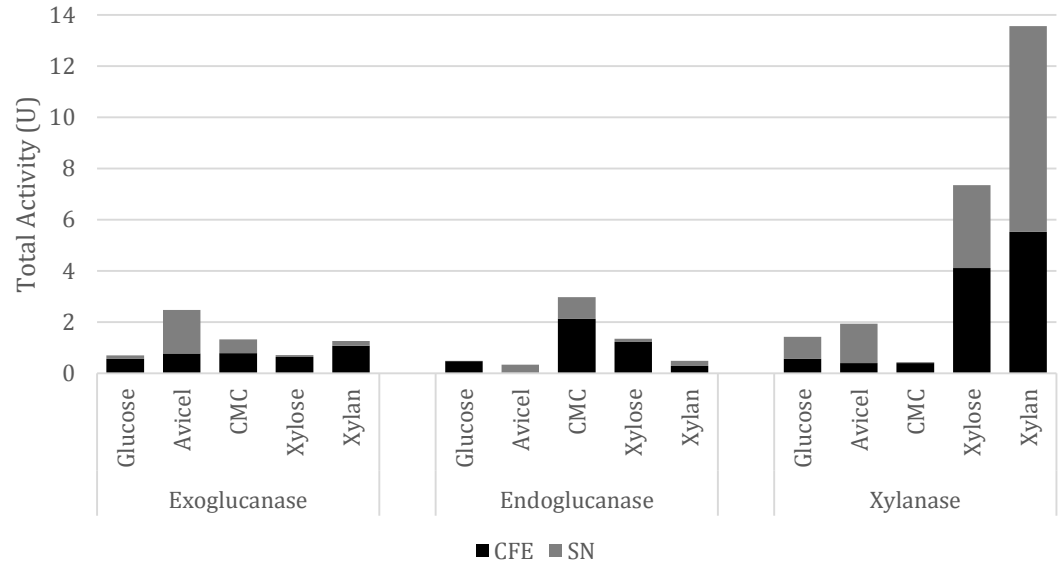


Figure 14. Total activity of cellulases in *C. kristjanssonii* cell-free extracts and supernatant. This stacked column graph shows the total activity of exoglucanases, endoglucanases, and xylanases measured in *C. saccharolyticus* samples from both intracellular (CFE) and extracellular (SN) environments. Note that CFE stands for cell-free extract and SN stands for supernatant.

When considering the amounts of protein in each sample, the proportion of active cellulase enzymes in relation to total protein is much higher in the supernatant than the CFE, suggesting that the specific activity of exoglucanases, endoglucanases, and xylanases are greater extracellularly. Figures 15 and 16 show that more cellulase specific activity occurs in the supernatant portion of the culture for both *C. saccharolyticus* and *C. kristjanssonii*.

The highest exoglucanase specific activity for *C. saccharolyticus* was observed in the supernatant at 4.9 U/mg when grown on xylan. The highest endoglucanase specific activity for *C. saccharolyticus* was similar but slightly higher in the CFE at 1.8 U/mg compared to the supernatant which had 1.7 U/mg. For xylanase activity, the highest specific activity was observed from the supernatant of *C. saccharolyticus* grown on Avicel, reaching a substantial 28.4 U/mg.

For *C. kristjanssonii*, the greatest specific activities were recorded from the growth condition that corresponded to the substrate that they react on. For example, exoglucanase acts on Avicel, endoglucanase acts on CMC, and xylanase acts on xylan. *C. kristjanssonii* had the greatest exoglucanase specific activity (5.5 U/mg) in the supernatant sample when grown on Avicel and xylanase specific activity (6.5 U/mg) in the supernatant portion of the xylan culture as well. Once again, the endoglucanase specific activity was similar between the inside and outside of the cell where the CFE (1.8 U/mg) was slightly higher than the supernatant (1.5 U/mg).

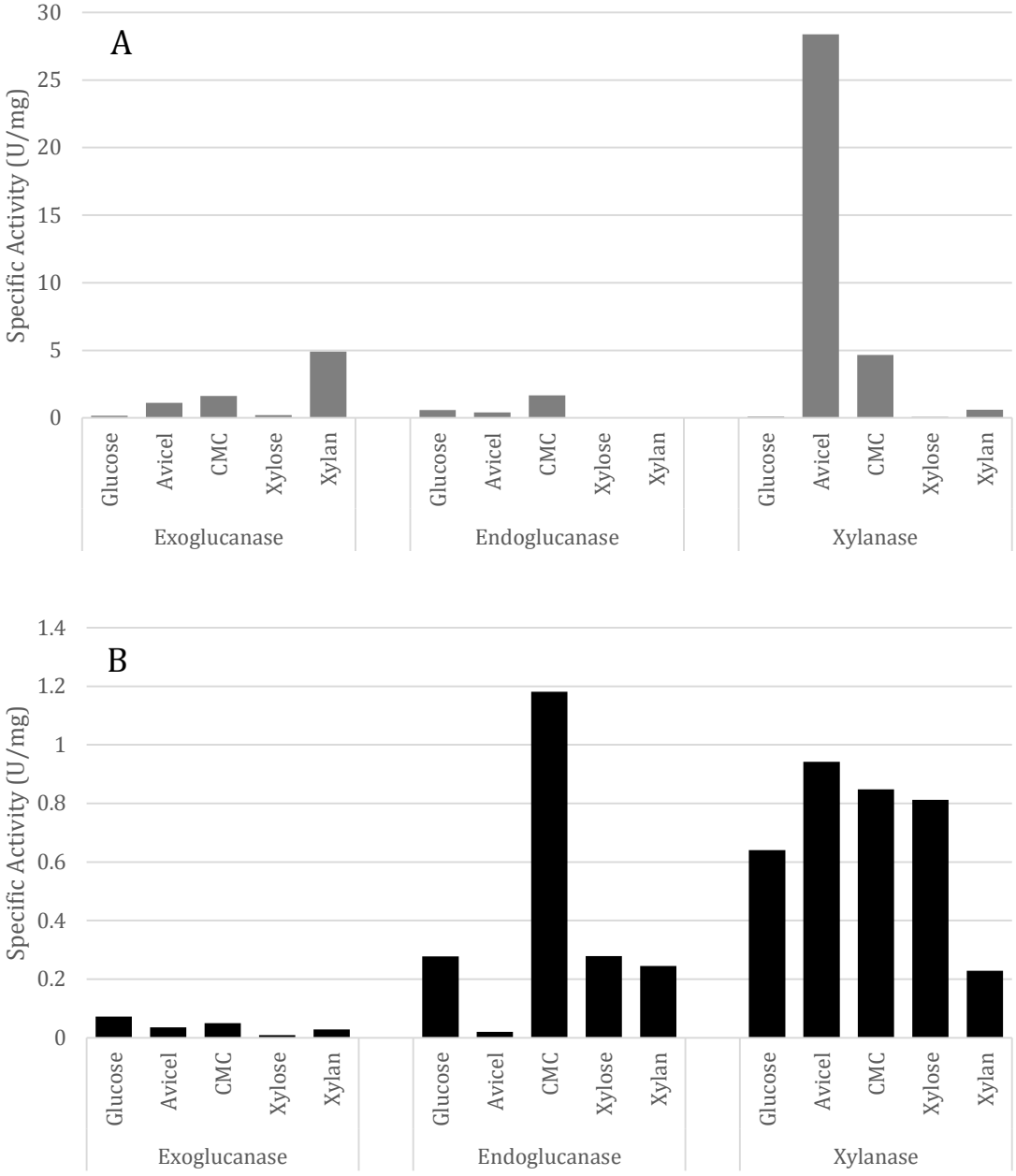


Figure 15. Exoglucanase, endoglucanase, and xylanase specific activities of *C. saccharolyticus*. These bar graphs summarize the specific activities observed from culture (A) supernatants and (B) cell-free extracts when *C. saccharolyticus* was grown on five substrates: glucose, xylose, Avicel, CMC and xylan.

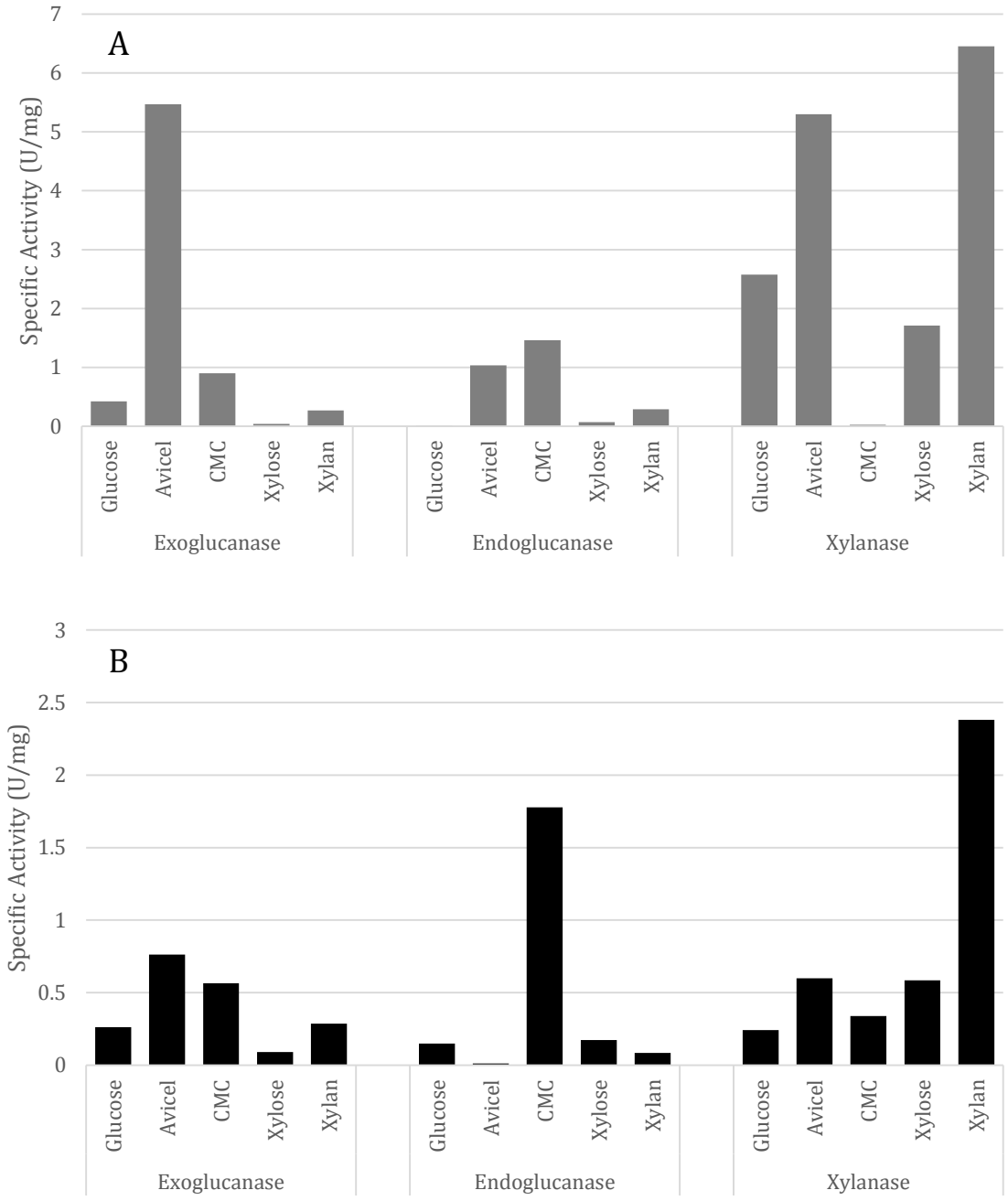


Figure 16. Exoglucanase, endoglucanase, and xylanase specific activities of *C. kristjanssonii*. These bar graphs summarize the specific activities observed from culture (A) supernatants and (B) cell-free extracts when *C. kristjanssonii* was grown on five substrates: glucose, xylose, Avicel, CMC and xylan.

3.4 Genome analysis

As a view of the currently known enzymes related to carbohydrate hydrolysis and metabolism, a thorough search was conducted through the NCBI database with results tabulated here. Table 5 lists the locus tags of genomic sequences that were annotated as some of the main enzymes of interest in this study. This table shows some key enzymes found on NCBI while a metabolic map with locus tags and proteins expressed will be presented in the discussion section. The following key enzymes were searched for: β-glucosidase, xylanases, ABC and PTS transporters, phosphofructokinase, KDPG aldolase, pyruvate kinase, transketolase, transaldolase, pyruvate ferredoxin oxidoreductase, hydrogenase, lactate dehydrogenase, and alcohol dehydrogenase. Multiple loci were found for ABC transporters, pyruvate ferredoxin oxidoreductase, and alcohol dehydrogenases, indicating that there are many different channels for the uptake of sugars, both organisms are able to ferment the sugars, and that there may be several different pathways possible for producing ethanol.

The locus tags listed in Table 5 show the theoretical presence of enzymes that hydrolyze cellulosic materials and transport and ferment the carbohydrates in each organism. To experimentally prove that these theoretical enzymes exist and are in fact produced by each organism, biomass samples were sent to have proteomic and RNA-seq analyses done.

Table 5. Genomic sequence locus tags for key enzymes in carbohydrate metabolic pathways found in *C. saccharolyticus* and *C. kristjanssonii*

Key Enzymes	Locus Tags	
	C. saccharolyticus	*C. kristjanssonii*
Cellulase	Csac_0678 Csac_1079	Calkr_2522
β-glucosidase	Csac_1089	Calkr_2513
Xylanase	N/A	Calkr_0572
Transporters (ABC)	Csac_0238 Csac_0240 Csac_0297 Csac_0358 Csac_0392 Csac_0681 Csac_2324 Csac_2505 Csac_2696	Calkr_2413 Calkr_2435
PTS Transporters	Csac_2439	Calkr_0283
Phosphofructokinase	Csac_1830 Csac_2366	Calkr_0133 Calkr_0849 Calkr_1246
KDPG aldolase	Csac_0354	Calkr_1981
Pyruvate kinase	Csac_1831	Calkr_1247
Transketolase	Csac_0874 Csac_1351 Csac_1352	Calkr_0504 Calkr_0505 Calkr_1273 Calkr_2357
Transaldolase	Csac_2036	Calkr_1474
Pyruvate ferredoxin oxidoreductase	Csac_1459 Csac_1460 Csac_1549 Csac_1551 Csac_2115 Csac_2248	Calkr_0643 Calkr_0644 Calkr_1085 Calkr_1722
Hydrogenase	Csac_1864	Calkr_1284
Lactate dehydrogenase	N/A	Calkr_1983
Alcohol dehydrogenase	Csac_0395 Csac_0407 Csac_0622 Csac_0711 Csac_0763 Csac_0869 Csac_1226 Csac_1500	Calkr_0097 Calkr_0372 Calkr_0544 Calkr_0933 Calkr_2248

The locus tags listed here were obtained from the NCBI protein database accessible online.
In particular, the enzymes of interest include cellulases, β-glucosidases, and xylanases which help to break down cellulosic materials; transporters such as ABC and PTS which bring carbohydrates into the cells; phosphofructokinase, KDPG aldolase and pyruvate kinase from the EMP and ED pathways of glycolysis; transketolase and transaldolase from the PPP pathway; and enzymes from other fermentative pathways.

3.5 Proteomic analysis

Proteomic analyses provide a snapshot of the proteins expressed at a given point in the culture's growth and it also shows the comparative levels of expression. More specifically, when studying the effects of different growth substrates, the specific amounts of protein translated could be compared to see which enzymes may be upregulated or downregulated in each case. Five pairwise comparisons were done to examine the difference in protein expression levels when each organism was grown on: a C6-sugar versus a C5-sugar (glucose vs xylose), a C6-monosaccharide versus a C6-disaccharide (glucose vs cellobiose), a C5-monosaccharide versus a C5-polymer (xylose vs xylan), a C6-monosaccharide versus a C6-polymer (glucose vs Avicel), and two C6-polymers (Avicel vs CMC).

Peptides were quantified through 2D LC-MS/MS and the proteomics experiments were normalized using the relative abundance index, which can be calculated by dividing the number of spectral counts by the molecular mass of the protein. Expression levels were obtained from proteomic experiments, showing the $\log_2$ signal intensity from an average of duplicates. The following tables highlight some of the highly upregulated enzymes of interest which are involved in carbohydrate transport and metabolism. Although Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest, other proteins with a z-score of ≥1.2 are included for consideration. Expression ratios were determined by finding the difference between the $\log_2$ values of each protein and the average of these expression ratios was used to find the Z-score by calculating the deviation from the mean when comparing two samples. It should also be noted that false discovery rates do not apply here because each sample was run in duplicates.

For *C. saccharolyticus* and *C. kristjanssonii* respectively, the upregulated proteins are shown between glucose and xylose (Tables 6 and 11), cellobiose and glucose (Tables 7 and 12), xylan and xylose (Tables 8 and 13), Avicel and glucose (Tables 9 and 14), and CMC and Avicel (Tables 10 and 15). The proteins that were most seen to be upregulated include flagellar proteins, GHs, ABC transporters, and central metabolic pathway enzymes including xylose isomerase from the PPP, phosphofructokinase from the EMP, acetate kinase, hydrogenase, and alcohol dehydrogenase from fermentative pathways. While GHs are considered for their ability to degrade cellulose, ABC transporters for sugar uptake, and central metabolic pathway enzymes for sugar metabolism, flagella were also included in this analysis because of their presence in previous studies and their possible function of substrate adhesion.

Notably, ABC transporters were upregulated in glucose and xylan (Csac_2504) and in cellobiose, Avicel, CMC, and xylan (Csac_0681). For *C. kristjanssonii*, xylose isomerase (Calkr_1997) was upregulated in different conditions including glucose, xylose, and CMC. Please see Table 22 for enzymes that were upregulated in both the proteomic and RNA-seq datasets.

Table 6. Upregulated proteins of interest from *C. saccharolyticus* grown on xylose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Xylose	Glucose	
Upregulated with Xylose			
Csac_0586 glycoside hydrolase, family 3 domain protein	21.8	17.9	3.49
Csac_0763 Alcohol dehydrogenase GroES domain protein	18.3	15.7	2.35
Csac_1154 Xylose isomerase domain protein TIM barrel	22.9	20.9	1.90
Csac_1541 hydrogenase accessory protein HypB	18.9	17.0	1.75
Csac_0204 glycoside hydrolase, family 10	17.8	16.0	1.70
Csac_2249 pyruvate/ketoisovalerate oxidoreductase, gamma	22.0	20.3	1.57
Csac_1257 flagellar M–ring protein FliF	17.0	15.4	1.47
Csac_2248 pyruvate flavodoxin/ferredoxin oxidoreductase	22.6	21.2	1.35
Csac_1551 pyruvate ferredoxin/flavodoxin oxidoreductase	20.0	18.6	1.32
Upregulated with Glucose			
Csac_2441 PfkB domain protein	13.7	20.6	5.72
Csac_0203 alpha amylase, catalytic sub domain	15.9	21.2	4.34
Csac_1173 Xylose isomerase domain protein TIM barrel	16.7	20.6	3.19
Csac_0426 alpha amylase, catalytic region	16.8	20.0	2.54
Csac_2711 6–phosphogluconate dehydrogenase, NAD–binding	16.1	18.8	2.14
Csac_0359 glycoside hydrolase, family 43	14.5	17.1	2.02
Csac_1090 glycosyltransferase 36	16.3	18.4	1.70
Csac_1116 ABC transporter related	15.8	17.8	1.57
Csac_1118 glycoside hydrolase, clan GH–D	18.1	19.9	1.45
Csac_1102 glycoside hydrolase, family 3 domain protein	18.2	20.0	1.33
Csac_1226 Alcohol dehydrogenase GroES domain protein	17.1	18.7	1.22

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on xylose as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 7. Upregulated proteins of interest from *C. saccharolyticus* grown on cellobiose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Cellobiose	Glucose	
Upregulated with Cellobiose			
Csac_0681 ABC–type sugar transport system periplasmic	27.9	17.1	11.3
Csac_1091 glycosyltransferase 36	24.0	16.8	7.59
Csac_1090 glycosyltransferase 36	24.5	18.4	6.36
Csac_1089 Beta–glucosidase	19.0	16.8	2.44
Csac_1257 flagellar M–ring protein FliF	17.1	15.4	1.87
Upregulated with Glucose			
Csac_2441 PfkB domain protein	13.9	20.6	6.61
Csac_1102 glycoside hydrolase, family 3 domain protein	16.1	20.0	3.74
Csac_2504 Monosaccharide–transporting ATPase	16.3	19.5	3.14
Csac_0395 Alcohol dehydrogenase, zinc–binding domain	17.7	20.6	2.85
Csac_0203 alpha amylase, catalytic sub domain	18.5	21.2	2.52
Csac_0396 Xylose isomerase domain protein TIM barrel	16.9	19.3	2.31
Csac_2734 glycoside hydrolase family 2, TIM barrel	14.5	16.6	1.92
Csac_1698 flagellar protein FlaG protein	15.1	16.7	1.43
Csac_0360 glycosyl hydrolase, family 88	16.1	17.7	1.42

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on cellobiose as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 8. Upregulated proteins of interest from *C. saccharolyticus* grown on xylan versus xylose

Locus tag and enzyme name	Expression value		Z–score
	Xylan	Xylose	
Upregulated with Xylan			
Csac_0681 ABC–type sugar transport system periplasmic	20.4	15.9	3.63
Csac_1090 glycosyltransferase 36	20.2	16.3	3.18
Csac_0203 alpha amylase, catalytic sub domain	19.8	15.9	3.13
Csac_0204 glycoside hydrolase, family 10	21.4	17.8	2.93
Csac_1173 Xylose isomerase domain protein TIM barrel	20.1	16.7	2.79
Csac_0586 glycoside hydrolase, family 3 domain protein	24.9	21.8	2.51
Csac_0798 xylulokinase	23.4	20.7	2.19
Csac_1154 Xylose isomerase domain protein TIM barrel	25.5	22.9	2.12
Csac_1459 pyruvate ferredoxin/flavodoxin oxidoreductase,	23.4	20.8	2.11
Csac_0400 carbon–monoxide dehydrogenase, catalytic subunit	21.7	19.3	2.01
Csac_2686 glycoside hydrolase family 2, sugar binding	20.4	18.0	1.95
Csac_1118 glycoside hydrolase, clan GH–D	20.1	18.1	1.71
Csac_0362 glycoside hydrolase family 2, TIM barrel	20.1	18.0	1.71
Csac_0359 glycoside hydrolase, family 43	16.5	14.5	1.64
Csac_0426 alpha amylase, catalytic region	18.8	16.8	1.61
Csac_1226 Alcohol dehydrogenase GroES domain protein	18.7	17.1	1.37
Csac_1189 fructose–1,6–bisphosphate aldolase, class II	24.6	23.0	1.35
Csac_2504 Monosaccharide–transporting ATPase	19.9	18.5	1.22
Upregulated with Xylose			
Csac_1258 flagellar motor switch protein FliG	13.9	17.4	2.58
Csac_1541 hydrogenase accessory protein HypB	15.9	18.9	2.25
Csac_2734 glycoside hydrolase family 2, TIM barrel	14.4	16.3	1.32
Csac_1268 flagellar basal body–associated protein FliL	17.3	19.1	1.29

