Integration of Genome Content, Enzyme Activities, and Expression Profiles in Assessing Changes in End-Product Yields in *Clostridium thermocellum*

by

Thomas Rydzak

A Thesis Submitted to the Faculty of Graduate Studies of
The University of Manitoba
In partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Microbiology
University of Manitoba
Winnipeg

Integration of Genome Content, Enzyme Activities, and Expression Profiles in Assessing Changes in End-Product Yields in *Clostridium thermocellum*

by

Thomas Rydzak

University of Manitoba, 2013

Supervisory Committee

Dr. Richard Sparling (Department of Microbiology, University of Manitoba)

Supervisor

Dr. Ivan Oresnik (Department of Microbiology, University of Manitoba)

Departmental Member

Dr. Deborah Court (Department of Microbiology, University of Manitoba)

Departmental Member

Dr. Nazim Cicek (Department of Biosystems Engineering, University of Manitoba)

Department Member

Dr. David Wu (Department of Chemical Engineering, University of Rochester)

External Examiner

Supervisory Committee

Dr. Richard Sparling (Department of Microbiology, University of Manitoba)

Supervisor

Dr. Ivan Oresnik (Department of Microbiology, University of Manitoba)

Departmental Member

Dr. Deborah Court (Department of Microbiology, University of Manitoba)

Departmental Member

Dr. Nazim Cicek (Department of Biosystems Engineering, University of Manitoba)

Department Member

Dr. David Wu (Department of Chemical Engineering, Universiy of Rochester)

External Examiner

Abstract

Clostridium thermocellum is a fermentative, Gram-positive, thermophile capable of cellulosome-mediated breakdown of hemicellulose and simultaneous biofuels (ethanol and H_2) production, and is thus an excellent candidate for consolidated bioprocessing. However, ethanol and/or H_2 production yields are below theoretical maxima due to branched product pathways. Biofuel yields may be improved by manipulation of fermentation conditions or implementation of rational metabolic engineering strategies. However, the latter relies on a thorough understanding of gene content, gene product expression, enzyme activity, and intracellular metabolite levels, which can all influence carbon and electron flux. The thesis work represents the first large-scale attempt in combining bioinformatic, enzymatic, proteomic, and culture perturbation approaches to systematically understand these interactions.

C. thermocellum was used to investigate how these parameters affect end-product yields. Enzyme activities involved in conversion of pyruvate to end-products were consistent with end-product profiles and draft genome annotation. NADH and NADPH-dependent alcohol dehydrogenase (ADH) activities were comparable, whereas NADPH-dependent hydrogenase activities were higher than NADH and ferredoxin-dependent hydrogenase activities. While product yields changed in response to exogenous end-product additions, most core fermentative enzyme activities did not, suggesting that these changes may be governed by thermodynamics. The lack of major changes (>2-fold) in expression in response to growth and gas sparging was further confirmed by proteomics and RT-qPCR, respectively, although the latter revealed that ADH expression changes in response to gas sparging.

Improved genome curation allowed refinement of metabolic pathways. A genomic and end-product meta-analysis of ethanol and/or H_2 producing fermentative bacteria revealed that presence/absence of genes encoding hydrogenases and aldehyde dehydrogenases/ADHs had the greatest impacts on biofuel yields. However, genome content alone did not necessarily explain end-product yields.

Given that genomic analysis of *C. thermocellum* revealed the presence of redundant genes encoding enzymes with analogous functions, shotgun and multiple reaction monitoring proteomics was used to refine which proteins are expressed. Absence/low expression of aldehyde dehydrogenase, ferredoxin-dependent hydrogenase and NADH:ferredoxin oxidoreductase suggest that these enzymes may not play a significant role in metabolism. An alternative electron flow pathway is proposed to explain end-product synthesis patterns in response to pyruvate addition or presence of protein inhibitors (CO, hypophosphite).

Acknowledgements

The support, guidance, and friendships of so many people have made this endeavour possible. Firstly, I would like to thank my advisor, Dr. Richard Sparling, for his guidance, unwavering support, and the opportunity to persue this research. Thank you for challenging me, believing in me, and allowing me the flexibility and independence to branch out and peruse new challenges. I would also like to thank Dr. David B. Levin for his dedication and tireless efforts to ensure the success of this project. Your endless optimism has taught me to 'persevere' through challenging times.