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on xylan as opposed to xylose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 9. Upregulated proteins of interest from *C. saccharolyticus* grown on Avicel versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Avicel	Glucose	
Upregulated with Avicel			
Csac_0681 ABC–type sugar transport system periplasmic	24.7	17.1	4.55
Csac_0204 glycoside hydrolase, family 10	21.3	16.0	3.38
Csac_1091 glycosyltransferase 36	21.6	16.8	3.06
Csac_0296 glycosidase, PH1107–related	22.0	17.6	2.84
Csac_1090 glycosyltransferase 36	22.3	18.4	2.53
Csac_0586 glycoside hydrolase, family 3 domain protein	21.6	17.9	2.45
Csac_0129 glycoside hydrolase family 2, sugar binding	18.6	15.1	2.33
Csac_1163 Phosphotransferase system, phosphocarrier	19.2	16.3	2.06
Csac_1257 flagellar M–ring protein FliF	17.9	15.4	1.80
Csac_0798 xylulokinase	22.6	20.2	1.74
Csac_2734 glycoside hydrolase family 2, TIM barrel	18.6	16.6	1.57
Csac_1154 Xylose isomerase domain protein TIM barrel	22.7	20.9	1.44
Upregulated with Glucose			
Csac_2439 PTS system, fructose subfamily, IIC subunit	16.7	22.0	2.43
Csac_2437 phosphoenolpyruvate–protein phosphotransferase	15.7	20.9	2.36
Csac_1460 pyruvate flavodoxin/ferredoxin oxidoreductase	18.4	23.4	2.28
Csac_2040 acetate kinase	19.5	24.4	2.19
Csac_1458 pyruvate/ketoisovalerate oxidoreductase, gamma	19.6	24.3	2.09
Csac_0622 iron–containing alcohol dehydrogenase	22.7	26.0	1.35

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on Avicel as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 10. Upregulated proteins of interest from *C. saccharolyticus* grown on CMC versus Avicel

Locus tag and enzyme name	Expression value		Z–score
	CMC	Avicel	
Upregulated with CMC			
Csac_1173 Xylose isomerase domain protein TIM barrel	25.0	19.8	2.75
Csac_1460 pyruvate flavodoxin/ferredoxin oxidoreductase	22.9	18.4	2.36
Csac_1458 pyruvate/ketoisovalerate oxidoreductase, gamma	23.7	19.6	2.10
Csac_0681 ABC–type sugar transport system periplasmic	28.3	24.7	1.83
Csac_2040 acetate kinase	22.6	19.5	1.53
Csac_2748 glycoside hydrolase, family 4	21.9	19.1	1.38
Csac_1090 glycosyltransferase 36	24.9	22.3	1.26
Csac_1091 glycosyltransferase 36	24.2	21.6	1.26
Csac_0396 Xylose isomerase domain protein TIM barrel	21.5	18.8	1.25
Upregulated with Avicel			
Csac_2506 D–xylose ABC transporter, periplasmic	20.3	24.5	2.79
Csac_0204 glycoside hydrolase, family 10	17.4	21.3	2.63
Csac_2734 glycoside hydrolase family 2, TIM barrel	15.1	18.6	2.38
Csac_1268 flagellar basal body–associated protein FliL	16.3	18.8	1.76
Csac_0129 glycoside hydrolase family 2, sugar binding	16.4	18.6	1.59
Csac_1257 flagellar M–ring protein FliF	15.8	17.9	1.55
Csac_0426 alpha amylase, catalytic region	15.8	17.7	1.41
Csac_1551 pyruvate ferredoxin/flavodoxin oxidoreductase	16.4	18.3	1.38
Csac_0598 Xylose isomerase domain protein TIM barrel	16.7	18.5	1.35
Csac_0400 carbon–monoxide dehydrogenase, catalytic subunit	17.2	18.8	1.30
Csac_0586 glycoside hydrolase, family 3 domain protein	19.9	21.6	1.29

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on CMC as opposed to Avicel. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 11. Upregulated proteins of interest from *C. kristjanssonii* grown on xylose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Xylose	Glucose	
Upregulated with Xylose			
Calkr_1997 xylose isomerase	25.9	20.4	4.96
Calkr_0175 carbon–monoxide dehydrogenase, catalytic subunit	22.7	17.9	4.40
Calkr_2026 xylulokinase	21.6	17.2	4.07
Calkr_2396 glycoside hydrolase family 3 domain protein	19.4	15.2	3.89
Calkr_0097 iron–containing alcohol dehydrogenase	27.2	23.2	3.68
Calkr_0505 transketolase central region	22.7	21.1	1.63
Calkr_1722 pyruvate flavodoxin/ferredoxin oxidoreductase	18.4	17.2	1.34
Calkr_0504 transketolase domain–containing protein	21.6	20.6	1.20
Upregulated with Glucose			
Calkr_1677 flagellin domain protein	21.8	25.1	2.57
Calkr_2248 iron–containing alcohol dehydrogenase	19.8	23.1	2.56
Calkr_0311 xylose isomerase domain–containing protein tim	17.5	20.0	1.84
Calkr_1980 pfkb domain protein	13.0	15.5	1.78
Calkr_2435 d–xylose abc transporter, periplasmic	25.6	27.9	1.66
Calkr_0218 glycoside hydrolase family 2 immunoglobulin	16.9	18.9	1.45
Calkr_2207 beta–galactosidase	16.6	18.4	1.29
Calkr_2205 glycosyltransferase 36	20.5	22.3	1.25

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on xylose as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 12. Upregulated proteins of interest from *C. kristjanssonii* grown on cellobiose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Cellobiose	Glucose	
Upregulated with Cellobiose			
Calkr_2026 xylulokinase	20.0	17.2	2.85
Calkr_2206 glycosyltransferase 36	23.9	21.3	2.65
Calkr_0097 iron–containing alcohol dehydrogenase	25.7	23.2	2.56
Calkr_1722 pyruvate flavodoxin/ferredoxin oxidoreductase	19.2	17.2	2.16
Calkr_2205 glycosyltransferase 36	24.3	22.3	2.14
Calkr_0175 carbon–monoxide dehydrogenase, catalytic subunit	19.9	17.9	2.09
Calkr_2239 glycosyltransferase 28–like protein	19.1	17.3	1.94
Calkr_1713 xylose isomerase domain–containing protein tim	18.2	16.7	1.69
Calkr_2471 glycosyl transferase group 1	15.7	14.5	1.38
Upregulated with Glucose			
Calkr_1997 xylose isomerase	16.3	20.4	3.84
Calkr_1666 flagellar hook–associated 2 domain–containing	14.8	18.8	3.73
Calkr_2435 d–xylose abc transporter, periplasmic	23.9	27.9	3.71
Calkr_0415 flagellar hook–basal body protein	15.6	18.7	2.87
Calkr_1677 flagellin domain protein	22.8	25.1	2.06
Calkr_2248 iron–containing alcohol dehydrogenase	20.9	23.1	2.06
Calkr_2204 glycosyl transferase group 1	14.8	16.6	1.66
Calkr_0331 glycosyl transferase family 2	17.7	19.5	1.55

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on cellobiose as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 13. Upregulated proteins of interest from *C. kristjanssonii* grown on xylan versus xylose

Locus tag and enzyme name	Expression value		Z–score
	Xylan	Xylose	
Upregulated with Xylan			
Calkr_2396 glycoside hydrolase family 3 domain protein	24.1	19.4	3.44
Calkr_1980 pfkb domain protein	16.8	13.0	2.76
Calkr_2205 glycosyltransferase 36	23.6	20.5	2.30
Calkr_0311 xylose isomerase domain–containing protein tim	20.5	17.5	2.21
Calkr_2206 glycosyltransferase 36	22.8	20.5	1.71
Calkr_2239 glycosyltransferase 28–like protein	19.8	17.7	1.57
Calkr_1677 flagellin domain protein	23.8	21.8	1.54
Upregulated with Xylose			
Calkr_0175 carbon–monoxide dehydrogenase, catalytic subunit	16.0	22.7	5.03
Calkr_0097 iron–containing alcohol dehydrogenase	23.0	27.2	3.14
Calkr_0782 l–lactate dehydrogenase	16.7	18.6	1.39
Calkr_1635 glycosyl transferase group 1	19.2	21.1	1.38

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on xylan as opposed to xylose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 14. Upregulated proteins of interest from *C. kristjanssonii* grown on Avicel versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Avicel	Glucose	
Upregulated with Avicel			
Calkr_2396 glycoside hydrolase family 3 domain protein	22.5	15.2	5.45
Calkr_2245 cellulose 1,4–beta–cellobiosidase	21.2	15.5	4.38
Calkr_2026 xylulokinase	21.0	17.2	2.95
Calkr_1997 xylose isomerase	23.2	20.4	2.27
Calkr_0097 iron–containing alcohol dehydrogenase	25.2	23.2	1.68
Calkr_0462 pfkb domain protein	18.5	16.6	1.62
Calkr_0218 glycoside hydrolase family 2 immunoglobulin	20.8	18.9	1.57
Calkr_2206 glycosyltransferase 36	23.0	21.3	1.47
Calkr_0283 phosphotransferase system, phosphocarrier	18.5	16.9	1.41
Upregulated with Glucose			
Calkr_2435 d–xylose abc transporter, periplasmic	25.0	27.9	1.83
Calkr_1693 flagellar operon protein yvyf	16.2	18.7	1.52
Calkr_1666 flagellar hook–associated 2 domain–containing	16.5	18.8	1.42
Calkr_2248 iron–containing alcohol dehydrogenase	20.8	23.1	1.38
Calkr_1478 acetate kinase	21.5	23.8	1.36
Calkr_0415 flagellar hook–basal body protein	16.5	18.7	1.28

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on Avicel as opposed to glucose. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

Table 15. Upregulated proteins of interest from *C. kristjanssonii* grown on CMC versus Avicel

Locus tag and enzyme name	Expression value		Z–score
	CMC	Avicel	
Upregulated with CMC			
Calkr_0175 carbon–monoxide dehydrogenase, catalytic subunit	20.2	16.8	3.11
Calkr_0311 xylose isomerase domain–containing protein tim	23.8	20.6	2.92
Calkr_1722 pyruvate flavodoxin/ferredoxin oxidoreductase	18.9	16.7	2.00
Calkr_1997 xylose isomerase	25.2	23.2	1.84
Calkr_0143 alpha amylase catalytic region	22.1	20.3	1.73
Calkr_2205 glycosyltransferase 36	25.1	23.4	1.65
Calkr_2207 beta–galactosidase	21.1	19.4	1.57
Calkr_2206 glycosyltransferase 36	24.7	23.0	1.53
Calkr_1514 glycosidase related protein	18.7	17.3	1.28
Upregulated with Avicel			
Calkr_2396 glycoside hydrolase family 3 domain protein	19.1	22.5	3.00
Calkr_0017 type 3a cellulose–binding domain protein	15.6	18.3	2.30
Calkr_0097 iron–containing alcohol dehydrogenase	22.9	25.2	2.02
Calkr_0283 phosphotransferase system, phosphocarrier	16.8	18.5	1.40
Calkr_0218 glycoside hydrolase family 2 immunoglobulin	19.3	20.8	1.23

This table lists the protein expression values beside their corresponding locus tags and enzyme names as well as their z-scores from comparing the expression of the proteins when grown on CMC as opposed to Avicel. Expression values represent the $\log_2$ signal intensity reads. Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but others with a z-score of 1.2 or more are included here so that other relevant proteins can be examined.

3.6 RNA sequencing

RNA sequencing examines the transcriptome, or all of the mRNA sequences that are transcribed in each organism. More specifically, when examining eight different growth substrates, the amounts of mRNA transcribed could be compared to see which enzymes may be upregulated or downregulated. Three pairwise comparisons were done to examine the difference in expression levels of mRNA when each organism was grown on: a C6-sugar versus a C5-sugar (glucose vs xylose), a C6-monosaccharide versus a C6-disaccharide (glucose vs cellobiose), and a C5-monosaccharide versus a C5-polymer (xylose vs xylan).

The following tables highlight some of the highly upregulated enzymes of interest which are involved in carbohydrate transport and metabolism. Upregulated RNAs were compared between xylose and glucose, cellobiose and glucose, and xylan and xylose for *C. saccharolyticus* in Tables 16, 17, and18 and for *C. kristjanssonii* in Tables19, 20, and 21. Expression levels were obtained from RNA-seq analyses, showing the $\log_2$ signal intensity in the Illumina expression array. The expression values shown represent an average of duplicates. Although Z-scores of 1.96 or greater are used as a threshold to identify sequences of interest, other enzymes with a z-score of ≥1.2 are included for consideration. Expression ratios were determined by finding the difference between the $\log_2$ values of each enzyme and the average of these expression ratios was used to find the Z-score by calculating the deviation from the mean when comparing two sample conditions.

Several flagellar proteins, glycoside hydrolases, ABC transporters, as well as glycolytic and fermentative pathway enzymes were observed to be upregulated. Specifically, family 3 and family 5 GHs were observed to be upregulated in several conditions in *C. saccharolyticus*

while family 2 and family 10 GHs were upregulated in more than one growth condition for *C. kristjanssonii*. An ABC transporter (Csac_0297) was upregulated in both cellobiose and xylan for *C. saccharolyticus*. Meanwhile, for *C. kristjanssonii*, xylose isomerase (Calkr_1173) was upregulated in both glucose and cellobiose. More specific loci and enzyme regulation response will be discussed in the following chapter. Please see Table 22 for enzymes that were upregulated in both the proteomic and RNA-seq datasets.

Table 16. Upregulated RNA sequences for enzymes of interest from *C. saccharolyticus* grown on xylose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Xylose	Glucose	
Upregulated with Xylose			
Csac_0869 Alcohol dehydrogenase, zinc–binding domain	7.97	3.36	3.17
Csac_0241 Monosaccharide–transporting ATPase	11.1	7.04	2.82
Csac_2366 phosphofructokinase	12.8	10.2	1.82
Csac_2720 PfkB domain protein	8.25	5.69	1.76
Csac_0586 glycoside hydrolase, family 3 domain protein	10.9	8.43	1.71
Csac_0762 glycosidase, PH1107–related	8.77	6.43	1.61
Csac_1541 hydrogenase accessory protein HypB	11.6	9.57	1.37
Csac_0696 Endo–1,4–beta–xylanase	12.0	10.2	1.23
Upregulated with Glucose			
Csac_2439 PTS system, fructose subfamily, IIC subunit	8.75	15.0	4.30
Csac_2438 Phosphotransferase system, phosphocarrier	3.71	9.05	3.70
Csac_2441 PfkB domain protein	7.73	12.2	3.08
Csac_2437 phosphoenolpyruvate–protein phosphotransferase	9.36	13.2	2.65
Csac_2506 D–xylose ABC transporter, periplasmic	13.5	16.5	2.10
Csac_1700 flagellar protein FliS	7.94	10.8	1.99
Csac_1698 flagellar protein FlaG protein	9.06	11.9	1.98
Csac_0408 alpha amylase, catalytic region	7.99	10.8	1.94
Csac_1699 flagellar hook–associated 2 domain protein	10.8	13.5	1.84
Csac_2504 Monosaccharide–transporting ATPase	11.1	13.6	1.78
Csac_0134 glycosyl transferase, group 1	10.2	12.7	1.75
Csac_0395 Alcohol dehydrogenase, zinc–binding domain	9.59	11.9	1.64
Csac_2249 pyruvate/ketoisovalerate oxidoreductase, gamma	12.3	14.6	1.58
Csac_0350 Xylose isomerase domain protein TIM barrel	8.73	10.7	1.38
Csac_0259 glycosidase, PH1107–related	8.75	10.6	1.26
Csac_1264 flagellar hook capping protein	8.19	9.98	1.24
Csac_1173 Xylose isomerase domain protein TIM barrel	9.29	11.0	1.21

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on xylose or glucose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

Table 17. Upregulated RNA sequences for enzymes of interest from *C. saccharolyticus* grown on cellobiose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Cellobiose	Glucose	
Upregulated with Cellobiose			
Csac_0681 ABC–type sugar transport system periplasmic	17.0	9.05	7.81
Csac_0678 Cellulase	12.9	7.46	5.32
Csac_1864 hydrogenase, Fe–only	15.2	13.1	2.14
Csac_1541 hydrogenase accessory protein HypB	11.7	9.57	2.09
Csac_1459 pyruvate ferredoxin/flavodoxin oxidoreductase,	11.6	9.67	1.94
Csac_2040 acetate kinase	14.3	12.4	1.90
Csac_1089 Beta–glucosidase	9.58	7.65	1.90
Csac_0297 ABC–type sugar transport system periplasmic	10.0	8.16	1.84
Csac_1953 glyceraldehyde–3–phosphate dehydrogenase, type I	16.2	14.3	1.83
Csac_1173 Xylose isomerase domain protein TIM barrel	12.8	11.0	1.79
Csac_1076 glycoside hydrolase, family 48	9.86	8.07	1.77
Csac_1460 pyruvate flavodoxin/ferredoxin oxidoreductase	14.5	12.8	1.68
Csac_1502 FAD dependent oxidoreductase	9.76	8.09	1.65
Csac_1172 oxidoreductase domain protein	11.4	9.86	1.51
Csac_0258 glycosidase, PH1107–related	7.50	5.98	1.50
Csac_2539 glycoside hydrolase, family 20	9.82	8.46	1.34
Csac_1078 Cellulose 1,4–beta–cellobiosidase	9.78	8.46	1.31
Csac_1077 glycoside hydrolase, family 5	9.70	8.46	1.23
Csac_1027 L–lactate dehydrogenase	14.9	13.7	1.21
Csac_1079 Cellulase., Cellulose 1,4–beta–cellobiosidase	10.1	8.93	1.20
Upregulated with Glucose			
Csac_2439 PTS system, fructose subfamily, IIC subunit	8.93	15.0	4.84
Csac_2441 PfkB domain protein	8.93	12.2	4.14
Csac_2438 Phosphotransferase system, phosphocarrier	5.23	9.05	3.75
Csac_2696 ABC–type sugar transport system periplasmic	5.84	9.42	3.51
Csac_2437 phosphoenolpyruvate–protein phosphotransferase	9.72	13.2	3.40
Csac_0390 Xylose isomerase domain protein TIM barrel	6.08	9.27	3.12
Csac_2504 Monosaccharide–transporting ATPase	10.6	13.6	3.02
Csac_0393 Monosaccharide–transporting ATPase	6.65	9.68	2.97
Csac_2505 ABC transporter related	9.58	12.5	2.86
Csac_0392 ABC transporter related	6.61	9.30	2.63
Csac_2506 D–xylose ABC transporter, periplasmic	13.9	16.5	2.57
Csac_1102 glycoside hydrolase, family 3 domain protein	8.11	10.6	2.49
Csac_0394 Monosaccharide–transporting ATPase	4.46	6.99	2.48
Csac_0137 glycoside hydrolase, family 5	7.44	9.89	2.40
Csac_0400 carbon–monoxide dehydrogenase, catalytic subunit	9.23	11.4	2.15
Csac_2436 Acetyl xylan esterase	7.66	9.78	2.07
Csac_0696 Endo–1,4–beta–xylanase	8.13	10.2	1.99
Csac_1354 glycoside hydrolase, family 31	7.43	9.07	1.61
Csac_2404 glycoside hydrolase, family 39	7.35	8.78	1.39
Csac_2408 Endo–1,4–beta–xylanase	7.01	8.40	1.36
Csac_0721 L–ribulokinase	6.21	7.59	1.34
Csac_0203 alpha amylase, catalytic sub domain	10.0	11.3	1.27
Csac_2202 ABC transporter related	8.50	9.77	1.24
Csac_2686 glycoside hydrolase family 2, sugar binding	8.36	9.62	1.23
Csac_1018 Beta–galactosidase	11.1	12.4	1.22
Csac_0408 alpha amylase, catalytic region	9.54	10.8	1.21

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on cellobiose or glucose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

Table 18. Upregulated RNA sequences for enzymes of interest from *C. saccharolyticus* grown on xylan versus xylose

Locus tag and enzyme name	Expression value		Z–score
	Xylan	Xylose	
Upregulated with Xylan			
Csac_2696 ABC–type sugar transport system periplasmic	15.5	8.17	4.50
Csac_2435 ABC transporter related	12.2	8.99	1.97
Csac_0205 polysaccharide deacetylase	11.0	7.87	1.94
Csac_1172 oxidoreductase domain protein	10.6	7.63	1.85
Csac_0420 oxidoreductase domain protein	12.4	9.48	1.79
Csac_2249 pyruvate/ketoisovalerate oxidoreductase, gamma	14.9	12.3	1.63
Csac_1953 glyceraldehyde–3–phosphate dehydrogenase, type I	15.3	12.8	1.57
Csac_1699 flagellar hook–associated 2 domain protein	13.3	10.8	1.57
Csac_1700 flagellar protein FliS	10.4	7.94	1.56
Csac_1698 flagellar protein FlaG protein	11.5	9.06	1.55
Csac_0297 ABC–type sugar transport system periplasmic	10.4	7.91	1.54
Csac_0429 alpha–glucan phosphorylase	13.2	10.8	1.52
Csac_0259 glycosidase, PH1107–related	11.2	8.75	1.50
Csac_0782 glucose–1–phosphate adenylyltransferase, GlgD	13.5	11.1	1.49
Csac_0783 glucose–1–phosphate adenylyltransferase	13.5	11.1	1.48
Csac_1154 Xylose isomerase domain protein TIM barrel	15.3	12.9	1.48
Csac_1027 L–lactate dehydrogenase	14.8	12.6	1.40
Csac_1955 pyruvate, phosphate dikinase	15.0	12.9	1.32
Csac_0426 alpha amylase, catalytic region	12.4	10.3	1.32
Csac_0586 glycoside hydrolase, family 3 domain protein	13.0	10.9	1.29
Csac_2036 putative transaldolase	12.1	10.1	1.24
Csac_0798 xylulokinase	12.5	10.5	1.23
Csac_0204 glycoside hydrolase, family 10	12.5	10.5	1.22
Upregulated with Xylose			
Csac_2366 phosphofructokinase	7.56	12.8	3.23
Csac_0241 Monosaccharide–transporting ATPase	6.77	11.1	2.67
Csac_0762 glycosidase, PH1107–related	5.55	8.77	1.96
Csac_1096 aldo/keto reductase	4.38	7.49	1.89
Csac_1141 short–chain dehydrogenase/reductase SDR	4.89	7.92	1.84
Csac_0869 Alcohol dehydrogenase, zinc–binding domain	4.96	7.97	1.83
Csac_2720 PfkB domain protein	5.35	8.25	1.76
Csac_1077 glycoside hydrolase, family 5	6.80	9.43	1.60
Csac_0400 carbon–monoxide dehydrogenase, catalytic subunit	9.63	12.2	1.57
Csac_0258 glycosidase, PH1107–related	5.33	7.64	1.41
Csac_2527 glycosidase, PH1107–related	5.83	8.13	1.40
Csac_1080 glycoside hydrolase, family 5	5.85	8.00	1.31
Csac_1079 Cellulase., Cellulose 1,4–beta–cellobiosidase	7.45	9.54	1.27

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on xylan or xylose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

Table 19. Upregulated RNA sequences for enzymes of interest from *C. kristjanssonii* grown on xylose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Xylose	Glucose	
Upregulated with Xylose			
Calkr_1447 polysaccharide deacetylase	13.4	7.66	3.06
Calkr_0052 abc–2 type transporter	12.0	8.04	2.12
Calkr_0051 abc–2 type transporter	12.4	8.97	1.82
Calkr_1997 xylose isomerase	14.1	10.9	1.66
Calkr_0736 xylose isomerase domain–containing protein tim	12.4	9.26	1.66
Calkr_0844 polysaccharide deacetylase	13.5	10.4	1.62
Calkr_0097 iron–containing alcohol dehydrogenase	15.7	12.8	1.53
Calkr_2204 glycosyl transferase group 1	12.7	9.90	1.49
Calkr_0532 glycoside hydrolase family 4	11.3	8.61	1.45
Calkr_1588 glycosyl transferase group 1	10.3	7.85	1.28
Calkr_0498 glycosyl transferase group 1	12.5	10.1	1.27
Upregulated with Glucose			
Calkr_1677 flagellin domain protein	10.4	15.0	2.41
Calkr_1722 pyruvate flavodoxin/ferredoxin oxidoreductase	9.74	13.5	1.97
Calkr_1723 pyruvate/ketoisovalerate oxidoreductase, gamma	9.77	13.4	1.93
Calkr_1691 flgn family protein	6.31	9.94	1.92
Calkr_1690 flagellar hook–associated protein flgk	8.37	11.8	1.84
Calkr_0117 alpha amylase catalytic region	7.47	10.9	1.81
Calkr_1693 flagellar operon protein yvyf	6.85	10.2	1.78
Calkr_1666 flagellar hook–associated 2 domain–containing	10.8	13.6	1.46
Calkr_1689 flagellar hook–associated protein 3	8.10	10.7	1.40
Calkr_0228 glycoside hydrolase 15–related protein	6.52	9.14	1.39
Calkr_1676 glycosyl transferase family 2	7.40	10.0	1.39
Calkr_2248 iron–containing alcohol dehydrogenase	11.1	13.4	1.26
Calkr_1514 glycosidase related protein	8.82	11.2	1.24
Calkr_0337 phosphoenolpyruvate carboxykinase (gtp)	11.9	14.3	1.23

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on xylose or glucose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

Table 20. Upregulated RNA sequences for enzymes of interest from *C. kristjanssonii* grown on cellobiose versus glucose

Locus tag and enzyme name	Expression value		Z–score
	Cellobiose	Glucose	
Upregulated with Cellobiose			
Calkr_0229 glycosyl transferase group 1	14.1	9.41	4.62
Calkr_1929 glycoside hydrolase family 10	11.9	8.80	3.08
Calkr_0120 alpha–glucan phosphorylase	13.2	10.4	2.77
Calkr_0849 phosphofructokinase	15.2	12.5	2.70
Calkr_2206 glycosyltransferase 36	14.8	12.1	2.68
Calkr_2205 glycosyltransferase 36	14.5	12.1	2.42
Calkr_0218 glycoside hydrolase family 2 immunoglobulin	12.5	10.1	2.32
Calkr_1664 flagellar protein flis	12.6	10.5	2.09
Calkr_0311 xylose isomerase domain–containing protein tim	11.1	9.11	2.04
Calkr_0130 polysaccharide deacetylase	11.6	9.80	1.75
Calkr_0782 l–lactate dehydrogenase	13.7	12.1	1.58
Calkr_1928 polysaccharide deacetylase	9.59	8.14	1.46
Calkr_1677 flagellin domain protein	16.3	15.0	1.32
Upregulated with Glucose			
Calkr_0175 carbon–monoxide dehydrogenase, catalytic subunit	8.39	11.6	3.21
Calkr_0543 glycosidase related protein	7.02	9.15	2.09
Calkr_1977 glycosyl hydrolase family 88	7.22	9.02	1.77
Calkr_1976 glycoside hydrolase family 28	7.16	8.68	1.49
Calkr_1981 2–dehydro–3–deoxyphosphogluconate	6.81	8.32	1.48
Calkr_0572 endo–1,4–beta–xylanase	7.61	9.09	1.46
Calkr_1650 aldose 1–epimerase	8.81	10.2	1.36
Calkr_0995 pyruvate/ketoisovalerate oxidoreductase	7.06	8.45	1.36
Calkr_1855 glycoside hydrolase family 2 sugar binding	7.95	9.28	1.31
Calkr_1980 pfkb domain protein	7.55	8.82	1.24

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on cellobiose or glucose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

Table 21. Upregulated RNA sequences for enzymes of interest from *C. kristjanssonii* grown on xylan versus xylose

Locus tag and enzyme name	Expression Value		Z–score
	Xylan	Xylose	
Upregulated with Xylan			
Calkr_1677 flagellin domain protein	16.9	10.4	2.73
Calkr_2396 glycoside hydrolase family 3 domain protein	14.4	8.26	2.58
Calkr_2245 cellulose 1,4–beta–cellobiosidase	16.3	10.3	2.52
Calkr_0572 endo–1,4–beta–xylanase	14.3	8.39	2.50
Calkr_1858 alpha–glucuronidase	12.8	7.10	2.41
Calkr_0229 glycosyl transferase group 1	12.8	7.52	2.22
Calkr_1723 pyruvate/ketoisovalerate oxidoreductase, gamma	14.9	9.77	2.19
Calkr_1722 pyruvate flavodoxin/ferredoxin oxidoreductase	14.6	9.74	2.07
Calkr_0117 alpha amylase catalytic region	12.3	7.47	2.03
Calkr_1693 flagellar operon protein yvyf	11.1	6.85	1.81
Calkr_1855 glycoside hydrolase family 2 sugar binding	12.5	8.38	1.75
Calkr_2026 xylulokinase	13.9	9.87	1.70
Calkr_1666 flagellar hook–associated 2 domain–containing	14.7	10.8	1.63
Calkr_1690 flagellar hook–associated protein flgk	12.1	8.37	1.58
Calkr_1514 glycosidase related protein	12.5	8.82	1.55
Calkr_1676 glycosyl transferase family 2	10.9	7.40	1.50
Calkr_1929 glycoside hydrolase family 10	12.2	8.68	1.49
Calkr_1928 polysaccharide deacetylase	10.9	7.52	1.42
Calkr_1689 flagellar hook–associated protein 3	11.2	8.10	1.32
Upregulated with Xylose			
Calkr_1447 polysaccharide deacetylase	5.95	13.4	3.12
Calkr_1588 glycosyl transferase group 1	4.72	10.3	2.31
Calkr_0844 polysaccharide deacetylase	10.1	13.5	1.39
Calkr_0543 glycosidase related protein	6.86	10.1	1.33
Calkr_0532 glycoside hydrolase family 4	8.13	11.3	1.33

Upregulated gene locus tags and annotated enzyme names are listed here along with their corresponding expression levels when grown on xylan or xylose and their comparative z-scores. Expression values represent the $\log_2$ signal intensity reads while Z-scores of 1.96 or greater are used as a threshold to identify proteins of interest but z-scores of 1.2 or more are also included here to be considered.

4.0 Discussion

Discoveries regarding the regulation and effects of growth on different substrates are discussed here. The results were analyzed with respect to their importance in carbohydrate catabolism, uptake, and utilization with a focus on understanding the metabolism of *C. saccharolyticus* and *C. kristjanssonii*. The influence of different substrates on their growth, physiology, and enzyme regulation is also examined.

4.1 Growth characteristics of *Caldicellulosiruptor* spp.

In order to examine the differences in growth characteristics resulting from varying substrates, *C. saccharolyticus* and *C. kristjanssonii* were monitored through cell counts and metabolite production. The expression of different cellulase enzymes as well as transport proteins also contribute to their ability to grow in each condition.

4.1.1 Growth characteristics on eight conditions

C. saccharolyticus was able to grow solely on all of the eight tested substrates: glucose, xylose, cellobiose, Avicel, CMC, switchgrass, xylan, and yeast extract. *C. kristjanssonii* was also observed to grow on these eight substrates with the exception of CMC which required the supplementation of yeast extract to grow. These results were consistent with findings for *C. saccharolyticus* (Rainey *et al*., 1994; Yang *et al*., 2009) and for *C. kristjanssonii*, (Bredholt *et al*., 1999; Ling, 2012) except for CMC. Culture supernatants from *C. kristjanssonii* have been shown to be able to hydrolyze CMC (Blumer-Schuette *et al*., 2010) but the organism itself has not been shown to grown on CMC.

In general, when the cultures were grown in large scale, there were similar growth curves to the small scale cultures, but with a significantly longer lag phase. One reason for this is due to the setup of the growth apparatus and incubation equipment. The small cultures were grown in air-tight bottles while the large scale cultures were grown in degassed flasks which were clamped shut and placed inside a large incubator. The large flasks appeared to be air-tight, but were suspected to have some leakage, causing oxygen to get into the flasks. Since these organism are strict anaerobes, even a small amount of oxygen, undetectable by the colour indicator resazurin, could affect the growth. Note that the large-scale growth curves were only monitored until late log phase because this was when the cultures were quenched and processed to be analyzed for RNA-sequencing and proteomics at external labs.

As shown in Figure 4, *C. saccharolyticus* grew the best on xylose with the quickest generation time and relatively high final cell densities while xylan gave the highest cell density at about 3.8×10^8 cells/mL. Although *C. saccharolyticus* is known to grow on a broad range of substrates (Vanfossen *et al*., 2009), the results observed here suggest that it prefers pentose sugars such as xylose. *C. kristjanssonii* on the other hand grew the best on cellobiose with the fastest generation times and the highest cell density of about 3.24×10^8 cells/mL, as shown in Figure 6. The worst growth for both organisms was observed on switchgrass (Figures 5 and 7). These results corroborate the findings of previous studies done on these organisms in this lab (Ling, 2012). When comparing the growth rate constants and the generation times with literature findings, the results in this study are much slower due to several reasons. Firstly, in many published papers, the cultures are

grown in nutrient-rich media containing excess carbon sources (Bielen *et al*., 2013), often greater than 10 g/L of sugars (Panagiotopoulos *et al*., 2010), while studies done in this project used the minimal amount of nutrients possible yet maintaining good growth. *C. saccharolyticus* was previously found to have a generation time of approximately 95 minutes (Vanfossen *et al*., 2009) and *C. kristjanssonii* had a generation time of 2 hours (Bredholt *et al*., 1999). However, in this study, the fastest generation times reached were 3 hours for *C. saccharolyticus* and 2.5 hours for *C. kristjanssonii*. The growth rates are much slower due to the exclusion or minimal use of yeast extract in this study. While other studies generally used about 1 g/L to 2 g/L yeast extract (Kádár *et al*., 2004; Yamamoto *et al*., 2011), it was completely omitted from the cultures in this study containing soluble sugars or xylan and only 0.1 g/L was used in cultures with insoluble substrates to maintain a manageable growth rate. The cultures were instead supplemented with a defined vitamin solution (see Appendix A for composition). It was determined that the amount of yeast extract added had a direct correlation on the generation time and the final cell densities of *Caldicellulosiruptor* cultures (data not shown) which is why the growth rates were not comparable to literature values. Even just a small amount of yeast extract, such as the 0.1 g/L use here, was found to vastly reduce the lag phase. Nonetheless, the final cell densities reached were similar to literature values of about 3 x 10^8 cells/mL for both *C. saccharolyticus* (Vanfossen *et al*., 2009) and *C. kristjanssonii* (Blumer-Schuette *et al*., 2010) albeit after a longer time. Table 1 summarizes the growth rates, generation times and final cell densities of both organisms grown under each condition. When looking at this table, it may also be noted that in some growth conditions, there were two noticeable phases of exponential growth. This could be either due to an extended lag phase showing

up as a slight increase in cell numbers, or due to a slow transition into stationary phase. One study mentioned that there is an exponential and a linear growth phase observed for when it increases in hydrogen productivity and has constant hydrogen productivity, respectively (Martinez-Porqueras, Wechselberger, & Herwig, 2013). For a few of the complex substrates, this can be explained by the fact that perhaps certain sugars are used up first and then the growth rate changes as the polymers break down into single sugars. For example, Avicel or crystalline cellulose, is made up of long polysaccharide chains of glucose units. When exoglucanases act on these chains, oligosaccharides or disaccharides may break off and be utilized for growth while other enzymes may further catabolize these sugars into monosaccharides which are used for growth when available.

Overall, the results of this experiment confirmed the fact that *C. saccharolyticus* could grow on a broad range of cellulosic substrates (Vanfossen *et al*., 2009) but it also shows the capability of *C. kristjanssonii* to grow on these as well. It also contributes research for growth on CMC, which is important with regards to endo-acting cellulases. Comparative growth also showed that both organisms grew equally well on simple sugars including glucose, xylose, and cellobiose, as well as hemicellulose (xylan), but growth on insoluble substrates was relatively poor because carbohydrates were not as readily available. These results merit enzyme assays for comparing cellulase activity as well as proteomics and RNA-sequencing to examine the regulation of transport and metabolic pathway enzymes when grown on different substrates.

4.1.2 Endoglucanase, exoglucanase, and xylanase activity

Cellulase enzymes are an important characteristic of *Caldicellulosiruptor* spp. as they allow these organisms to hydrolyze cellulosic materials to release sugars (saccharification) which they can also ferment simultaneously (Talluri, Raj, & Christopher, 2013). Enzyme assays were performed to experimentally determine the presence of functional cellulases native to *C. saccharolyticus* and *C. kristjanssonii*. The goals of this experiment were to determine whether enzymes persist inside or are secreted outside of cell as well as to study their comparative activity depending on growth conditions including: glucose, xylose, Avicel, CMC, and xylan. It should be noted that the enzyme assays in this study do not include the cellulases that may be bound to the cell walls. The results of the enzyme assays demonstrate that exoglucanase, endoglucanase, and xylanases are expressed by both of these organisms and that they are functional. While this has been done for *C. saccharolyticus* (VanFossen *et al*., 2011), enzyme activity studies have not yet been performed *on C. kristjanssonii*. Previous cellulase enzyme studies for *C. saccharolyticus* have been on recombinant enzymes, overexpressed in *Escherichia coli* rather than the native enzymes studied here. For this reason, the specific activity values may not be comparable.

In a study on *Caldocelum saccharolyticum* (Sissons, Sharrock, Daniel, & Morgan, 1987), prior to it being renamed *C. saccharolyticus* in 1994 (Rainey *et al*., 1994), recombinant xylanase specific activity from cell-free extracts was observed to be 2.7 U/mg (Lüthi, Jasmat, & Bergquist, 1990) while the native xylanases seen here only had about half of that at 1.3 U/mg specific activity at most in the CFE when grown on CMC. The recombinant xylanase study by Lüthi, *et al*. did not look at the secretome, but the xylanase activity from

the supernatant samples in this study showed formidable activity at 28.4 U/mg when grown on Avicel and still relatively high activity at 4.6 U/mg when grown on CMC. The remarkably high xylanase activity in the supernatant suggests that these enzymes are secreted outside of the cell and that growth on Avicel stimulates xylanase production. This is not unusual since Avicel has been recorded to induce the synthesis of both endoglucanases and xylanases in some microorganisms (Stafford, 1995).

Another study showed the endoglucanase (GH5) activity from recombinant CelB genes that were cloned, expressed, and purified from *E. coli* as well. The GH5 specific activity was measured to be about 0.71 U/mg (Park *et al*., 2011) while the specific activity measured in this experiment was greater at about 1.8 U/mg in the CFE when grown on CMC. The supernatant samples were also similar at 1.6 U/mg when *C. saccharolyticus* was grown on CMC (see Table 4). Though there were no exoglucanase enzyme activity studies done on *C. saccharolyticus*, another closely related organism, *Anaerocellum thermophilum*, now renamed *Caldicellulosiruptor bescii* (Yang *et al*., 2010), had its CelA gene cloned into *E. coli* to test for exoglucanase activity. It was found that this recombinant exoglucanase had about 0.6 U/mg specific activity from the culture supernatant (Zverlov, Mahr, Riedel, & Bronnenmeier, 1998) but *C. saccharolyticus* in this study showed much higher activity at up to 4.9 U/mg exoglucanase specific activity in the supernatant samples grown on xylan (Table 4).

When examining the total activities, the activity observed tends to be quite proportional between the CFE and supernatant samples, with the supernatant showing more activity in some cases and CFE in others (see Figure 13 and 14). However, when the amounts of

protein in the samples are taken into consideration, much more specific activity was observed in the supernatant of the cultures than the cell free extracts, suggesting that the cellulase enzymes are secreted. Other studies have also found that many of the cellulases are excreted with the secretome in *C. saccharolyticus* (Andrews, Lewis, Notey, Kelly, & Muddiman, 2010; van de Werken *et al*., 2008; VanFossen *et al*., 2011). Nevertheless, it is important to keep in mind that the CFE contains many other proteins released from inside the cell, other than the cellulases being examined here. Since the samples are not purified proteins, the specific activities stated here are likely underestimated.

When looking at the specific activity for *C. saccharolyticus*, the highest exoglucanase activity was seen in the supernatant when grown on xylan. This suggests that growth on xylan stimulates exoglucanases. This phenomenon was also observed with *Thermotoga petrophila* (Chen, 2011). As expected, endoglucanase activity was the highest in the supernatant samples when *C. saccharolyticus* was grown on CMC but xylanase activity was the highest in the Avicel supernatant sample. As mentioned earlier, growth on crystalline cellulose may induce xylanase activity (Stafford, 1995). When compared to genome annotations found on NCBI, the proportion of exoglucanase, endoglucanase, and xylanase activity corresponds to the number of sequences found for each protein.

In *C. kristjanssonii* CFE samples, each of the enzymes had the highest activity in the expected corresponding growth conditions; exoglucanase activity was highest in Avicel, endoglucanase activity was highest in CMC, and xylanase activity was the highest in the xylan growth condition (Ozioko, Ikeyi, & Ugwu, 2013). However, even though exoglucanase and xylanase activity is higher in the supernatant, endoglucanase activity was similar but

slightly higher in the CFE. When grown on CMC, *C. kristjanssonii* had a specific endoglucanase activity of about 1.8 U/mg in the CFE while in the supernatant it was about 1.5 U/mg, Further studies must be done to examine the cellulolytic ability of *C. kristjanssonii* and to determine whether or not cellulase enzymes are secreted. When compared to genome annotations found on NCBI, there were more xylanases than exoglucanases or endoglucanase, and this corresponds to the proportion of xylanase activity which was generally higher than the exoglucanase and endoglucanase activities in this assay.

4.1.3 Growth affected by ABC transporters and glycoside hydrolases

RNA-sequencing was done to examine the transcriptomes of each organism and how it is affected when grown on different cellulosic substrates. The three pairwise comparisons discussed here are: glucose versus xylose, glucose versus cellobiose, and xylose versus xylan, for each organism. These conditions grew the best and were easier to extract high-quality RNA from. Although there were 2740 different sequences found for *C. saccharolyticus* and 1923 sequences found for *C. kristjanssonii*, their expression levels were statistically analyzed and those with the highest upregulation in comparison to another growth condition were considered. Proteomics testing was done to examine the presence and regulation of all the proteins expressed when grown on different substrates. The presence of all the enzymes shown in the metabolic pathways in Figure 17 was confirmed experimentally through inspection of the locus tags expressed in the RNA-seq and proteomics datasets. For proteomics, 5 pairwise comparisons were chosen including: glucose versus xylose, glucose versus cellobiose, xylose versus xylan, glucose versus Avicel,

and Avicel versus CMC. These comparisons were done on both *C. saccharolyticus* and *C. kristjanssonii* and in effect compare C6 versus C5 monomers, C6 monosaccharide versus C6 disaccharide, C5 monosaccharide vs C5 polymer, C6 monosaccharide versus C6 polymer, and two C6 polymers, insoluble versus soluble, respectively. The following table (Table 22) summarizes the enzymes that were seen to be upregulated in both the RNA-seq and proteomics datasets in the same conditions for easy viewing.

Table 22. Common upregulated enzymes in RNA-sequencing and proteomics datasets

			Proteomics		RNA-seq	
Condition	**Locus Tag**	**Enzyme Name**	**Expression Value**	**Z-score**	**Expression Value**	**Z-score**
C. saccharolyticus						
Xylose vs Glucose	Csac_0586	glycoside hydrolase, family 3 domain protein	21.8	3.49	10.9	1.71
Glucose vs Xylose	Csac_2441	PfkB domain protein	20.6	5.72	12.2	3.08
	Csac_1173	Xylose isomerase domain protein TIM barrel	20.6	3.19	11	1.21
Cellobiose vs Glucose	Csac_0681	ABC−type sugar transport system periplasmic	27.9	11.3	17	7.81
Glucose vs Cellobiose	Csac_2441	PfkB domain protein	20.6	6.61	12.2	4.14
	Csac_1102	glycoside hydrolase, family 3 domain protein	20.0	3.74	10.6	2.49
	Csac_2504	Monosaccharide−transporting ATPase	19.5	3.14	13.6	3.02
	Csac_0203	alpha amylase, catalytic sub domain	21.2	2.52	11.3	1.27
Xylan vs Xylose	Csac_0204	glycoside hydrolase, family 10	21.4	2.93	12.5	1.22
	Csac_0586	glycoside hydrolase, family 3 domain protein	24.9	2.51	13	1.29
	Csac_0798	xylulokinase	23.4	2.19	12.5	1.23
	Csac_1154	Xylose isomerase domain protein TIM barrel	25.5	2.12	15.3	1.48
	Csac_0426	alpha amylase, catalytic region	18.8	1.61	12.4	1.32
C. kristjanssonii						
Xylose vs Glucose	Calkr_1997	xylose isomerase	25.9	4.96	14.1	1.66
	Calkr_0097	iron−containing alcohol dehydrogenase	27.2	3.68	15.7	1.53
Glucose vs Xylose	Calkr_1677	flagellin domain protein	25.1	2.57	15	2.41
	Calkr_2248	iron−containing alcohol dehydrogenase	23.1	2.56	13.4	1.26
Cellobiose vs Glucose	Calkr_2205	glycosyltransferase 36	24.3	2.14	14.5	2.42
Xylan vs Xylose	Calkr_2396	glycoside hydrolase family 3 domain protein	24.1	3.44	14.4	2.58
	Calkr_1677	flagellin domain protein	23.8	1.54	16.9	2.73

This table summarizes all of the relevant enzymes there were seen to be upregulated in both the RNA-seq and proteomic datasets when grown in the same condition. Note that the enzymes and values stated are upregulated in the first growth condition in comparison to the second growth condition stated in the “Condition” column.

When *C. saccharolyticus* was grown on glucose, PTS transporters (Csac_2437, Csac_2438, and Csac_2439) as well as a monosaccharide-transporting ATPase (Csac_2504) were upregulated compared to xylose and cellobiose, suggesting that these transport 6-carbon sugars into the cell, whereas another monosaccharide-transporting ATPase (Csac_0241) is responsible for transporting 5-carbon sugars as it was upregulated on xylose compared to glucose and xylan. While grown on glucose, a xylose ABC-transporter (Csac_2506) was also upregulated. However, as mentioned in the introduction, Csac_2504 and Csac_2506 were found to be able to transport both glucose and xylose (Vanfossen *et al*., 2009).
C. saccharolyticus produced more transcripts for an ABC-type sugar transporter (Csac_0681) when grown on cellobiose in comparison to glucose, suggesting that this protein transports cellobiose or other disaccharides. This protein was also upregulated when *C. saccharolyticus* was grown on cellobiose, xylan, Avicel, and CMC compared to their monosaccharide counterparts suggesting that it transports oligosaccharides or polysaccharides. It was also observed that ABC-type transporter (Csac_0297) was upregulated in both cellobiose and xylan conditions, indicating that this transporter brings in disaccharides or polysaccharides. ABC-type transporters (Csac_2696 and Csac_2435) may be responsible for transporting xylan or polysaccharides into the cell as these transcripts were upregulated on xylan compared to xylose.