I would like to extend my appreciation to my supervisory committee members. Dr. Nazim Cicek, thank you for your worldly advice and showing me the importance of balancing life. Dr. Ivan Oresnik, thank you for always challenging me and teaching me to stand up for what I believe in. It has been both humbling and empowering. Finally, Dr. Deborah Court, thank you for your kindness and mentorship. Your classes were one of the reasons why I continued in this endeavour.

Last but not least, my sincerest thanks to all of my collaborators, mentors, and lab members at the University of Manitoba. There have been too many to mention. Special thanks to John A. Wilkins and Oleg V. Krokhin for allowing me to pursue proteomics and work in your lab, Peyman Ezzati for your mentorship and technical support, Vic Spicer, for helping me sort through and make sense vast fields of data, Patric Chong, for teaching me the ropes, and Sharon Berg for all your help throughout the years. Thank you to all the members of our lab(s) for your friendship, help, support, and ideas: Carlo Carere, Tobin Verbeke, Marcel Taillefer, Scott Wushke, Umesh Ramachandran, Rumana Islam, Marina Grigoryan, Warren Blunt, Nathan Wrana, John Schellenberg, Peter McQueen, Zack Cunningham, Ian Cunningham, David Yu, and so many others that are too numerous to mention. Here's to late nights and early mornings.

Dedication

To my parents, *Grace* and *Henryk*

Thank you for teaching me the importance of hard work, challenging me, and having unconditional and unwavering faith in me, even when logic would disagree.

And to my partner, *Monika*

Thank you for your unbelievable patience and support. You have given me reason, stability, serenity, and sanity. You are my rock in an ocean full of woes.

Table of Contents

List of Tables

List of Figures

List of Supplementary Tables and Figures

List of Abbreviations

2D-HPLC-MS/MS	Two dimensional high pressure liquid chromatography tandem mass spectrometry
ABC	ATP-binding cassette
ACK	Acetate kinase
ADH	Alcohol dehydrogenase
AdhE	Acetaldehyde/alcohol dehydrogenase (bifunctional)
ADP	Adenosine diphosphate
AldH	Aldehyde dehydrogenase
AMP	Adenosine monophosphate
asRNA	Anti-sense RNA
ATCC	American Type Culture Collection
ATK	Acetate thiokinase
ATP	Adenosine triphosphate
BadH	Butyrylaldehyde dehydrogenase
BCA	Bicinchononic acid
BDH	Butanol dehydrogenase
BK	Butyrate kinase
BLAST	Basic Local Alignment Search Tool
BP	British Petroleum
BSA	Bovine serum albumin
C_1/C_2	Ratio of single carbon (CO_2 and formate) versus two carbon (acetate and ethanol) end-products
CAZyme	Carbohydrate active enzymes
CB phosphorylase	Cellobiose phosphorylase
CBP	Consolidated bioprocessing
CD phosphorylase	Cellodextrin phosphorylase
CFC	Chlorofluorocarbon
CipA	Cellulosome-integrating protein

CN	Cyanide
CO	Carbon monoxide
CO_2	Carbon dioxide
COGs	Clusters of orthologous groups
DOE	Department of Energy
DTT	Dithiothreitol
DSM(Z)	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German Collection of Microorganisms and Cell cultures)
E:A	Ethanol-to-acetate ratio
Ech	Enery conserving hydrogenase
EDTA	Ethylenediaminetetraacetic acid
EISA	Energy Independency and Security Act
EMP	Embdem-Meyerhof-Parnas
$E^{o'}$	Standard reduction potential
Fd	Ferredoxin
FDH	Formate dehydrogenase
FDP	Fructose-1,6-bisphosphate
FHL	Formate hydrogen lyase
G-1-P	Glucose-1-phosphate
G-6-P	Glucose-6-phosphate
GADPH	Glyceraldehyde-3-phosphate dehydrogenase
GAPFOR	Glyceraldehyde-3-phosphate:ferredoxin oxidoreductase
GHG	Greenhouse gasses
H_2	Hydrogen gas
H_2ase	Hydrogenase
HBD	β-hydroxybutyryl-CoA dehydrogenase
HK	Hexokinase
HPLC	High pressure liquid chromatography
HPP	Hypophosphite
IMG	Integrated Microbial Genomes