Like *C. saccharolyticus*, *C. kristjanssonii* grown on glucose had a xylose ABC transporter (Calkr_2435) upregulated as well. It may be possible for ABC transporters to act on both glucose and xylose since it was found that other bacterial ABC transporters can take up multiple sugars (Ajdić & Pham, 2007). Likewise, several ABC-type sugar transport systems

(Csac_2504, Csac_2506, and Csac_0681) are able to transport many different polysaccharides. This may be similar to *C. saccharolyticus* transporters that are able to uptake more than one type of sugar.

In terms of cellulases, *C. saccharolyticus* produced more RNAs for β-glucosidase (Csac_1089), cellulose 1,4-beta-cellobiosidase (Csac_1078 and Csac_1079), and glycoside hydrolases family 48 (Csac_1076) and family 20 (Csac_2539) transcripts when grown on cellobiose compared to glucose which makes sense because these enzymes are used to hydrolyze this disaccharide into monosaccharide units. As for proteomic analyses, many cellulases are expectedly upregulated according to their corresponding growth condition but several glycoside hydrolases listed do not match their growth condition. It is suspected that these only appear to be upregulated in relation to the other growth condition or it may just be constitutively expressed as mentioned in literature (VanFossen *et al*., 2011). These enzymes include: alpha amylase (Csac_2441 and Csac_0426) and glycoside hydrolases family 43 (Csac_0359), clan GH-D (Csac_1118) and family 3 (Csac_1102) when grown on glucose compared to xylose which expressed glycoside hydrolase family 3 (Csac_0586) and family 10 (Csac_0204); glycoside hydrolase family 3 (Csac_1102), family 2 (Csac_2734) and family 88 (Csac_0360) and alpha amylase (Csac_0203) when grown on glucose in comparison to cellobiose; and glycoside hydrolase family 2 (Csac_2734) on xylose compared to xylan which expressed alpha amylases (Csac_0203 and Csac_0426) as well as glycoside hydrolases family 10 (Csac_0204), family 3 (Csac_0586), family 2 (Csac_2686 and Csac_0362), clan GH-D (Csac_1118) and family 43 (Csac_0359) for *C. saccharolyticus*. Other cellulolytic enzymes were upregulated as expected in response to their growth substrate.

A β-glucosidase (Csac_1089) was upregulated when *C. saccharolyticus* grew on cellobiose compared to glucose, which was expected since β-glucosidase can break the bond between two glucose units in cellobiose. Glycoside hydrolase family 10 (Csac_0204), family 3 (Csac_0586) and family 43 (Csac_0359) were upregulated in xylan compared to xylose which makes sense since each of these GH families contain enzymes with xylanase and β-xylosidase activity. Moreover, *C. saccharolyticus* expressed upregulated glycoside hydrolases from family 10 (Csac_0204), family 3 (Csac_0586), and family 2 (Csac_0129 and Csac_2734) when grown on Avicel compared to both glucose and CMC which makes sense since GH family 10 contains enzymes with cellobiohydrolase activity, family 3 has exoglucanase activity, and family 2 has β-glucosidase activity.

Furthermore, flagella were observed in many previous studies and are thought to aid in substrate adhesion (Blumer-Schuette *et al*., 2012; Ling, 2012; Lochner *et al*., 2011). An interesting discovery is that flagellar proteins were upregulated when grown on glucose (Csac_1698, Csac_1699, Csac_1700, and Csac_1264) in comparison to xylose, but not when compared to cellobiose, suggesting that either hexose sugars promote flagella production or pentoses may possibly inhibit its production. *C. saccharolyticus* has previously been reported to have a set of flagellar genes even though they are non-motile bacteria (van de Werken *et al*., 2008). When grown on xylan, many of the same flagellar proteins (Csac_1698, Csac_1699, and Csac_1700) were upregulated but in this case, it is thought that these proteins are used for substrate attachment rather than motility (Blumer-Schuette *et al*., 2012). When examining the flagellar proteins in proteomic data, the results differ from the RNA-seq data. Flagellar protein (Csac_1698) was upregulated when grown on glucose

which was also seen in RNA-seq data, but surprisingly, both cellobiose and xylose had a flagellar M-ring protein (Csac_1257) upregulated in comparison to glucose. This contradicts the idea that hexoses stimulate or pentoses inhibit flagellar production. Furthermore, the fact that flagellar proteins were actually translated, opposes the previous thought that they were interrupted by a stop codon (van de Werken *et al*., 2008). Although the true function of these proteins is still unknown since these bacteria are observed to be non-motile, it has been suggested that flagellar structures are produced for cell adhesion or attachment to substrates to form biofilms (Lochner *et al*., 2011).

When examining the RNA transcripts of *C. kristjanssonii*, another phenomenon similar to that of *C. saccharolyticus*, was observed where flagellar proteins (Calkr_1666, Calkr_1677, Calkr_1689, Calkr_1690, Calkr_1691, and Calkr_1693) are upregulated when grown on glucose when compared to xylose, but not in comparison to cellobiose. In fact, when grown on cellobiose, flagellar proteins (Calkr_1664 and Calkr_1677) are highly expressed in comparison to glucose. Similarly, *C. kristjanssonii* produced several flagellar proteins (Calkr_1666, Calkr_1677, Calkr_1689, Calkr_1690, and Calkr_1693) when grown on xylan, signifying attachment to insoluble substrates. Upon examination of *C. kristjanssonii* proteomic data, the flagellar protein expression is more similar to the RNA-seq data where flagellin (Calkr_1677) and flagellar hook proteins (Calkr_1666 and Calkr_0415) were upregulated on glucose compared to xylose and cellobiose. However, this data still does not conclude that hexoses stimulate flagella production, and perhaps could be explained by the fact that these samples represent a snapshot of the organisms' physiology at the mid-log phase of growth. Since some flagellar proteins may already be translated, they may no

longer need to produce RNAs, or vice versa where some RNAs are transcribed here in this data, but not yet translated into proteins. On the contrary, in a study on *C. obsidiansis*, another *Caldicellulosiruptor* species, it was found that growth on cellobiose repressed flagella production and that more flagella were synthesized for attachment when grown on crystalline cellulose (Lochner *et al*., 2011). This was also observed with the flagellar protein (Csac_1257 and Csac_1268) expression when *C. saccharolyticus* was grown on Avicel, a crystalline cellulose, in comparison to glucose or CMC which are both soluble. *C. kristjanssonii* did not synthesize more flagella when grown on Avicel, but it did highly express a flagellin protein (Calkr_1677) when grown on xylan, another insoluble substrate, in comparison to xylose. This was also seen in the RNA-seq data.

Regarding cellulose catabolism, RNAs for polysaccharide deacetylase (Calkr_1447) was strangely upregulated in xylose compared to glucose and xylan. This enzyme is an endoxylanase that is capable of hydrolyzing glycosidic bonds in xylan but it is even upregulated when grown on xylose in comparison to xylan. In nature, this may make sense since these carbohydrates, cellulose and hemicellulose are always together in plant materials. Glycoside hydrolase family 10 (Calkr_1929) was observed to be upregulated in both cellobiose and xylan compared to glucose and xylose respectively. This suggests that glycoside hydrolase 10 may be capable of breaking down both 6-carbon and 5-carbon polysaccharides. *C. kristjanssonii* also expressed proteins in glycoside hydrolase family 2 (Calkr_0218) when grown on glucose compared to xylose which had glycoside hydrolase family 3 (Calkr_2396). Nonetheless, a majority of the other cellulases expressed do make sense and correspond to their supplied substrate. In proteomic datasets, glycoside

hydrolase family 3 (Calkr_2396) was upregulated in xylan compared to xylose because this family of GHs contains enzymes with β-xylosidase activity. Glycoside hydrolase family 3 (Calkr_2396) and family 2 (Calkr_0218) were also upregulated on Avicel in comparison to both glucose and CMC since these GHs have β-glucosidase and exoglucanase activity. CMC on the other hand had alpha amylase (Calkr_0143), beta-galactosidase (Calkr_2207) and glycosidase (Calkr_1514) upregulated in comparison to Avicel which make sense as well since these enzymes help to catabolize the complex CMC polymer. Remarkably, even though *C. kristjanssonii* is classified as a weakly cellulolytic since it is missing GH48 (Blumer-Schuette *et al*., 2012), it was observed to utilize cellulosic substrates just as well as *C. saccharolyticus* in the growth and end product detection experiments of this project.

4.2 Metabolism of *Caldicellulosiruptor* spp.

The metabolism of an organism can be affected by the type of substrates they are grown on. In this section, the metabolism of *C. saccharolyticus* and *C. kristjanssonii* are examined through the end products they produce and the regulation on the enzymes and biochemical pathways resulting from growth in different conditions.

4.2.1 The effect of substrates on metabolism and end product formation

Besides examining growth, metabolites and end products were examined through HPLC and GC to better understand the physiology of *C. saccharolyticus* and *C. kristjanssonii*. When looking at the graphs in Figures 8-11, the expected trends are seen where the substrates decrease as the cell densities increase along with metabolite production.

As seen in Table 2, the hydrogen to acetate ratio was generally 1:1 for both organisms with some variation in individual conditions. However, in literature, it was observed that the hydrogen to acetate ratio for *C. saccharolyticus* was 2:1 or even greater (Shen, Zhang, Song, Wang, & Zeng, 2013). It is expected that 2 moles of acetate and 4 moles of hydrogen can theoretically be produced from fermenting 1 mole of glucose (de Vrije *et al*., 2007). Conversely, Table 2 shows that acetate to hydrogen yields in most growth conditions yield a 1:1 ratio with some cases where acetate yield is even greater than hydrogen. One cause of this error may be due to gas leakage from the culture bottles while the liquid medium composition remains the same. Moreover, it was observed that the acetate to carbon dioxide ratio is 2:1 for *C. saccharolyticus* although it is expected to be 1:1 (Kádár *et al*., 2004) whereas *C. kristjanssonii* samples had the expected 1:1 ratio. *C. saccharolyticus* had a higher acetate to carbon dioxide ratio because the cultures bottles possibly had some gas

leakage as mentioned previously. When comparing these ratios in terms of growth substrates, soluble sugars tended to be more consistent and produced more end products, but on complex substrates such as Avicel, CMC, and switchgrass, where poor growth was evident, the ratios fluctuated erratically without a trend to be seen. For *C. saccharolyticus*, hydrogen and carbon dioxide gas were observed to be accumulating at a 2:1 ratio which was expected from complete sugar metabolism (Van Niel *et al*., 2002). This was generally true for *C. saccharolyticus* with some of the carbon dioxide being dissolved into the culture medium. However, for *C. kristjanssonii* cultures, the hydrogen to carbon dioxide ratio appeared to almost be 1:1. One possible reason for this is that some of the CO_2 could be dissolved in the liquid medium. Another explanation for this is that *C. kristjanssonii* is suspected to utilize alternate pathways for carbohydrate metabolism and fermentation. While *C. saccharolyticus* is known to use the EMP pathway for glycolysis (de Vrije *et al*., 2007), it may be possible that *C. kristjanssonii* uses the ED pathway or other modified pathways. Other thermophilic microorganisms have been found to use either the EM or ED pathways for glycolysis and in fact some use both EM and ED pathways for glucose conversion (Selig, Xavier, Santos, & Schonheit, 1997). Furthermore, there have been genomic annotations of enzymes in the ED pathway, such as KDPG aldolase, that were found in *C. kristjanssonii*.

When examining the end products, it is also observed that *C. saccharolyticus* did not produce detectable amounts of lactate while *C. kristjanssonii* did. According to other papers, *C. saccharolyticus* shifts its metabolism from acetate to lactate production when it reaches stationary phase (Willquist *et al*., 2010). Although from the growth curves, the cultures did

appear to reach stationary phase, the final cell densities are somewhat lower than expected and compared to small scale growth. This suggests that growth stalled even though it may not have reached stationary phase. Culturing *C. saccharolyticus* is sometimes unpredictable (Verhaart, 2010) but this behaviour is likely caused by end product inhibition or another unknown issue with the cultures. One explanation for this discrepancy could be due to the fact that the culture conditions used were different than those in the literature. In this experiment, no yeast extract or trypticase was used and the carbon sources were minimized to 2 g/L as opposed to 10 or 20 g/L, resulting in an altered metabolism and end products. Another possible source of error that may have contributed to this is the fact that the culture bottles used for end product detection had a greater disturbance than the ones where only cell counts were done. These bottles were sampled for cell counts, 1 mL of media removed for HPLC, and the gas phase sampled for GC. Not only is there a higher chance of temperature decrease during sampling intervals, but there are also more holes punctured into the bottle stoppers, which could increase the risk of oxygen getting into the anaerobic bottles. Moreover, the holes in the stoppers could have contributed to a release of pressure which normally causes the cultures to shift from acetate to lactate production. It has been reported that *C. saccharolyticus* shifts to lactate production when the H_2 partial pressure is greater than 10kPa (van Niel *et al*., 2003).

A carbon balance was performed as displayed in Table 3. The results show that for each of the soluble sugars, the carbon balance was near 100% for both *C. saccharolyticus* and *C. kristjanssonii*. This was expected as other studies have shown a near 100% recovery as well (Kádár *et al*., 2004). Nonetheless, it should be noted that in some conditions, the yield was

slightly greater than 100%. This error can be attributed to an error in the cell biomass measurements. Since the pellets were still wet, the biomass weight was likely overestimated.

Furthermore, several sugars were observed as detectable residual sugars after fermentation. These sugars include cellobiose converted from CMC, glucose released from switchgrass, and xylose released from xylan. Each of these sugars being released make sense as they are units that make up the corresponding complex cellulosic polymers. Detecting these sugars shows that these organisms have cellulases to hydrolyze the complex polymers. Additionally, arabinose is detected when *C. kristjanssonii* is grown on xylose or xylan. This is unusual since there are no published papers regarding the ability of *Caldicellulosiruptor* spp. to convert xylose into arabinose, although the same phenomenon occurred in the previous results from this lab (Ling, 2012). Since there are no genome annotations for UDP-arabinose epimerase found, further research needs to be done in order to determine if there are other enzymes produced by *C. kristjanssonii* that can convert these two pentose isomers.

Another intriguing phenomenon observed in both *C. saccharolyticus* and *C. kristjanssonii* cultures was the appearance of citric acid in cultures that were grown on xylan. These organisms are not known to produce citric acid and it is suspected that instead of the microorganisms converting the xylan, the acid is released from the breakdown of xylan itself either due to heat or other reactions in the vessel. Additionally, the acid detected by the HPLC could be another unknown acid that happens to have the same retention time as citric acid. One highly plausible compound that was detected here is ferulic acid. Ferulic

acid is a component in the structure of xylan (Collins, Gerday, & Feller, 2005) and has a similar molecular weight to citric acid. It is also reported that *C. saccharolyticus* has a gene encoding ferulic acid esterase or feruloyl esterase which cleaves this site (Cao, 2012). This finding holds significance in future research on cellulose-degrading organisms since a recent study established that the addition of feruloyl esterases can improve the hydrolysis of other cellulosic materials (Braga *et al*., 2014).

In summary, major end products from *Caldicellulosiruptor* sp. include acetate, lactate, ethanol, hydrogen and carbon dioxide and this indicates the use of fermentative pathways for the production of these metabolites. Although the end products detected with *C. saccharolyticus* were inconclusive here, further insight was gained with *C. kristjanssonii*. Carbon balances were near 100% and it is recommended that ferulic acid should be examined when working with xylan and its metabolites.

4.2.2 Genome annotations and putative biochemical pathways

Key enzymes pertaining to cellulose degradation, carbohydrate transport and metabolism, as well as fermentation were searched for in the genomes of both *C. saccharolyticus* and *C. kristjanssonii*. From the results listed in Table 5, it can generally be observed that cellulases were annotated for both *C. saccharolyticus* and *C. kristjanssonii* but there were no locus tags listed for *C. saccharolyticus* xylanases. However, although there were no locus tags available from GenBank, there were 7 xylanases annotated as NCBI reference sequences. Both organisms have ABC and PTS transporters for uptaking sugars into the cell, but it appears that ABC transporters are more abundant. The upregulation of ABC transporters is also evident in the RNA-seq and proteomic datasets which are discussed later on. As for central

carbohydrate metabolism, phosphofructokinase from the EM pathway, KDPG aldolase from the ED pathway, and pyruvate kinase from glycolysis as well as transketolase and transaldolase from the PPP were all present in both microorganisms. Pyruvate ferredoxin oxidoreductase, hydrogenase and alcohol dehydrogenase were also annotated for both organisms, indicating fermentation and possible production of hydrogen and ethanol. Lactate dehydrogenase on the other hand, was found in *C. kristjanssonii* but not in *C. saccharolyticus*. Although this corresponds to the lack of lactate detected in the end product detection experiments, this is erroneous as it contradicts many other findings stating that *C. saccharolyticus* produces lactate in stationary phase (Willquist *et al*., 2010). Again, although there were no locus tags listed in GenBank, there was one lactate dehydrogenase protein listed as an NCBI reference sequence. Since most of these proteins exist theoretically from sequence analysis, the results listed here served as a guide for RNA sequencing and proteomics analyses.

From all of the experiments performed in this project, including end product detection, genomic analysis, RNA-seq., proteomics, and enzyme assays, a putative metabolic map could be generated. These biochemical pathways, shown in Figure 17, can help to better understand the carbohydrate metabolism of *C. saccharolyticus* and *C. kristjanssonii* so that they may be used for hydrogen production.

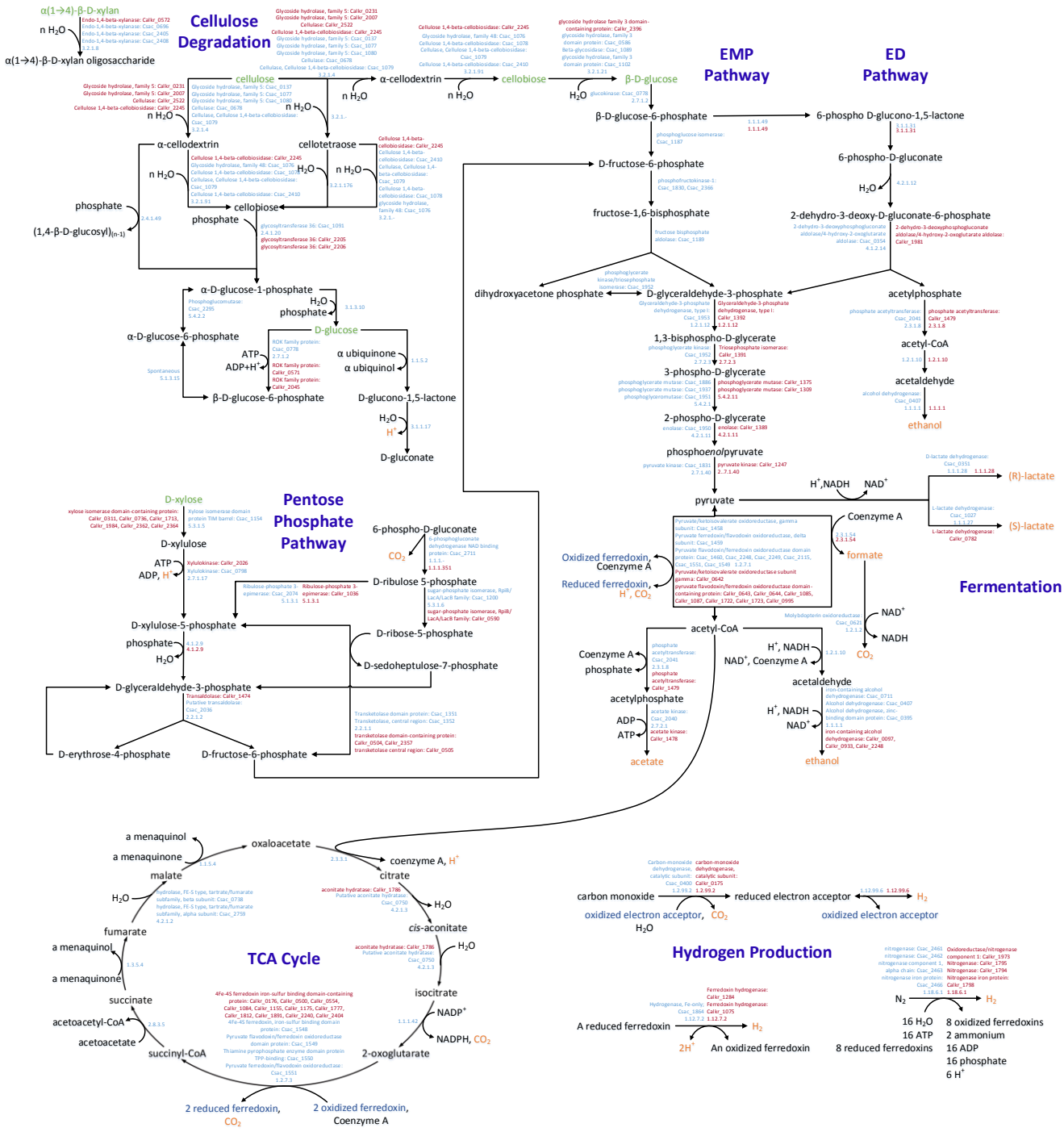


Figure 17. Metabolic map of central carbohydrate degradation and fermentation pathways in *C. saccharolyticus* and *C. kristjanssonii*. This diagram shows the biochemical pathways involved in cellulose degradation, carbohydrate metabolism, and fermentation. Genome annotations from *C. saccharolyticus* are shown in blue while annotations from *C. kristjanssonii* are show in red. Starting substrates are also shown in green while detected end products are shown in orange. Note that the TCA cycle is included here because there were several enzymes annotated in the genome that correspond with this pathway. Please see Appendix D for a closer look of these pathways.

4.2.3 Regulation of central metabolic and peripheral fermentative pathways

Depending on growth conditions, proteomic and transcriptomic datasets were analyzed to see the regulation of metabolic pathway enzymes. The metabolic map shown in in Figure 17 proposes the biochemical pathways taken by both *C. saccharolyticus* and *C. kristjanssonii* for cellulose degradation, carbohydrate metabolism, and fermentation. Such diagrams displaying the metabolic pathways for *C. saccharolyticus* have been shown previously (Bielen *et al*., 2013; van de Werken *et al*., 2008) but not for *C. kristjanssonii*. The locus tags displayed for the pathway enzymes were obtained from annotations of each organisms' genome available on NCBI and BioCyc. All of the locus tags presented have also been verified to be transcribed and translated into proteins as demonstrated with the RNA-seq and proteomic analyses. This map in Figure 17, along with enlarged pathways shown in Appendix D, can be used to help visualize the pathway enzymes discussed below. For the purposes of this project, data within the 95% confidence interval ($p < 0.05$) was considered but it is noted that it may not be statistically significant since many false positives may arise from high-throughput data. However, this data is still used since the objective of this research was to identify candidate genes and proteins, and to compare similar changes between the RNA and proteins.

Firstly, looking at substrate conversion and pathways feeding into central metabolism, several group 1 glycosyltransferases (Calkr_2204, Calkr_1588, and Calkr_0498) were upregulated when grown on xylose in comparison to glucose and Calkr_1588 is upregulated in xylose compared to xylan. However, this finding is unusual since this enzyme is normally used to transfer sugar groups and it allows cellobiose to be converted

into glucose-1-phosphate (as shown in Figure 17) which can later be introduced into glycolytic pathways. Xylose isomerases also contribute to substrate conversion in the pentose phosphate pathway. Strangely, several xylose isomerase proteins (Csac_ 1173, Csac_0350, and Csac_0390) were upregulated even when *C. saccharolyticus* was grown on glucose compared to growth on xylose itself. Xylose isomerases are usually used to channel in xylose into central metabolism via the PPP (van de Werken *et al*., 2008) but it can also be used to convert glucose into fructose (de Vrije *et al*., 2009). Previous studies have found that although *C. saccharolyticus* is able to utilize a broad range of substrates, it prefers fructose over xylose or glucose. Therefore, the production of xylose isomerase during growth on glucose may be occurring to feed this preference (Vanfossen *et al*., 2009). Again, it is interesting to see xylose isomerase (Csac_1154) and xylulokinase (Csac_0798) upregulated when grown on xylan even when compared to xylose. The same phenomenon is also seen in *C. kristjanssonii* where xylulokinase (Calkr_2026) is upregulated when grown on xylan in comparison to xylose. It is expected that these PPP enzymes would be present when either organism is grown on xylan but perhaps they are even more highly expressed than in xylose because each xylan molecule contains many xylose units (Collins *et al*., 2005). Therefore, when it is broken down, there is much more xylose available and the cells must accommodate by producing these PPP enzymes. For *C. saccharolyticus* grown on glucose, xylose isomerase (Csac_1173) was upregulated in proteomics just as it was in RNA-sequencing against xylose. Xylose isomerases (Csac_1173 and Csac_0396) were also upregulated on CMC compared to Avicel. With *C. kristjanssonii*, xylose isomerases (Calkr_0311) was upregulated in glucose compared to xylose, but when grown on xylose, PPP enzymes including xylose isomerase (Calkr_1997), xylulokinase (Calkr_2026), and

transketolases (Calkr_0505 and Calkr_0504) were upregulated as expected in comparison to growth on glucose. *C. kristjanssonii* also expressed xylulokinase (Calkr_2026) and xylose isomerase (Calkr_1997) when grown on Avicel in comparison to glucose and even more xylose isomerase (Calkr_0311 and Calkr_1997) enzymes when grown on CMC in comparison to Avicel. When grown on xylose, this made sense as xylulokinase (Calkr_2026) was also upregulated, indicating use of the PPP. As discussed earlier, the xylose isomerases may be expressed to convert glucose into fructose (de Vrije *et al*., 2009) which may be the case for the C6 substrates glucose, Avicel, and CMC. Xylulokinase (Csac_0798) and xylose isomerase (Csac_1173 and Csac_1154) were upregulated when *C. saccharolyticus* was grown on xylan compared to xylose which corresponds to the results found in RNA-sequencing. *C. kristjanssonii* grown on xylan also had an upregulated xylose isomerase (Calkr_0311) over growth on xylose. Again, these are enzymes belonging to the PP pathway and because each xylan polymer contains many xylose units, it is thought that these enzymes will help to funnel in the xylose into central metabolism (van de Werken *et al*., 2008).

Next, regarding glycolytic pathways, different RNA sequences were observed for phosphofructokinase proteins when grown on glucose (Csac_2441) and xylose (Csac_2366 and Csac_2720) suggesting that the EMP pathway is used in both cases. In terms of proteomic data, the same phosphofructokinase (Csac_2441) was upregulated when *C. saccharolyticus* was grown on glucose in comparison to both xylose and cellobiose. This makes sense since PFK catalyzes the conversion of fructose-6-phosphate into fructose-1,6-bisphosphate, both of which are glucose intermediates in the EM pathway. 6-

phosphogluconate dehydrogenase (Csac_2711) was also upregulated when grown on glucose in comparison to xylose which corresponds to the fact that this enzyme converts 6-phosphogluconate, a glucose intermediate, into ribulose-5-phosphate in the PP pathway as opposed to the route xylose would take (see Figure 17). Fructose-1,6-bisphosphate aldolase (Csac_1189), an EMP enzyme, was oddly upregulated in xylan compared to xylose. Since both of these substrates comprise of xylose units, it may just be that xylan has more sugars converted through the PPP, as demonstrated with the upregulated xylose isomerase enzymes described earlier, and then fed through glycolysis in central metabolism. Similarly, *C. kristjanssonii* grown on glucose and xylan both expressed more phosphofructokinase (Calkr_1980) in comparison to xylose. Another enzyme worth mentioning in the glycolytic pathway is pyruvate phosphate dikinase (PPDK). Although this enzyme was not considered to be upregulated by at least two-fold in any of the pairwise comparisons completed, it was observed to be expressed both as RNA transcripts and as proteins in all of the conditions. One locus tag for PPDK was found in each organism (Csac_1955 and Calkr_1394) but they were expressed at relatively high levels in each condition. This enzyme is of interest because it has been recently found to be dependent on inorganic pyrophosphate (PPi), an energy carrier, and if it is highly expressed, it means that the cultures are in exponential growth phase as opposed to low levels observed in stationary phase (Bielen *et al*., 2010).

Looking further past central carbohydrate metabolism, several RNAs for fermentative enzymes were seen upregulated in certain conditions. When grown on cellobiose, it was observed that lactate dehydrogenases (Csac_1027 and Calkr_0782) were upregulated in both *C. saccharolyticus* and *C. kristjanssonii* when compared to glucose which may be

indicative of substrate preference where growth reaches stationary phase quicker and they shift to lactate production as mentioned earlier. In proteomic datasets, lactate dehydrogenase (Calkr_0782) was shown to be upregulated when *C. kristjanssonii* was grown on xylose in comparison to xylan, although this does not correspond to the lactate produced in the HPLC data since xylose was almost the same as xylan (see Table 2). Several alcohol dehydrogenases were upregulated in glucose (Csac_1226) versus xylose (Csac_0763), glucose (Csac_0395) versus cellobiose, glucose (Csac_0622) versus Avicel, and xylan (Csac_1226) versus xylose for *C. saccharolyticus*. Although the amounts of ethanol were too minute to be detected by HPLC in most conditions, *C. saccharolyticus* did produce 0.847 mM of ethanol when grown on xylan. *C. kristjanssonii* also had different alcohol dehydrogenases upregulated on different conditions: glucose (Calkr_2248) versus xylose (Calkr_0097), glucose (Calkr_2248) versus cellobiose (Calkr_0097), xylose (Calkr_0782) versus xylan, Avicel (Calkr_0097) versus glucose (Calkr_2248), and Avicel (Calkr_0097) versus CMC. Even though ethanol was only detectable when *C. kristjanssonii* was grown on xylan as well, previous literature states that ethanol is one of the major products of *Caldicellulosiruptor* fermentation (Isern *et al*., 2013; Willquist *et al*., 2010). Ethanol may have been produced in other growth conditions but might have been in amounts that were too small to be quantified (<0.5 mM) or partially evaporated due to high temperatures of the cultures. Acetate kinase (Csac_2040) was upregulated when *C. saccharolyticus* was grown on both glucose and CMC compared to Avicel and this corresponds directly to the growth trends as well as the acetate detected by HPLC on their respective substrates. The same result was seen with upregulated acetate kinase (Calkr_1478) when *C. kristjanssonii* was grown on glucose compared to Avicel and the amounts of acetate produced correlates

as well (Table 2). Another interesting enzyme that appeared upregulated in several growth conditions was carbon monoxide dehydrogenase (CODH). Despite the fact that CODH is generally used for carbon fixation and cycling (Vorholt, Kunow, Stetter, & Thauer, 1995), this enzyme also produces hydrogen as a by-product (see Appendix D), which is of interest to this study. Carbon monoxide dehydrogenase (Csac_0400) was seen to be upregulated when *C. saccharolyticus* was grown on xylan compared to xylose and on Avicel compared to CMC. For *C. kristjanssonii*, carbon monoxide dehydrogenase (Calkr_0175) was seen to be upregulated on xylose and cellobiose compared to glucose, and on CMC compared to Avicel. While CODH expression correlated to the amounts of hydrogen detected in *C. kristjanssonii*, this is perhaps a coincidence since there was no correlation seen with *C. saccharolyticus*. Although CODH might contribute to some hydrogen production, there are many other biochemical pathways that produce hydrogen as well.

With the results shown in this research and the contributions of new insights on *C. kristjanssonii*, a better understanding of these organisms' metabolism is gained to help further studies and development in microbial hydrogen production.

5.0 Conclusions

The goals of this research were to gain a better understanding of the metabolism and regulation of *C. saccharolyticus* and *C. kristjanssonii* when grown on different substrates. From the many experiments conducted, these conclusions could be drawn:

1. Growth conditions were successfully designed for future studies on *C. saccharolyticus* and *C. kristjanssonii*. Both organisms are capable of growing on modified media without yeast extract or trypticase, containing 2 g/L glucose, xylose, or cellobiose, or 4 g/L Avicel, CMC, switchgrass or xylan as sole carbon sources with the exception that *C. kristjanssonii* requires yeast extract supplementation for growth on CMC.
2. The best growth was observed when *C. saccharolyticus* grew on xylose with a generation time of 4.2 hours while for *C. kristjanssonii*, the best growth was on cellobiose with a generation time of 2.5 hours. The major end products observed include acetate, lactate, hydrogen, and carbon dioxide.
3. Cellulase enzyme assays reveal that the majority of exoglucanase, endoglucanases, and xylanases are secreted extracellularly. Interestingly, Avicel appears to stimulate xylanase production for *C. saccharolyticus*.
4. Enzymes involved in cellulose degradation, transport, and carbohydrate metabolic pathways that were previously annotated in the genome are now confirmed to be expressed. Flagellar proteins, ABC transporters, glycoside hydrolases, and metabolic pathway enzymes, namely xylose isomerase, fructose-1,6-bisphosphate aldolase, phosphofructokinase, lactate dehydrogenase, and alcohol dehydrogenase, were upregulated when both organisms were grown on certain conditions. The EMP and the PPP are the main pathways used for central metabolism.

While *C. saccharolyticus* is still considered one of the best candidates for hydrogen production, *C. kristjanssonii* can be considered another contender because of its tolerance for higher temperatures and potential for comparable or even higher hydrogen yields.

5.1 Future Perspectives

Although many studies have already been done for *C. saccharolyticus* and the research completed in this project contributes to a better understanding of the metabolism of both *C. saccharolyticus* and *C. kristjanssonii*, there is still a lot more that can be explored in terms of understanding and engineering both organisms. As not much is known about *C. kristjanssonii*, substrate preferences as well as hydrogen production capabilities can be further examined. A step was already taken to improve the growth media composition in this project, but it can still be optimized to achieve higher cell densities and hence greater hydrogen productivity with the lowest amount of substrate required. Particular focus should be placed on reducing yeast extract, if not eliminated entirely, to reduce costs. Furthermore, a large number of hypothetical proteins were observed in both the proteomics and RNA-sequencing data in this study and since these have unknown functions at this time, they could possibly have great potential for optimizing biofuel production once they are identified and manipulated. Genetic engineering can be developed to increase hydrogen production or even just to collect desirable enzymes. The results of this research can be implemented in industrial applications but research and development must also be done to scale up the cultures in commercial reactors. Finally, hydrogen is a clean and renewable resource but before it can be put into practical use, new technology is needed for storage and safe combustion of this biofuel.

References

1. Ajdić, D., & Pham, V.T.T. (2007). Global transcriptional analysis of *Streptococcus mutans* sugar transporters using microarrays. *Journal of Bacteriology*, *189*(14), 5049–5059.

2. Andrews, G., Lewis, D., Notey, J., Kelly, R., & Muddiman, D. (2010). Part I: characterization of the extracellular proteome of the extreme thermophile *Caldicellulosiruptor saccharolyticus* by GeLC-MS2. *Analytical and Bioanalytical Chemistry*, *398*(1), 377–389.

3. Antoni, D., Zverlov, V.V., & Schwarz, W.H. (2007). Biofuels from microbes. *Applied Microbiology and Biotechnology*, *77*(1), 23–35.

4. Bergey, D. H. (1919). Thermophilic Bacteria. *Journal of Bacteriology*, *IV*(4), 301–306.

5. Berka, R.M., Grigoriev, I.V., Otillar, R., Salamov, A., Grimwood, J., Reid, I., Ishmael, N., John, T., Darmond, C., Moisan, M.C., Henrissat, B., Coutinho, P.M., Lombard, V., Natvig, D.O., Lindquist, E., Schmutz, J., Lucas, S., Harris, P., Powlowski, J., Bellemare, A., Taylor, D., Butler, G., de Vries, R.P., Allijn, I.E., van den Brink, J., Ushinsky, S., Storms, R., Powell, A.J., Paulsen, I.T., Elbourne, L.D.H., Baker, S.E., Magnuson, J., LaBoissiere, S., Clutterbuck, A.J., Martinez, D., Wogulis, M., de Leon, A.L., Rey, M.W., & Tsang, A. (2011). Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nature Biotechnology*, *29*(10), 922–927.

6. Bielen, A.A.M., Verhaart, M.R.A., van der Oost, J., & Kengen, S.W.M. (2013). Biohydrogen Production by the Thermophilic Bacterium *Caldicellulosiruptor saccharolyticus*: Current Status and Perspectives. *Life*, *3*(1), 52–85.

7. Bielen, A.A.M., Verhaart, M.R.A., VanFossen, A.L., Blumer-Schuette, S.E., Stams, A.J.M., van der Oost, J., Kelly, R.M. Kengen, S.W.M. (2013). A thermophile under pressure: Transcriptional analysis of the response of *Caldicellulosiruptor saccharolyticus* to different H_2 partial pressures. *International Journal of Hydrogen Energy*, *38*(4), 1837–1849.

8. Bielen, A.A.M., Willquist, K., Engman, J., van der Oost, J., van Niel, E.W.J., & Kengen, S.W.M. (2010). Pyrophosphate as a central energy carrier in the hydrogen-producing extremely thermophilic *Caldicellulosiruptor saccharolyticus*. *FEMS Microbiology Letters*, *307*(1), 48–54.

9. Bing, W., Wang, H., Zheng, B., Zhang, F., Zhu, G., Feng, Y., & Zhang, Z. (2015). *Caldicellulosiruptor changbaiensis* sp. nov., a cellulolytic and hydrogen-producing bacterium from a hot spring. *International Journal of Systematic and Evolutionary Microbiology*, *65*(1), 293–297.

10. Blumer-Schuette, S.E., Giannone, R.J., Zurawski, J.V., Ozdemir, I., Ma, Q., Yin, Y., Yin, Y., Xu, Y., Kataeva, I., Poole, F.L., Adams, M.W.W., Hamilton-Brehm, S.D., Elkins, J.G., Larimer, F.W., Land, M.L., Hauser, L.J., Cottingham, R.W., Hettich, R.L., & Kelly, R.M. (2012). *Caldicellulosiruptor* Core and Pangenomes Reveal Determinants for Noncellulosomal Thermophilic Deconstruction of Plant Biomass. *Journal of Bacteriology*, *194*(15), 4015–28.

11. Blumer-Schuette, S.E., Kataeva, I., Westpheling, J., Adams, M.W.W., & Kelly, R.M. (2008). Extremely thermophilic microorganisms for biomass conversion: status and prospects. *Current Opinion in Biotechnology*, *19*(3), 210–217.

12. Blumer-Schuette, S.E., Lewis, D.L., & Kelly, R.M. (2010). Phylogenetic, Microbiological, and Glycoside Hydrolase Diversities within the Extremely Thermophilic, Plant Biomass-Degrading Genus *Caldicellulosiruptor*. *Applied and Environmental Microbiology*, *76*(24), 8084–92.

13. Blumer-Schuette, S.E., Ozdemir, I., Mistry, D., Lucas, S., Lapidus, A., Cheng, J.F., Goodwin, L.A., Pitluck, S., Land, M.L., Hauser, L.J., Woyke, T., Mikhailova, N., Pati, A., Kyrpides, N.C., Ivanova, N., Detter, J.C., Walston-Davenport, K., Han, S., Adams, M.W.W., & Kelly, R.M. (2011). Complete genome sequences for the anaerobic, extremely thermophilic plant biomass-degrading bacteria *Caldicellulosiruptor hydrothermalis*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor kronotskyensis*, *Caldicellulosiruptor owensensis*, and *Caldicellulosiruptor lactoaceticus*. *Journal of Bacteriology*, *193*(6), 1483–1484.

14. Braga, C.M.P., Delabona, P.D.S., Lima, D.J.D.S., Paixão, D.A.A., Pradella, J.G.D.C., & Farinas, C.S. (2014). Addition of feruloyl esterase and xylanase produced on-site improves sugarcane bagasse hydrolysis. *Bioresource Technology*, *170*, 316–324.

15. Bratbak, G., & Dundas, I.A.N. (1984). Bacterial Dry Matter Content and Biomass Estimations. *Applied and Environmental Microbiology, 48*(4), 755–757.

16. Bredholt, S., Sonne-Hansen, J., Nielsen, P., Mathrani, I.M., & Ahring, B.K. (1999). *Caldicellulosiruptor kristjanssonii* sp. nov., a cellulolytic, extremely thermophilic, anaerobic bacterium. *International Journal of Systematic Bacteriology*, *49*(3), 991–996.

17. Brynjarsdottir, H., Scully, S. M., & Orlygsson, J. (2013). Production of biohydrogen from sugars and lignocellulosic biomass using *Thermoanaerobacter* GHL15. *International Journal of Hydrogen Energy*, 1–9.

18. Cao, J. (2012). *Biochemical characterization of β-xylan acting glycoside hydrolases from the thermophilic bacterium Caldicellulosiruptor saccharolyticus*. University of Dayton.

19. Chandel, A.K., Giese, E.C., Singh, O.V., & Silva, S.S.D. (2013). Sustainable Role of Thermophiles in the Second Generation of Ethanol Production. *Extremophiles: Sustainable Resources and Biotechnological Implications*, *1*, 267–289.

20. Chandra, R., Takeuchi, H., & Hasegawa, T. (2012). Methane production from lignocellulosic agricultural crop wastes: A review in context to second generation of biofuel production. Renewable and Sustainable Energy Reviews, 16(3), 1462–1476.

21. Chen, L. (2011). *Utilization of Cellulosic Materials by Thermotoga petrophila by*. University of Waterloo.

22. Chomczynski, P., & Sacchi, N. (2006). The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature Protocols*, *1*(2), 581–5.

23. Collins, T., Gerday, C., & Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews*, *29*, 3–23.

24. De Vrije, T., Bakker, R.R., Budde, M.A.W., Lai, M.H., Mars, A.E., & Claassen, P.A.M. (2009). Efficient hydrogen production from the lignocellulosic energy crop *Miscanthus* by the extreme thermophilic bacteria *Caldicellulosiruptor saccharolyticus* and *Thermotoga neapolitana*. *Biotechnology for Biofuels*, *2*(1), 12.

25. De Vrije, T., Budde, M.A.W., Lips, S.J., Bakker, R.R., Mars, A.E., & Claassen, P.A.M. (2010). Hydrogen production from carrot pulp by the extreme thermophiles *Caldicellulosiruptor saccharolyticus* and *Thermotoga neapolitana*. *International Journal of Hydrogen Energy*, *35*(24), 13206–13213.

26. De Vrije, T., Mars, A.E., Budde, M.A.W., Lai, M.H., Dijkema, C., de Waard, P., & Claassen, P.A.M. (2007). Glycolytic pathway and hydrogen yield studies of the extreme thermophile *Caldicellulosiruptor saccharolyticus*. *Applied Microbiology and Biotechnology*, *74*(6), 1358–1367.

27. Durand, H., Soucaille, P., & Tiraby, G. (1984). Comparative study of cellulases and hemicellulases from four fungi: mesophiles *Trichoderma reesei* and *Penicillium* sp. and thermophiles *Thielavia terrestris* and *Sporotrichum cellulophilum*. *Enzyme and Microbial Technology*, *6*, 175–180.

28. Gowen, C.M., & Fong, S.S. (2010). Exploring Biodiversity for Cellulosic Biofuel Production. *Chemistry & Biodiversity*, *7*(5), 1086–1097.

29. Hamilton-Brehm, S.D., Mosher, J.J., Vishnivetskaya, T., Podar, M., Carroll, S., Allman, S., Phelps, T.J., Keller, M., & Elkins, J.G. (2010). *Caldicellulosiruptor obsidiansis* sp. nov., an Anaerobic, Extremely Thermophilic, Cellulolytic Bacterium Isolated from Obsidian Pool, Yellowstone National Park. *Applied and Environmental Microbiology*, *76*(4), 1014–1020.

30. Huang, C., Patel, B.K., Mah, R.A., & Baresi, L. (1998). *Caldicellulosiruptor owensensis* sp. nov., an anaerobic, extremely thermophilic, xylanolytic bacterium. *International Journal of Systematic Bacteriology*, *48*, 91–97.

31. Isern, N.G., Xue, J., Rao, J.V, Cort, J.R., & Ahring, B.K. (2013). Novel monosaccharide fermentation products in *Caldicellulosiruptor saccharolyticus* identified using NMR spectroscopy. *Biotechnology for Biofuels*, *6*(1), 47.

32. Islam, R., Özmihçi, S., Cicek, N., Sparling, R., & Levin, D.B. (2013). Enhanced cellulose fermentation and end-product synthesis by *Clostridium thermocellum* with varied nutrient compositions under carbon-excess conditions. *Biomass and Bioenergy*, *48*, 213–223.

33. Islam, R., Sparling, R., Cicek, N., & Levin, D.B. (2015). Optimization of Influential Nutrients during Direct Cellulose Fermentation into Hydrogen by *Clostridium thermocellum*. *International Journal of Molecular Sciences*, *16*(2), 3116–3132.

34. Ivanova, G., Rákhely, G., & Kovács, K.L. (2009). Thermophilic biohydrogen production from energy plants by *Caldicellulosiruptor saccharolyticus* and comparison with related studies. *International Journal of Hydrogen Energy*, *34*(9), 3659–3670.

35. Kádár, Z., de Vrije, T., van Noorden, G.E., Budde, M.A.W., Szengyel, Z., Réczey, K., & Claassen, P.A.M. (2004). Yields from glucose, xylose, and paper sludge hydrolysate during hydrogen production by the extreme thermophile *Caldicellulosiruptor saccharolyticus*. *Applied Biochemistry and Biotechnology*, *113-116*, 497–508.

36. Kanafusa-Shinkai, S., Wakayama, J., Tsukamoto, K., Hayashi, N., Miyazaki, Y., Ohmori, H., Tajima, K., & Yokoyama, H. (2013). Degradation of microcrystalline cellulose and non-pretreated plant biomass by a cell-free extracellular cellulase/hemicellulase system from the extreme thermophilic bacterium *Caldicellulosiruptor bescii*. *Journal of Bioscience and Bioengineering*, *115*(1), 64–70.

37. Kapdan, I.K., & Kargi, F. (2006). Bio-hydrogen production from waste materials. *Enzyme and Microbial Technology*, *38*(5), 569–582.

38. Kridelbaugh, D.M., Nelson, J., Engle, N.L., Tschaplinski, T.J., & Graham, D.E. (2013). Nitrogen and sulfur requirements for *Clostridium thermocellum* and *Caldicellulosiruptor bescii* on cellulosic substrates in minimal nutrient media. *Bioresource Technology*, *130*, 125–135.

39. 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.

40. Levin, D.B., Carere, C.R., Cicek, N., & Sparling, R. (2009). Challenges for biohydrogen production via direct lignocellulose fermentation. *International Journal of Hydrogen Energy*, *34*(17), 7390–7403.

41. Ling, X. (2012). *Utilization of Cellulosic Materials by Caldicellulosiruptor Species*. University of Waterloo.

42. Lochner, A., Giannone, R.J., Rodriguez, M., Shah, M.B., Mielenz, J.R., Keller, M., Antranikian, G., Graham, D.E., & Hettich, R.L. (2011). Use of label-free quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. *Applied and Environmental Microbiology*, *77*(12), 4042–4054.

43. Lüthi, E., Jasmat, N.B., & Bergquist, P.L. (1990). Xylanase from the extremely thermophilic bacterium "*Caldocellum saccharolyticum*": overexpression of the gene in *Escherichia coli* and characterization of the gene product. *Applied and Environmental Microbiology*, *56*(9), 2677–2683.

44. Lynd, L.R., Weimer, P.J., Zyl, W.H. Van, & Isak, S. (2002). Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*, *66*(3), 506–577.

45. Martinez, D., Larrondo, L.F., Putnam, N., Gelpke, M.D.S., Huang, K., Chapman, J., Helfenbein, K.G., Ramaiya, P., Detter, J.C., Larimer, F., Coutinho, P.M., Henrissat, B., Berka, R., Cullen, D., & Rokhsar, D. (2004). Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnology*, *22*(6), 695–700.

46. Martinez-Porqueras, E., Wechselberger, P., & Herwig, C. (2013). Effect of medium composition on biohydrogen production by the extreme thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *International Journal of Hydrogen Energy*, *38*(27), 11756–11764.

47. Miroshnichenko, M.L., Kublanov, I.V., Kostrikina, N.A., Tourova, T.P., Kolganova, T.V., Birkeland, N.K., & Bonch-Osmolovskaya, E.A. (2008). *Caldicellulosiruptor kronotskyensis* sp. nov. and *Caldicellulosiruptor hydrothermalis* sp. nov., two extremely thermophilic, cellulolytic, anaerobic bacteria from Kamchatka thermal springs. *International Journal of Systematic and Evolutionary Microbiology*, *58*(Pt 6), 1492–1496.

48. Mladenovska, Z., Mathrani, I.M., & Ahring, B.K. (1995). Isolation and characterization of *Caldicellulosiruptor lactoaceticus* sp. nov., an extremely thermophilic, cellulolytic, anaerobic bacterium. *Archives of Microbiology*, *163*(3), 223–230.

49. Onyenwoke, R.U., Lee, Y.J., Dabrowski, S., Ahring, B.K., & Wiegel, J. (2006). Reclassification of *Thermoanaerobium acetigenum* as *Caldicellulosiruptor acetigenus* comb. nov. and emendation of the genus description. *International Journal of Systematic and Evolutionary Microbiology*, *56*(Pt 6), 1391–1395.

50. Ozioko, P.C., Ikeyi, A.P., & Ugwu, O.P.C. (2013). Cellulases, their substrates, activity and assay methods. *International Journal of Science and Technology*, *12*(2), 778–785.

51. Panagiotopoulos, I.A., Bakker, R.R., de Vrije, T., Claassen, P.A.M., & Koukios, E.G. (2012). Dilute-acid pretreatment of barley straw for biological hydrogen production using *Caldicellulosiruptor saccharolyticus*. *International Journal of Hydrogen Energy*, *37*(16), 11727–11734.

52. Panagiotopoulos, I.A., Bakker, R. R., de Vrije, T., Koukios, E.G., & Claassen, P.A.M. (2010). Pretreatment of sweet sorghum bagasse for hydrogen production by *Caldicellulosiruptor saccharolyticus*. *International Journal of Hydrogen Energy*, *35*(15), 7738–7747.

53. Panagiotopoulos, J.A., Bakker, R.R., de Vrije, T., Urbaniec, K., Koukios, E.G., & Claassen, P.A.M. (2010). Prospects of utilization of sugar beet carbohydrates for biological hydrogen production in the EU. *Journal of Cleaner Production*, *18*, S9–S14.

54. Park, J.I., Kent, M.S., Datta, S., Holmes, B.M., Huang, Z., Simmons, B.A., Sale, K.L., & Sapra, R. (2011). Enzymatic hydrolysis of cellulose by the cellobiohydrolase domain of CelB from the hyperthermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Bioresource Technology*, *102*(10), 5988–94.

55. Pawar, S.S., Vongkumpeang, T., Grey, C., & van Niel, E.W. (2015). Biofilm formation by designed co-cultures of *Caldicellulosiruptor* species as a means to improve hydrogen productivity. *Biotechnology for Biofuels*, *8*(19), 1-12.

56. Rainey, F.A., Donnison, A.M., Janssen, P.H., Saul, D., Rodrigo, A., Bergquist, P.L., Daniel, R.M., Stakebrandt, E., & Morgan, H.W. (1994). Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: An obligately anaerobic, extremely thermophilic, cellulolytic bacterium. *FEMS Microbiology Letters*, *120*, 263–266.

57. Ren, N., Guo, W., Liu, B., Cao, G., & Ding, J. (2011). Biological hydrogen production by dark fermentation: challenges and prospects towards scaled-up production. *Current Opinion in Biotechnology*, *22*(3), 365–70.

58. Rittmann, S., & Herwig, C. (2012). A comprehensive and quantitative review of dark fermentative biohydrogen production. *Microbial Cell Factories*, *11*(1), 115.

59. Rothstein, J. (1993). Hydrogen and fossil fuels. *International Journal of Hydrogen Energy*, *20*(4), 283–286.

60. Rubin, E. M. (2008). Genomics of cellulosic biofuels. *Nature*, *454*(7206), 841–845.

61. Schwarz, W. H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology*, *56*(5-6), 634–649.

62. Selig, M., Xavier, K.B., Santos, H., & Schonheit, P. (1997). Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium *Thermotoga*. *Archives of Microbiology*, *167*(4), 217–232.

63. Shen, N., Zhang, F., Song, X.N., Wang, Y.S., & Zeng, R.J. (2013). Why is the ratio of H_2/acetate over 2 in glucose fermentation by *Caldicellulosiruptor saccharolyticus*? *International Journal of Hydrogen Energy*, *38*(26), 11241–11247.

64. Sissons, C.H., Sharrock, K.R., Daniel, R.M., & Morgan, H.W. (1987). Isolation of cellulolytic anaerobic extreme thermophiles from New Zealand thermal sites. *Applied and Environmental Microbiology*, *53*(4), 832–838.

65. Sizova, M.V, Izquierdo, J.A, Panikov, N.S., & Lynd, L.R. (2011). Cellulose- and xylan-degrading thermophilic anaerobic bacteria from biocompost. *Applied and Environmental Microbiology*, *77*(7), 2282–2291.

66. Sleat, R., Mah, R.A., & Robinson, R. (1984). Isolation and Characterization of an Anaerobic, Cellulolytic Bacterium, *Clostridium cellulovorans* sp. nov. *Applied and Environmental Microbiology*, *48*(1), 88–93.

67. Sorek, R., & Cossart, P. (2010). Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nature Reviews. Genetics, 11(1), 9–16.

68. Stafford, R. (1995). *Advances in Microbial Physiology* (37th ed.). San Diego: Academic Press Inc. pp. 1-376.

69. Svetlitchnyi, V.A., Kensch, O., Falkenhan, D.A., Korseska, S.G., Lippert, N., Prinz, M., Sassi, J., Schickor, A., & Curvers, S. (2013). Single-step ethanol production from lignocellulose using novel extremely thermophilic bacteria. *Biotechnology for Biofuels*, *6*(1), 31.

70. Talluri, S., Raj, S.M., & Christopher, L.P. (2013). Consolidated bioprocessing of untreated switchgrass to hydrogen by the extreme thermophile *Caldicellulosiruptor saccharolyticus* DSM 8903. *Bioresource Technology*, *139*, 272–279.

71. Van de Werken, H.J.G., Verhaart, M.R.A., VanFossen, A.L., Willquist, K., Lewis, D.L., Nichols, J.D., Goorissen, E.F., Mongodin, K.E.N., van Niel, E.W.J., Stams, A.J.M., Ward, D.E., Willem, M.D.V., van der Oost, J., Kelly, R.M., & Kengen, S. W. M. (2008). Hydrogenomics of the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Applied and Environmental Microbiology*, *74*(21), 6720–6729.

72. Van Niel, E.W.J., Budde, M.A.W., Haas, G.G.D., Van Der Wal, F.J., Claassen, P.A.M., & Stams, A.J.M. (2002). Distinctive properties of high hydrogen producing extreme thermophiles, *Caldicellulosiruptor saccharolyticus* and *Thermotoga elfii*. *International Journal of Hydrogen Energy*, *27*, 1391–1398.

73. Van Niel, E.W.J., Claassen, P.A.M., & Stams, A.J.M. (2003). Substrate and product inhibition of hydrogen production by the extreme thermophile, *Caldicellulosiruptor saccharolyticus*. *Biotechnology and Bioengineering*, *81*(3), 255–262.

74. VanFossen, A.L., Ozdemir, I., Zelin, S.L., & Kelly, R.M. (2011). Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Biotechnology and Bioengineering*, *108*(7), 1559–1569.

75. Vanfossen, A.L., Verhaart, M.R.A., Kengen, S.M.W., & Kelly, R.M. (2009). Carbohydrate utilization patterns for the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus* reveal broad growth substrate preferences. *Applied and Environmental Microbiology*, *75*(24), 7718–7724.

76. Verhaart, M.R.A. (2010). *Genomics and transcriptomics of the hydrogen producing extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus*. Wageningen University.

77. Vorholt, J., Kunow, J., Stetter, K.O., & Thauer, R.K. (1995). Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO_2 fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Archives of Microbiology*, *163*(2), 112–118.

78. Wagner, I.D., & Wiegel, J. (2008). Diversity of thermophilic anaerobes. *Annals of the New York Academy of Sciences*, *1125*, 1–43.

79. Warnick, T.A., Methé, B.A., & Leschine, S.B. (2002). *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. *International Journal of Systematic and Evolutionary Microbiology*, *52*, 1155–1160.

80. Willquist, K., Pawar, S.S., & Van Niel, E.W.J. (2011). Reassessment of hydrogen tolerance in *Caldicellulosiruptor saccharolyticus*. *Microbial Cell Factories*, *10*, 111–122.

81. Willquist, K., & van Niel, E.W.J. (2010). Lactate formation in *Caldicellulosiruptor saccharolyticus* is regulated by the energy carriers pyrophosphate and ATP. *Metabolic Engineering*, *12*(3), 282–290.

82. Willquist, K., & van Niel, E.W.J. (2012). Growth and hydrogen production characteristics of *Caldicellulosiruptor saccharolyticus* on chemically defined minimal media. *International Journal of Hydrogen Energy*, *37*(6), 4925–4929.

83. Willquist, K., Zeidan, A.A., & van Niel, E.W.J. (2010). Physiological characteristics of the extreme thermophile *Caldicellulosiruptor saccharolyticus*: an efficient hydrogen cell factory. *Microbial Cell Factories*, *9*(1), 89–106.

84. Yamamoto, T., Nishio-Kosaka, M., Izawa, S., Aga, H., Nishimoto, T., Chaen, H., & Fukuda, S. (2011). Enzymatic Properties of Recombinant Kojibiose Phosphorylase from *Caldicellulosiruptor saccharolyticus* ATCC43494. *Bioscience, Biotechnology, and Biochemistry*, *75*(6), 1208–1210.

85. Yang, S.J., Kataeva, I., Hamilton-brehm, S.D., Engle, N.L., Tschaplinski, T.J., Doeppke, C., Davis, M., Westpheling, J., & Adams, M.W.W. (2009). Efficient Degradation of Lignocellulosic Plant Biomass , without Pretreatment , by the Thermophilic Anaerobe *Anaerocellum thermophilum* DSM 6725. *Applied and Environmental Microbiology, 75*(14), 4762-4769.

86. Yang, S.J., Kataeva, I., Wiegel, J., Yin, Y., Dam, P., Xu, Y., Westpheling, J., & Adams, M.W.W. (2010). Classification of "*Anaerocellum thermophilum*" strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *60*(9), 2011–2015.

87. Zeidan, A.A., Rådström, P., & van Niel, E.W.J. (2010). Stable coexistence of two *Caldicellulosiruptor* species in a de novo constructed hydrogen-producing co-culture. *Microbial Cell Factories*, *9*(1), 102–115.

88. Zhang, F., Rodriguez, S., & Keasling, J.D. (2011). Metabolic engineering of microbial pathways for advanced biofuels production. *Current Opinion in Biotechnology*, *22*(6), 775–83.

89. Zverlov, V., Mahr, S., Riedel, K., & Bronnenmeier, K. (1998). Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile "*Anaerocellum thermophilum*" with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology*, *144*, 457–465.

Appendix A. Growth Media Composition

Modified DSMZ 640 – *Caldicellulosiruptor* Medium

Deionized water	1000.00 mL
NH_4Cl	0.90 g
NaCl	0.90 g
$MgCl_2 \times 6H_2O$	0.40 g
KH_2PO_4	0.75 g
K_2HPO_4	1.50 g
Trace element solution SL-10	1.00 mL
$FeCl_3 \times 6H_2O$	2.50 mg
Resazurin	0.50 mg
Vitamin solution *	10.00 mL
Cysteine-HCl × H_2O **	0.75 g
NaOH **	0.40 g

*Vitamin solution was added to the media to supplement the missing nutrients from removal of trypticase and yeast extract. The vitamin solution was added after autoclaving to prevent vitamins from being degraded by heat.
**Cysteine-HCl (about 0.25 mL of a 15% solution) was added after autoclaving and degassing to reduce the media and the NaOH (about 0.5 mL of 1M solution) was added after that to adjust the pH (ColorpHast pH Strips, EMD Millipore, USA) to 7.0.

Dissolve the components in the order listed in approximately 990 mL of DI water. Adjust pH (Accumet AB15 pH meter, Fisher Scientific, USA) of the completed medium to 7.2 using NaOH and top up to 1000 mL. Sterilize the media in the autoclave at 121°C for 30 minutes using a liquid cycle. Allow the media to cool to room temperature and then degas and flush the media using 100% N_2 gas.

Note: Trypticase and yeast extract were omitted from the original recipe since these provided complex sources of nutrients. Additionally, the cellobiose from the original recipe was replaced with the desired substrates. Soluble substrates were added after autoclaving from anoxic stock solutions prepared under N_2 gas atmosphere and sterilized by filtration. Insoluble substrates were added to the media prior to autoclaving.
Media recipe available from the DSMZ website:
(http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium640.pdf).

Trace element solution SL-10:

HCl (25%; 7.7 M)	10.00 mL
$FeSO_4$ *	1.15 g
$ZnCl_2$	70.00 mg
$MnCl_2 \times 4\ H_2O$	100.00 mg
H_3BO_3	6.00 mg
$CoCl_2 \times 6\ H_2O$	190.00 mg
$CuSO_4 \times 5\ H_2O$ *	3.00 mg
$NiCl_2 \times 6\ H_2O$	24.00 mg
$Na_2MoO_4 \times 2\ H_2O$	36.00 mg
Distilled water	990.00 mL

*$FeSO_4$ was used here to replace the $FeCl_2 \times 4\ H_2O$ and $CuSO_4 \times 5\ H_2O$ was used to replace $CuCl_2 \times 2\ H_2O$ in the original recipe due to availability in the lab.

First dissolve $FeSO_4$ in the HCl, then dilute in approximately 900 mL of DI water. Add and dissolve the other salts and then top up to 1000 mL.

Note: The original composition for Trace element solution SL-10 is available on the DSMZ website under "320. *Clostridium cellulovorans medium.*"
(http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium320.pdf)

Vitamin solution:

Biotin	2.00 mg
Folic acid	2.00 mg
Pyridoxine-HCl	10.00 mg
Thiamine-HCl $\times 2\ H_2O$	5.00 mg
Riboflavin	5.00 mg
Nicotinic acid	5.00 mg
D-Ca-pantothenate	5.00 mg
Vitamin B12	0.10 mg
p-Aminobenzoic acid	5.00 mg
Lipoic acid	5.00 mg
Distilled water	1000.00 mL

Dissolve all of the vitamins in the order listed and filter sterilize the solution using a 0.2 µm pore size filter.
Note: The Vitamin Solution was adapted from "141. Methanogenium Medium" from the DSMZ website and is the same vitamin mixture as used in the DSMZ 671 medium below.
(http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium141.pdf)

Modified DSMZ 671 – Modified BA Medium

Distilled water	1000.00 mL
NH_4Cl	1.00 g
NaCl	0.10 g
$MgCl_2 \times 6\ H_2O$	0.10 g
$CaCl_2 \times 2\ H_2O$	0.05 g
$K_2HPO_4 \times 3\ H_2O$	0.40 g
Resazurin	0.50 mg
Trace element solution (141)	10.00 mL
$NaHCO_3$	2.60 g
Vitamin solution *	10.00 mL
$Na_2S \times 9\ H_2O$ **	0.25 g
HCl **	0.29 g

*Vitamin solution was added to the media to supplement the missing nutrients from removal of yeast extract. The vitamin solution was added after autoclaving to prevent vitamins from being degraded by heat.
** $Na_2S \times 9\ H_2O$ (about 0.4 mL of a 3% solution) was added after autoclaving and degassing to reduce the media and the HCl (about 0.4 mL of a 1M solution) was added after that to adjust the pH to 7.0 (ColorpHast pH Strips, EMD Millipore, USA).

Dissolve the components in the order listed in approximately 900 mL of DI water. Adjust pH (Accumet AB15 pH meter, Fisher Scientific, USA) of the completed medium to 6.95 using NaOH and top up to 1000 mL. Sterilize the media in the autoclave at 121°C for 30 minutes using a liquid cycle. Allow the media to cool to room temperature and then degas and flush the media using 100% N_2 gas.

Note: Yeast extract was omitted from the original recipe since these provided complex sources of nutrients. Additionally, the cellobiose or cellulose from the original recipe was replaced with the desired substrates. Soluble substrates were added after autoclaving from anoxic stock solutions prepared under N_2 gas atmosphere and sterilized by filtration. Insoluble substrates were added to the media prior to autoclaving.
Media recipe available from the DSMZ website:
(http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium671.pdf).

Trace element solution (141):

Nitrilotriacetic acid	1.50 g
$MgSO_4 \times 7\ H_2O$	3.00 g
$MnSO_4 \times H_2O$	0.50 g
NaCl	1.00 g
$FeSO_4 \times 7\ H_2O$	0.10 g
$CoSO_4 \times 7\ H_2O$	0.18 g
$CaCl_2 \times 2\ H_2O$	0.10 g
$ZnSO_4 \times 7\ H_2O$	0.18 g
$CuSO_4 \times 5\ H_2O$	0.01 g
$KAl(SO_4)_2 \times 12\ H_2O$	0.02 g
H_3BO_3	0.01 g
$Na_2MoO_4 \times 2\ H_2O$	0.01 g
$NiCl_2 \times 6\ H_2O$	0.03 g
$Na_2SeO_3 \times 5\ H_2O$	0.30 mg
Distilled water	1000.00 mL

First dissolve the nitrilotriacetic acid in about 990 mL of DI water and adjust the pH (Accumet AB15 pH meter, Fisher Scientific, USA) to 6.5 with NaOH. Add the other minerals according to the list and adjust the final pH (Accumet AB15 pH meter, Fisher Scientific, USA) to 7.0 with NaOH. Finally, top up to 1000mL using DI water.
Note: The Trace element solution was adapted from "141. *Methanogenium* Medium" available on the DSMZ website.
(http://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium141.pdf)

Vitamin solution:

Biotin	2.00 mg
Folic acid	2.00 mg
Pyridoxine-HCl	10.00 mg
Thiamine-HCl $\times 2\ H_2O$	5.00 mg
Riboflavin	5.00 mg
Nicotinic acid	5.00 mg
D-Ca-pantothenate	5.00 mg
Vitamin B_{12}	0.10 mg
p-Aminobenzoic acid	5.00 mg
Lipoic acid	5.00 mg
Distilled water	1000.00 mL

Dissolve all of the vitamins in the order listed and filter sterilize the solution using a 0.2 μm pore size filter.
Note: The Vitamin Solution was adapted from "141. *Methanogenium* Medium" from the DSMZ website and is the same vitamin mixture as used in the DSMZ 640 medium above.

Substrates added to Growth Media

Soluble
Glucose (Dextrose, Anhydrous, GR ACS, from EMD Millipore)
Xylose (D-(+)-Xylose, 98+%, from Alfa Aesar)
Cellobiose (D-(+)-Cellobiose, ≥98%, from Sigma-Aldrich)
Yeast Extract (Yeast Extract, Granulated, from EMD Millipore)

Soluble substrates were added at a concentration of 2 g/L.

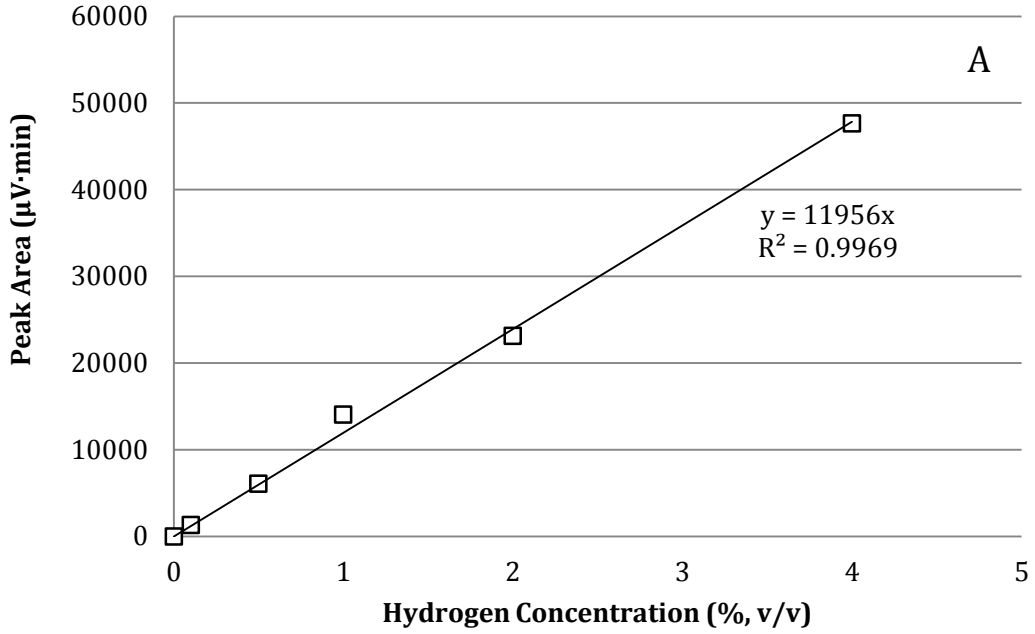
Insoluble
Avicel (Avicel PH 102, microcrystalline cellulose, NF, Ph. Eur., JP, BP, from FMC BioPolymer)
CMC (Carboxymethylcellulose sodium salt, low viscosity, from Sigma-Aldrich)
Switchgrass (Switchgrass, milled to 0.25 mm particle size using a UDY Cyclone sample milling machine (UDY 3010-080P, USA) and washed by stirring 20 g per 1L DI water incubated at 75°C for 24 hours and dried in a 50°C incubator overnight, produced and harvested from Nott Farms, Clinton, Ontario)
Xylan (Xylan, from beechwood, ≥90% (HPLC), cell wall polysaccharide, from Sigma-Aldrich)

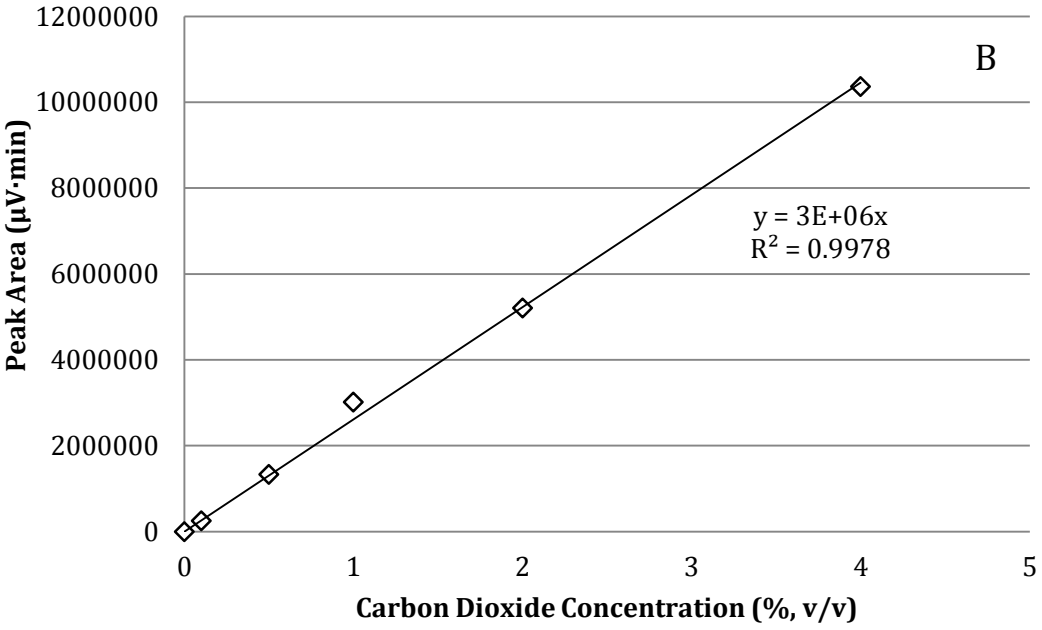
Insoluble substrates were added at a concentration of 4 g/L.

Note: Due to poor growth and long lag phases, insoluble substrates including Avicel, CMC, and switchgrass were supplemented with 0.1 g/L yeast extract. Xylan grew fine without the addition of yeast extract.

Appendix B. Standard Curves for GC and HPLC

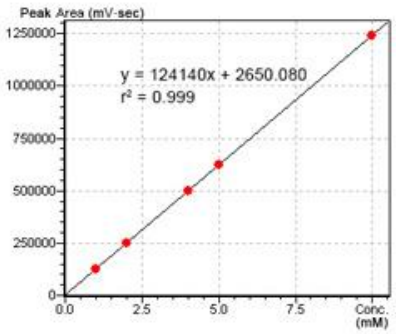
The following graphs display the standard curves used for calculating the hydrogen and carbon dioxide concentrations obtained from gas chromatography (GC).



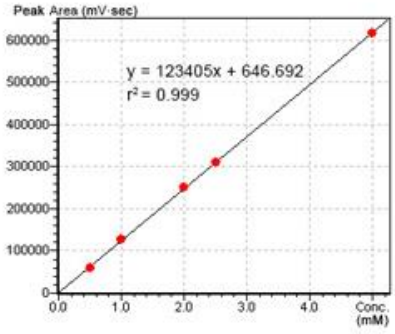


The following graphs display the standard curves used for calculating the cellobiose, citric acid, glucose, xylose, arabinose, succinate, lactate, acetate, and ethanol concentrations obtained from high performance liquid chromatography (HPLC).

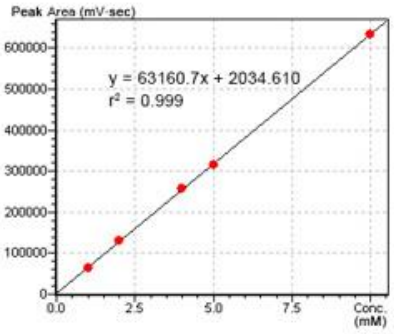
C



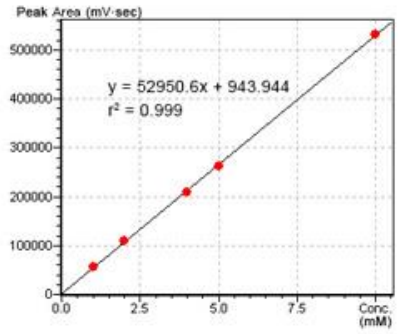
Cellobiose



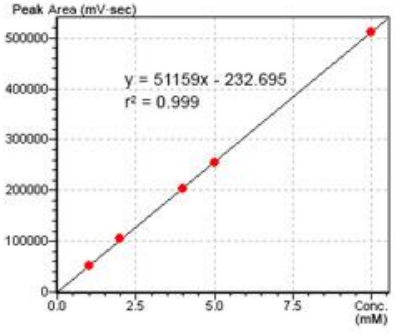
Citric Acid



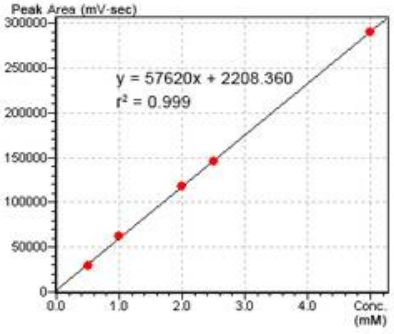
Glucose



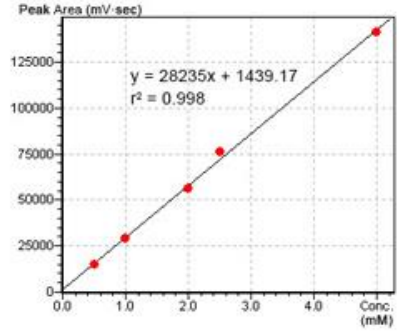
Xylose



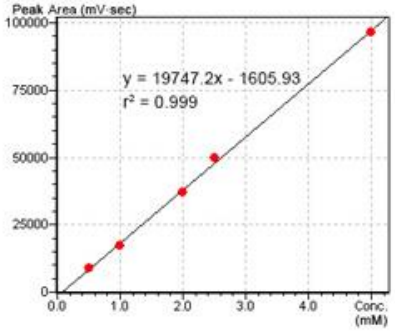
Arabinose



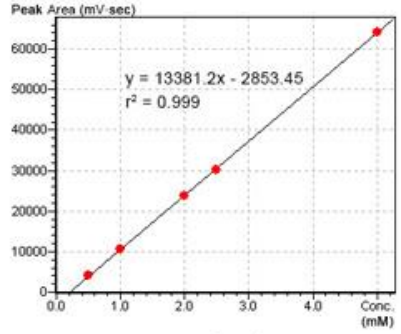
Succinate



Lactate



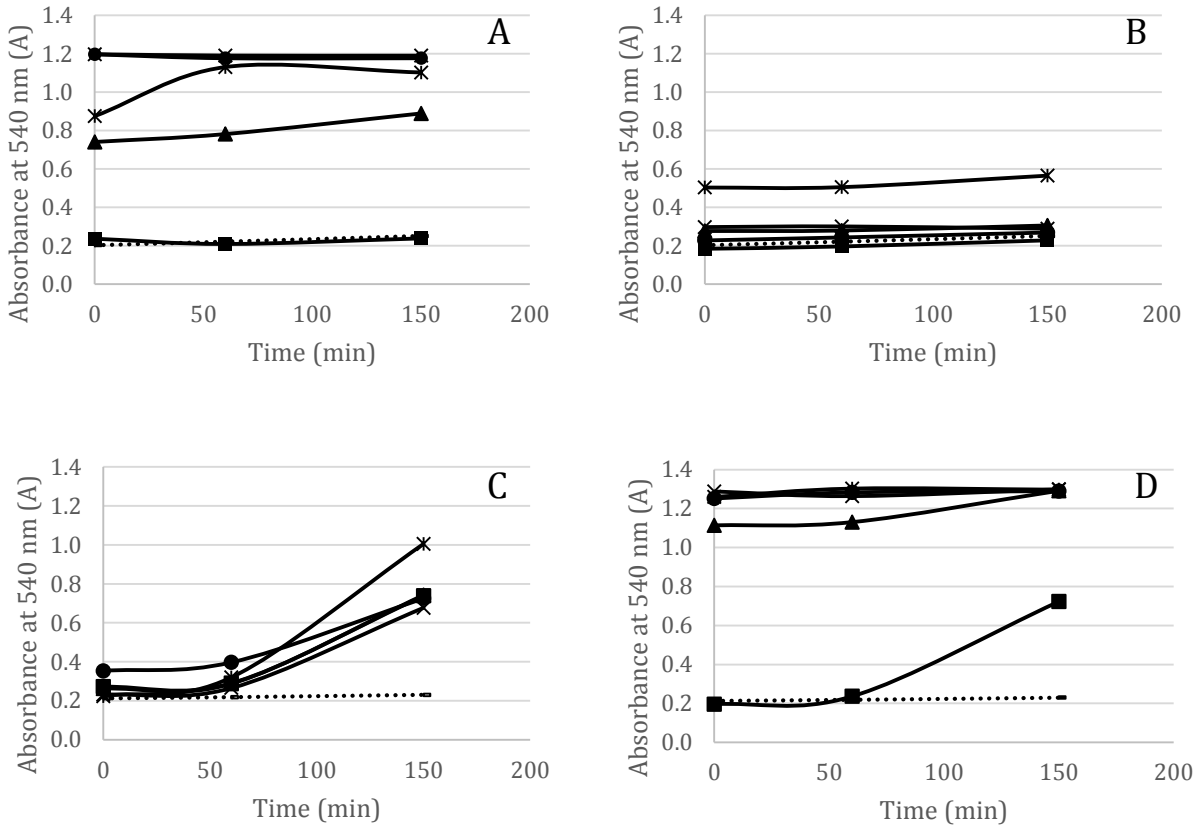
Acetate



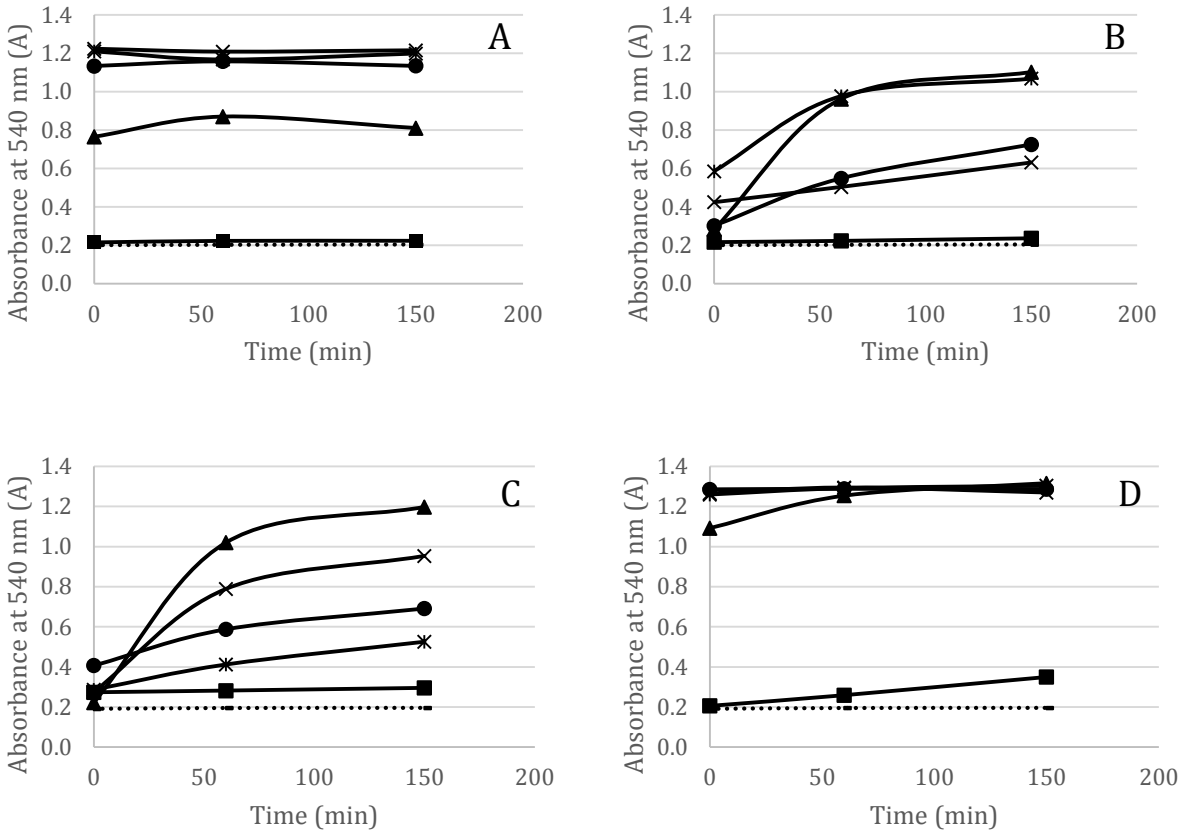
Ethanol

Appendix C. Kinetic Curves for Cellulase Enzyme Assays

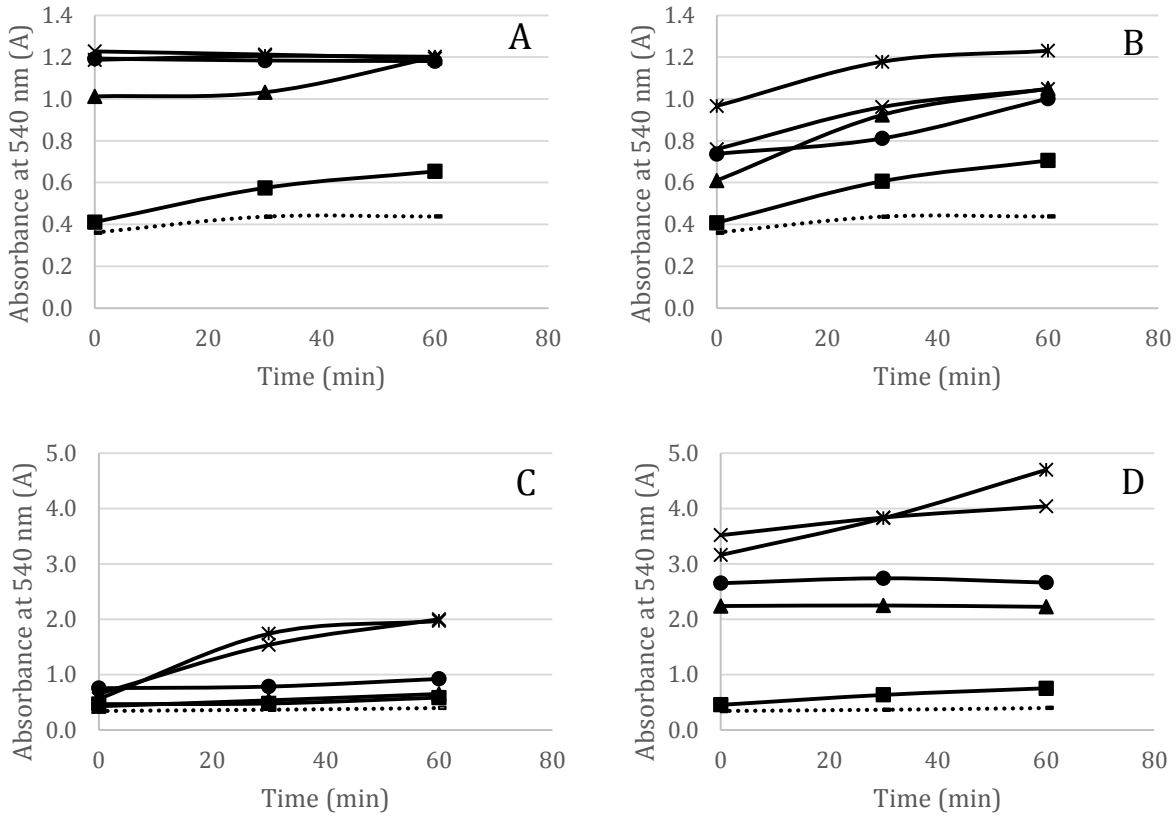
The following graphs show the kinetic curves recorded while performing the enzyme assays. The total activity and specific activity of each enzyme were calculated using these results.



Exoglucanase activity for *Caldicellulosiruptor* spp. grown on five substrates. These graphs show the results of enzyme assays performed on (A) *C. saccharolyticus* supernatant, (B) *C. saccharolyticus* CFE, (C) *C. kristjanssonii* supernatant, and (D) *C. kristjanssonii* CFE, each collected from respective cultures grown on glucose (●), xylose (✕), Avicel (■), CMC (▲), or xylan (✳), and a negative control (- - -). Note that activity observed in the supernatant represents enzymes that are secreted extracellularly while activity observed from the CFE represents intracellular enzymes.



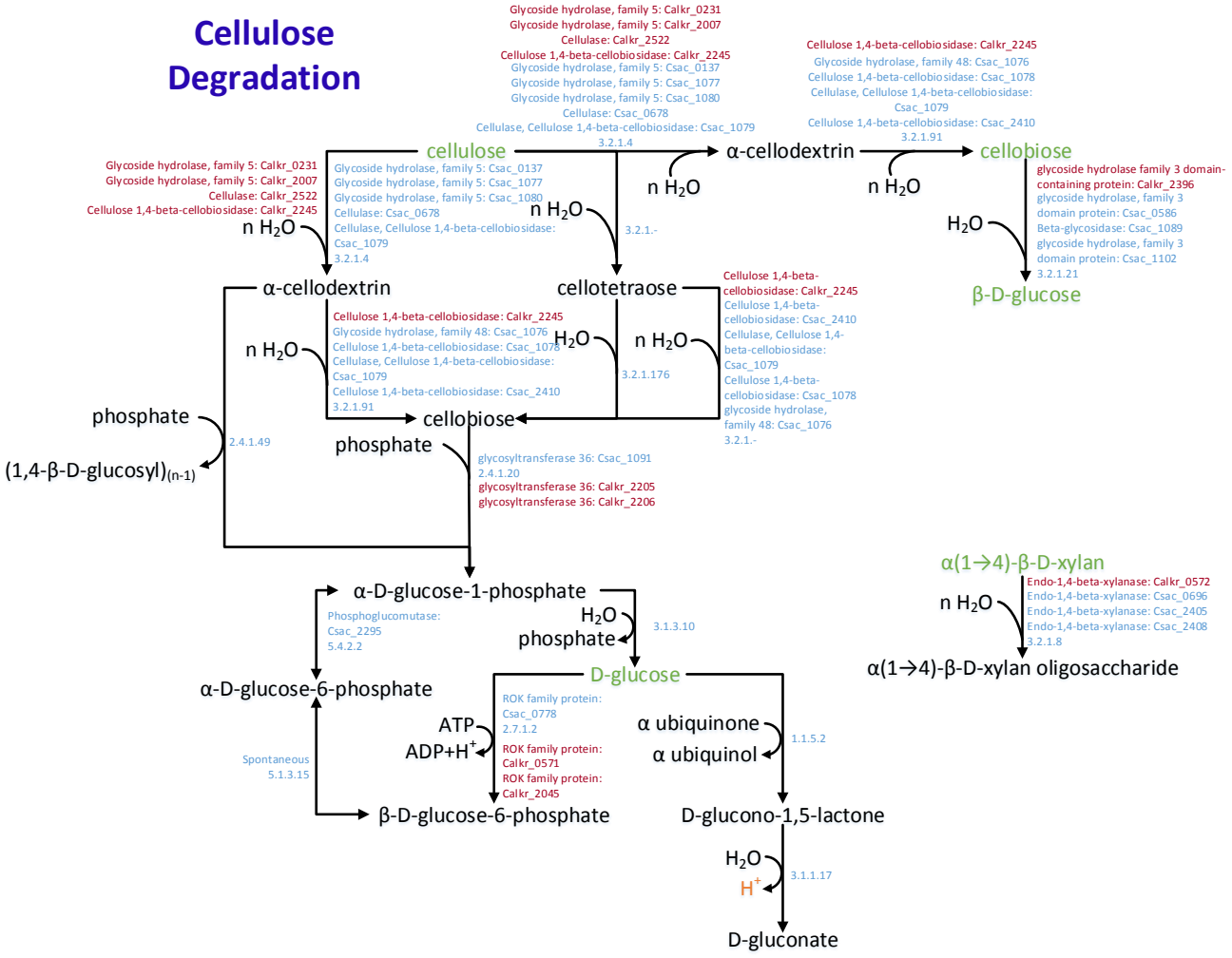
Endoglucanase activity for *Caldicellulosiruptor* spp. grown on five substrates. These graphs show the results of enzyme assays performed on (A) *C. saccharolyticus* supernatant, (B) *C. saccharolyticus* CFE, (C) *C. kristjanssonii* supernatant, and (D) *C. kristjanssonii* CFE, each collected from respective cultures grown on glucose (●), xylose (×), Avicel (■), CMC (▲), or xylan (✳), and a negative control (- - -). Note that activity observed in the supernatant represents enzymes that are secreted extracellularly while activity observed from the CFE represents intracellular enzymes.

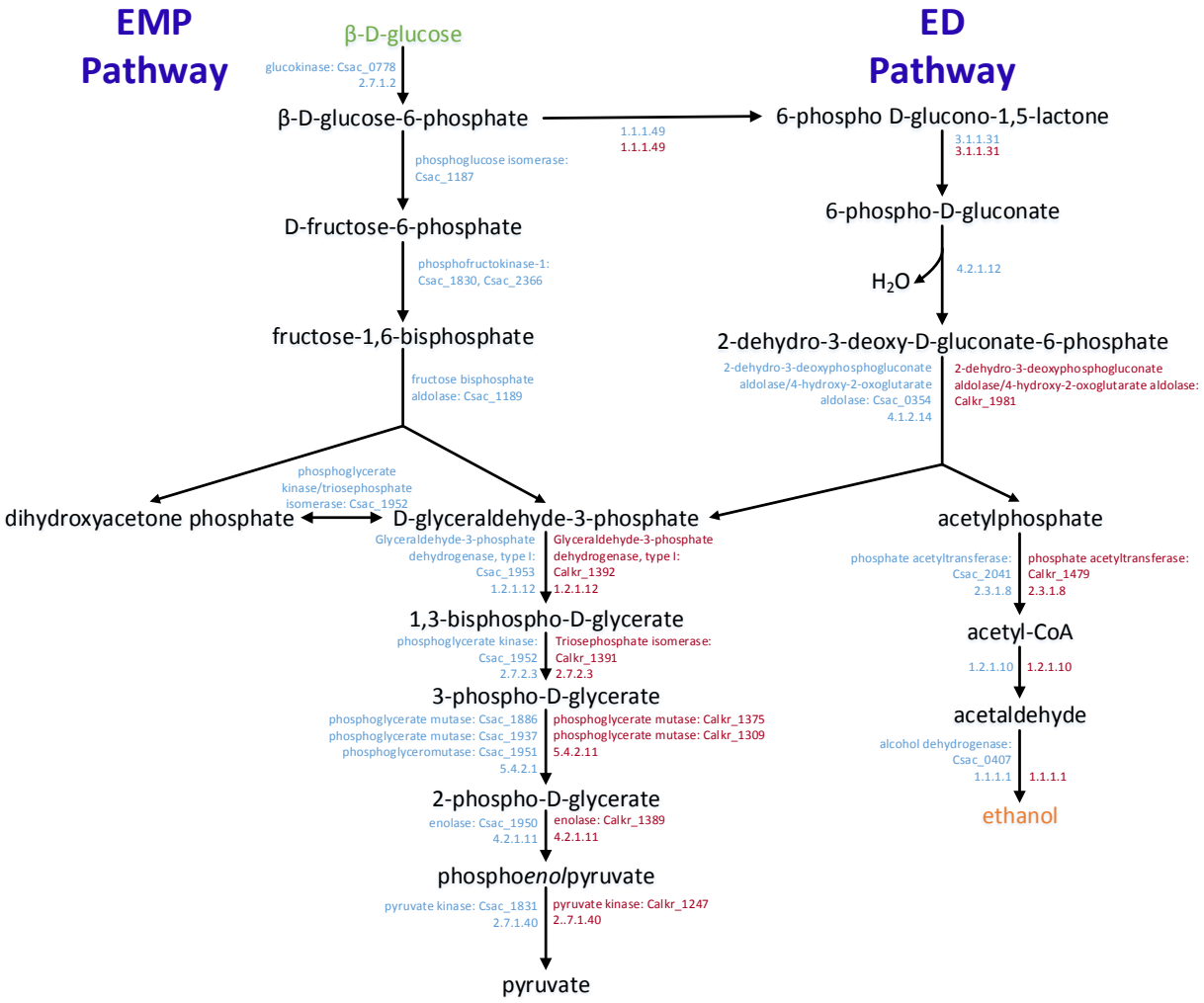


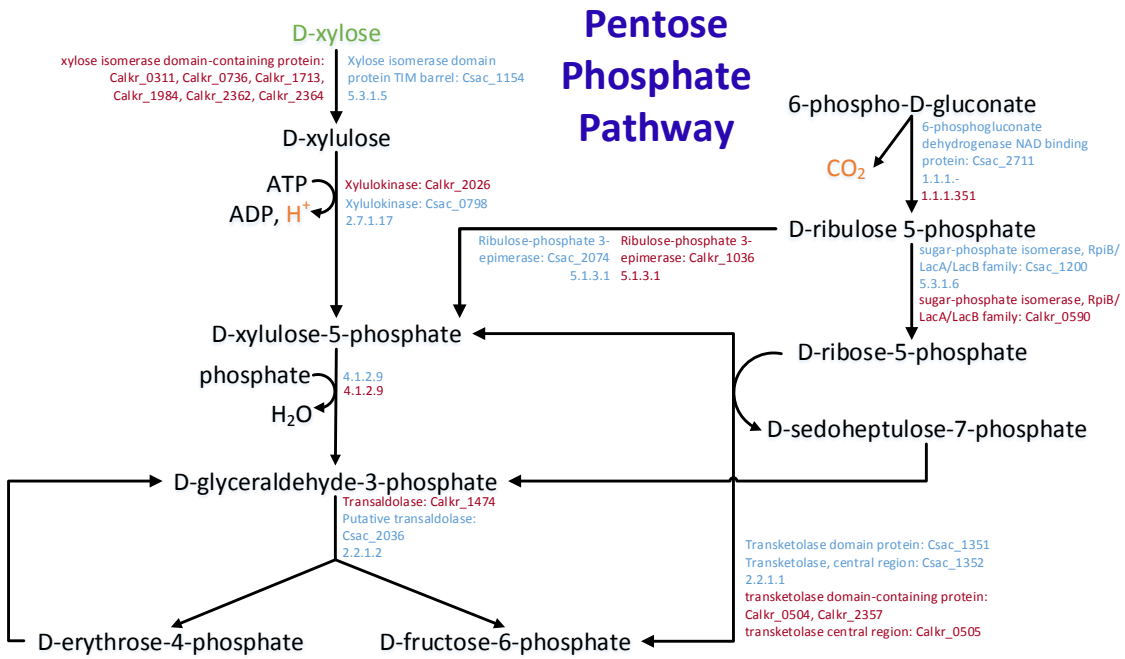
Xylanase activity for *Caldicellulosiruptor* spp. grown on five substrates.
These graphs show the results of enzyme assays performed on (A) *C. saccharolyticus* supernatant, (B) *C. saccharolyticus* CFE, (C) *C. kristjanssonii* supernatant, and (D) *C. kristjanssonii* CFE, each collected from respective cultures grown on glucose (●), xylose (✱), Avicel (■), CMC (▲), or xylan (✱), and a negative control (- - -). Note that activity observed in the supernatant represents enzymes that are secreted extracellularly while activity observed from the CFE represents intracellular enzymes.

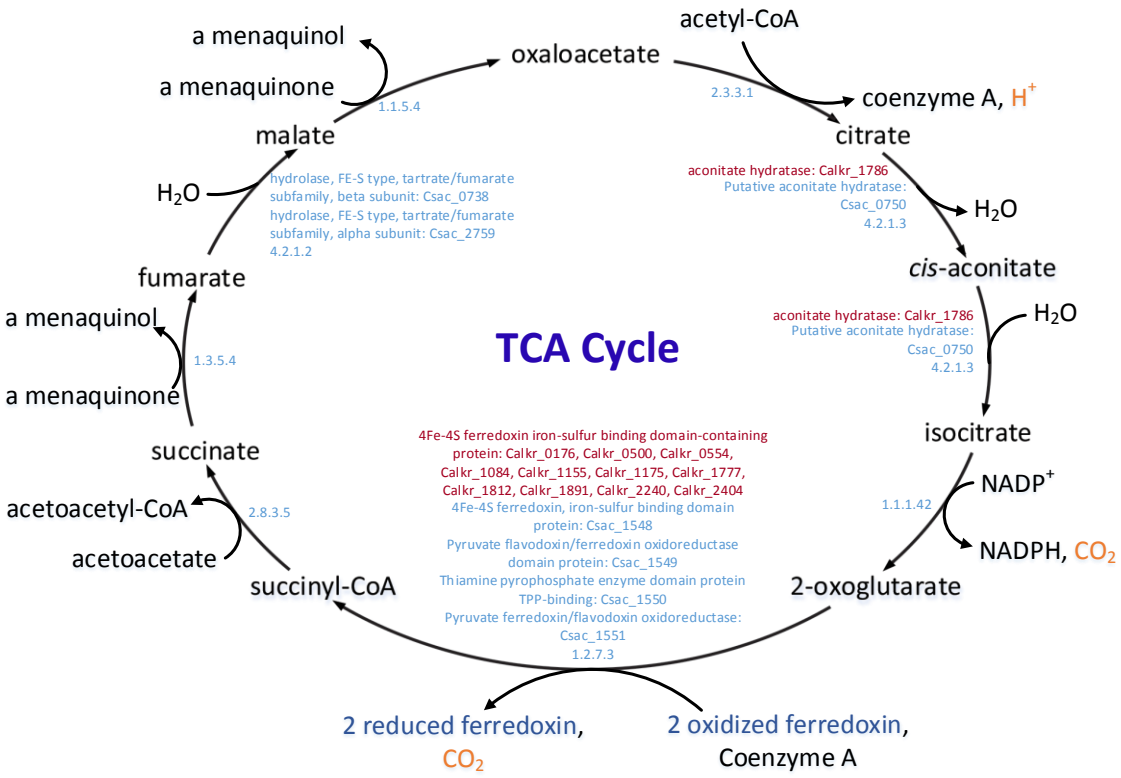
Appendix D. Biochemical Pathways in Metabolic Profile

Expanding from Figure 17, this is a closer look at the pathways involved in cellulose degradation, carbohydrate metabolism, and fermentation from *C. saccharolyticus* and *C. kristjanssonii*.

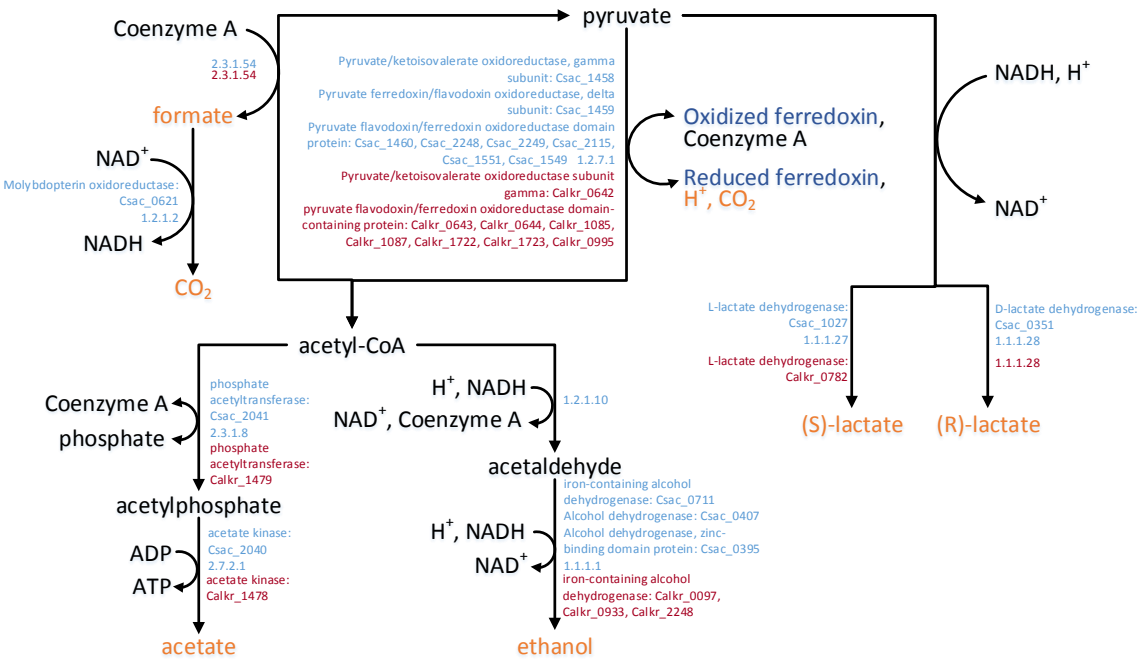



EMP Pathway
ED Pathway
β-D-glucose
glucokinase: Csac_0778
2.7.1.2
β-D-glucose-6-phosphate
1.1.1.49
1.1.1.49
6-phospho D-glucono-1,5-lactone
3.1.1.31
3.1.1.31
phosphoglucose isomerase: Csac_1187
D-fructose-6-phosphate
6-phospho-D-gluconate
phosphofructokinase-1: Csac_1830, Csac_2366
H2O
4.2.1.12
fructose-1,6-bisphosphate
2-dehydro-3-deoxy-D-gluconate-6-phosphate
2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase: Csac_0354
4.1.2.14
2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase: Calkr_1981
fructose bisphosphate aldolase: Csac_1189
phosphoglycerate kinase/triosephosphate isomerase: Csac_1952
dihydroxyacetone phosphate
D-glyceraldehyde-3-phosphate
acetylphosphate
Glyceraldehyde-3-phosphate dehydrogenase, type I: Csac_1953
1.2.1.12
Glyceraldehyde-3-phosphate dehydrogenase, type I: Calkr_1392
1.2.1.12
phosphate acetyltransferase: Csac_2041
2.3.1.8
phosphate acetyltransferase: Calkr_1479
2.3.1.8
1,3-bisphospho-D-glycerate
acetyl-CoA
phosphoglycerate kinase: Csac_1952
2.7.2.3
Triosephosphate isomerase: Calkr_1391
2.7.2.3
1.2.1.10
1.2.1.10
3-phospho-D-glycerate
acetaldehyde
phosphoglycerate mutase: Csac_1886
phosphoglycerate mutase: Csac_1937
phosphoglyceromutase: Csac_1951
5.4.2.1
phosphoglycerate mutase: Calkr_1375
phosphoglycerate mutase: Calkr_1309
5.4.2.11
alcohol dehydrogenase: Csac_0407
1.1.1.1
1.1.1.1
2-phospho-D-glycerate
ethanol
enolase: Csac_1950
4.2.1.11
enolase: Calkr_1389
4.2.1.11
phosphoenolpyruvate
pyruvate kinase: Csac_1831
2.7.1.40
pyruvate kinase: Calkr_1247
2..7.1.40
pyruvate

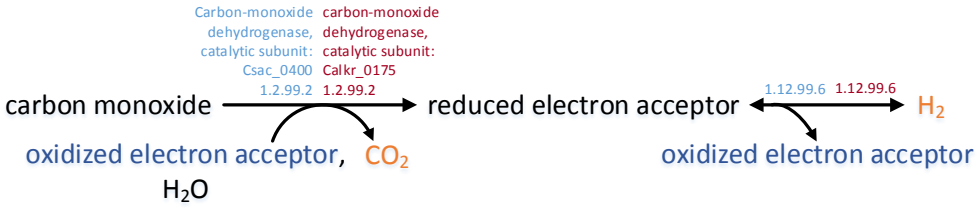

Pentose Phosphate Pathway
D-xylose
xylose isomerase domain-containing protein: Calkr_0311, Calkr_0736, Calkr_1713, Calkr_1984, Calkr_2362, Calkr_2364
Xylose isomerase domain protein TIM barrel: Csac_1154
5.3.1.5
6-phospho-D-gluconate
D-xylulose
6-phosphogluconate dehydrogenase NAD binding protein: Csac_2711
1.1.1.-
1.1.1.351
CO2
ATP
ADP, H+
Xylulokinase: Calkr_2026
Xylulokinase: Csac_0798
2.7.1.17
D-ribulose 5-phosphate
Ribulose-phosphate 3-epimerase: Csac_2074
5.1.3.1
Ribulose-phosphate 3-epimerase: Calkr_1036
5.1.3.1
sugar-phosphate isomerase, RpiB/LacA/LacB family: Csac_1200
5.3.1.6
sugar-phosphate isomerase, RpiB/LacA/LacB family: Calkr_0590
D-xylulose-5-phosphate
D-ribose-5-phosphate
phosphate
H2O
4.1.2.9
4.1.2.9
D-sedoheptulose-7-phosphate
D-glyceraldehyde-3-phosphate
Transaldolase: Calkr_1474
Putative transaldolase: Csac_2036
2.2.1.2
Transketolase domain protein: Csac_1351
Transketolase, central region: Csac_1352
2.2.1.1
transketolase domain-containing protein: Calkr_0504, Calkr_2357
transketolase central region: Calkr_0505
D-erythrose-4-phosphate
D-fructose-6-phosphate

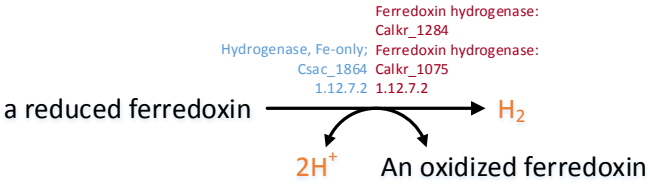


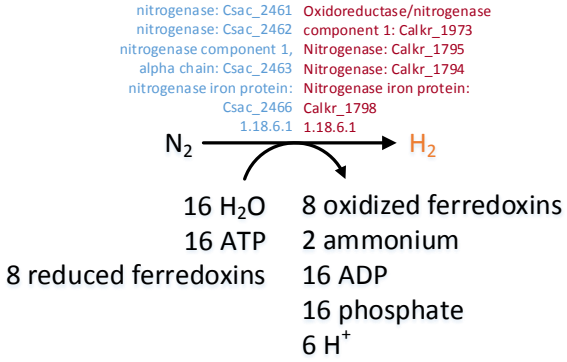
Fermentation



Hydrogen Production







Appendix E. Pathway enzymes expressed in RNA-seq and Proteomics data

The following tables outline the locus tags obtained from BioCyc genome annotations which are also illustrated in the metabolic pathways in Appendix C, as they are expressed in RNA-sequencing and proteomics data. RNA transcripts were observed in all conditions of growth while proteins were only expressed in some conditions.

C. saccharolyticus			
Process/ Pathway	**Enzyme**	**Locus Tags**	**Protein Expressed**
Cellulose Degradation	Cellulase	Csac_0678	Cellobiose, Avicel, CMC, SG
	Cellulase, Cellulose 1,4-beta-cellobiosidase	Csac_1079	Avicel
	Cellulose 1,4-beta-cellobiosidase	Csac_1078	Avicel
	Cellulose 1,4-beta-cellobiosidase	Csac_2410	Avicel, SG, Xylan
	Beta-glucosidase	Csac_1089	All except Xylose, Avicel and SG
	Endo-1,4-beta-xylanase	Csac_0696	All except Cellobiose and Glucose
	Endo-1,4-beta-xylanase	Csac_2405	None
	Endo-1,4-beta-xylanase	Csac_2408	Avicel, SG, CMC, Xylan
	Glycoside hydrolase, family 5	Csac_1077	Avicel
	Glycoside hydrolase, family 5	Csac_0137	Xylose
	Glycoside hydrolase, family 3	Csac_0586	All
	Glycoside hydrolase, family 3	Csac_1102	All
	Glycoside hydrolase, family 48	Csac_1076	Avicel, CMC, SG
	Glycosyltransferase 36	Csac_1091	All except xylose
	Phosphoglucomutase	Csac_2295	All
EM Pathway	Glucokinase	Csac_0778	All
	Phosphoglucose isomerase	Csac_1187	All
	6-phosphofructokinase	Csac_1830	All
	6-phosphofructokinase	Csac_2366	Xylose, Cellobiose, CMC
	Fructose-1,6-bisphosphate aldolase	Csac_1189	All
	Phosphoglycerate kinase	Csac_1952	All
ED Pathway	2-dehydro-3-deoxyphosphogluconate	Csac_0354	All
	Phosphate acetyltransferase	Csac_2041	All
Glycolysis	Glyceraldehyde-3-phosphate dehydrogenase, type I	Csac_1953	All
	Phosphoglycerate kinase	Csac_1952	All
	Phosphoglycerate mutase	Csac_1886	None
	Phosphoglycerate mutase	Csac_1937	None
	phosphoglycerate mutase	Csac_1951	All
	Enolase	Csac_1950	All
	Pyruvate kinase	Csac_1831	All
	Pyruvate/ketoisovalerate oxidoreductase, gamma	Csac_1458	All
	pyruvate ferredoxin/flavodoxin oxidoreductase	Csac_1459	All

	pyruvate flavodoxin/ferredoxin oxidoreductase	Csac_1460	All
	pyruvate flavodoxin/ferredoxin oxidoreductase	Csac_1549	All except SG
	pyruvate ferredoxin/flavodoxin oxidoreductase	Csac_1551	All except SG
Pentose Phosphate Pathway	Xylose isomerase domain protein	Csac_1154	All
	Xylulokinase	Csac_0798	All
	Transaldolase	Csac_2036	All
	6-phosphogluconate dehydrogenase, NAD-binding	Csac_2711	All
	Sugar-phosphate isomerase, RpiB/LacA/LacB family	Csac_1200	All except Avicel and SG
	Ribulose-phosphate 3-epimerase	Csac_2074	All except Cellobiose and SG
	Transketolase domain protein	Csac_1351	All
	Transketolase, central region	Csac_1352	All
Fermentation	Phosphate acetyltransferase	Csac_2041	All
	Acetate kinase	Csac_2040	All
	Iron-containing alcohol dehydrogenase	Csac_0711	All
	Alcohol Dehydrogenase	Csac_0395	All
	Alcohol dehydrogenase	Csac_0407	YE and CMC
	Alcohol dehydrogenase, zinc binding domain protein	Csac_0395	All
	Molybdopterin oxidoreductase	Csac_0621	All
	Lactate dehydrogenase	Csac_0351	All except Avicel and SG
	L-lactate dehydrogenase	Csac_1027	All
Hydrogen Production	Hydrogenase, Fe-only	Csac_1864	All
TCA Cycle	Aconitate hydratase	Csac_0750	All except SG
	4Fe-4S ferredoxin, iron-sulfur binding domain	Csac_1548	Glucose, xylose, YE, xylan
	Thiamine pyrophosphate enzyme domain protein	Csac_1550	All except SG
	Hydrolase, Fe-S type, tartrate/fumarate	Csac_0738	Glucose, xylose, cellobiose, Avicel
	Hydrolase, Fe-S type, tartrate/fumarate	Csac_2759	All except SG

C. kristjanssonii			
Process/ Pathway	**Enzyme**	**Locus Tags**	**Protein Expressed**
Cellulose Degradation	Cellulase	Calkr_2522	All
	Cellulose 1,4-beta-cellobiosidase	Calkr_2245	All except Xylose, YE, CMC
	Endo-1,4-beta-xylanase	Calkr_0572	Avicel, SG, CMC, Xylan
	Glycoside hydrolase family 5	Calkr_0231	Avicel
	Glycoside hydrolase family 5	Calkr_2007	All except Glucose, Xylose and YE
	Glycoside hydrolase family 3	Calkr_2396	All
	Glycosyltransferase 36	Calkr_2205	All
	Glycosyltransferase 36	Calkr_2206	All
	ROK family protein	Calkr_0571	All
	ROK family protein	Calkr_2045	All
ED Pathway	2-dehydro-3-deoxyphosphogluconate aldolase	Calkr_1981	All
	Phosphate acetyltransferase	Calkr_1479	All
Glycolysis	Glyceraldehyde-3-phosphate dehydrogenase	Calkr_1392	All
	Triosephosphate isomerase	Calkr_1391	All
	Enolase	Calkr_1389	All
	Pyruvate kinase	Calkr_1247	All
	Pyruvate/ketoisovalerate oxidoreductase, subunit gamma	Calkr_0642	All
	Pyruvate ferredoxin/flavodoxin oxidoreductase domain containing protein	Calkr_0643	All
	Pyruvate ferredoxin/flavodoxin oxidoreductase domain containing protein	Calkr_0644	All
	Pyruvate flavodoxin/ferredoxin oxidoreductase containing protein	Calkr_1085	All
	Pyruvate flavodoxin/ferredoxin oxidoreductase containing protein	Calkr_1087	All except CMC
	Pyruvate flavodoxin/ferredoxin oxidoreductase domain containing protein	Calkr_1722	All
	Pyruvate flavodoxin/ferredoxin oxidoreductase domain containing protein	Calkr_1723	All
Pentose Phosphate Pathway	Xylose isomerase domain protein	Calkr_0311	All
	Xylose isomerase domain protein	Calkr_0736	All
	Xylose isomerase domain protein	Calkr_1713	All
	Xylose isomerase domain protein	Calkr_1984	All
	Xylose isomerase domain protein	Calkr_2362	All
	Xylose isomerase domain protein	Calkr_2364	Xylan
	Xylulokinase	Calkr_2026	All
	Transaldolase	Calkr_1474	All
	Transketolase central region	Calkr_0505	All
	Transketolase domain-containing protein	Calkr_2357	CMC and SG
	Sugar-phosphate isomerase, rpib/laca/lacb family	Calkr_0590	All except SG
	Ribulose-phosphate 3-epimerase	Calkr_1036	Avicel and CMC

Fermentation	Acetate kinase	Calkr_1478	All
	Iron-containing alcohol dehydrogenase	Calkr_0933	All
	Iron-containing alcohol dehydrogenase	Calkr_2248	All
	L-lactate dehydrogenase	Calkr_0782	All
Hydrogen Production	Hydrogenase, Fe-only	Calkr_1284	All
	Ferredoxin hydrogenase	Calkr_1075	All
	Nitrogenase	Calkr_1794	None
	Nitrogenase	Calkr_1795	None
	Nitrogenase iron protein	Calkr_1798	None
	Oxidoreductase/nitrogenase component 1	Calkr_1973	YE and SG
	Carbon-monoxide dehydrogenase	Calkr_0175	All
TCA	aconitate hydratase	Calkr_1786	All
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_0176	Xylose
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_0500	None
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_0554	None
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1084	Xylan
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1155	None
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1175	None
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1777	None
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1812	CMC
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_1891	All except SG
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_2240	All except Glucose and SG
	4Fe-4S ferredoxin iron-sulfur binding domain containing protein	Calkr_2404	All except Glucose and SG

Note that although EMP enzymes were not annotated from the BioCyc genome, they were expressed in *C. kristjanssonii* RNA-seq and proteomic data.