IPCC	Intergovernmental Panel on Climate Change
iTRAQ	Isobaric tags for relative and absolute quantitation
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG ontology
LDH	Lactate dehydrogenase
MalE	Malic enzyme
Mbh	Membrane bound hydrogenase
MDH	Malate dehydrogenase
MGF	Mascot Generic File
Mr	Molecular mass
MRM	Multiple reaction monitoring
Mtt	Membrane targeting and translocating
MV	Methylviologen
N_2	Nitrogen gas
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
NfnAB	NADH-dependent reduced ferredoxin:$NADP^+$ oxidoreductase
NFOR	NADH:ferredoxin oxidoreductase
NSERC	Natural Sciences and Engineering Research Council of Canada
O/R	Oxidation/reduction ratio of oxidized to reduced end-products
OAADC	Oxaloacetate decarboxylase
OD	Optical density
$p(H_2)$	Hydrogen gas partial pressure
PAS	Per-ARNT-Sim
PBS	Phophate buffered saline
PCR	Polymerase chain reaction
PDC	Pyruvate decarboxylase
PDH	Pyruvate dehydrogenase
PEP	Phosphoenol pyruvate
PEPCK	Phosphoenolpyruvate carboxykinase
PEPS	Phosphoenol pyruvate dikinase

PET	Production of Ethanol (operon)
PFK	Phosphofructokinase
PFL	Pruvate:formate lyase
PFOR	Pyruvate:ferredoxin oxidoreductase
PGK	Phosphoglycerate kinase
PGM	Phosphoglucomutase
Pi	Inorganic phosphate
PMF	Proton motive force
PPase	Pyrophosphatase
PPDK	Pyruvate phosphate dikinase
PP_i	Pyrophosphate
PPK	Pyruvate phosphate kinase
PPP	Pentose phosphate pathway
PSM	Proton suicide method
PTA	Phosphotransacetylase
PTB	Phosphotransbutyrylase
PTS	Phosphotransferase system
RAI	Relative abundance index
Rex	Redox-sensing transcriptional regulator
△G°’	Standard free energy change
RNAseq	Whole transcriptome shotgun sequencing
Rnf	Rhodobacter nitrogen fixing
RPS	Redox sensing protein
RPS-BLAST	Reverse porition specific BLAST
RT	Reverse transcription
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
RV_{EP}	Total molar reduction values of reduced end-products (H_2 + ethanol).
SHL	Surface-layer homology
SpC	Spectral count

SSCF	Simultaneous saccharification and co-fermentation
SSF	Simultaneous saccharification and fermentation
Tat	Twin argenine translocation
TFA	Trifluoroacetic acid
THIL	Acetyl-CoA acetyltransferase
Vdiff	Vector difference

Chapter 1: Literature Review

1.1 Introduction

Fossil fuels, which include petroleum, natural gas, and coal, provide over 80% of the world's energy (The World Bank 2012). In addition to being a major contributor to air, soil and water pollution, the rate of the world's consumption of fossil fuels greatly exceeds their formation rates. The predicted increase in world energy consumption (Ruehl 2011) in conjunction with political pressures has driven the search for an alternative energy currency that is sustainable, environmentally benign, and that can be directly used in the transportation sector.

Biofuels, which include bioalcohols (ethanol, butanol), biogases (hydrogen, methane, syngas), and bio-oils (biodiesel, vegetable oils) provide such an alternative. These biofuels may be derived from sugar, starches, oil seeds, or cellulosic biomass. Unfortunately, first (1^{st}) generation biofuels-producing technologies primarily rely on the use of food crops including sugar cane or sugar beets, corn or wheat, and canola or soy seeds, which provide the necessary sugars, starches, and oils, respectively, required for the production of biofuels. The use of such food crops for biofuels production is not only costly, making biofuels less economically competitive with petroleum-based fuels, but has also been linked to increasing food costs and poses environmental concerns.

Second (2^{nd}) generation biofuels, produced from lignocellulosic biomass, provide a potentially low cost, environmentally and economically friendly alternative to 1^{st} generation biofuels. Lignonocellulose, which is the structural component of plant cell walls, is the most abundant raw material on earth and accounts for large amounts of forestry, agricultural, and municipal waste streams. While its high abundance and low cost makes it an appealing substrate

for fermentative biofuels production, processing of lignocellulose to liberate hexose and pentose sugars (derived from cellulose and hemicellulose, respectively) required for microbe-mediated biofuels production has been economically prohibitive. Conventional 2nd generation biofuels production technologies employ a muti-step process involving separate (i) cellulase production, (ii) cellulose and hemicelluloses hydrolysis, (iii) hexose fermentation, and (iv) pentose fermentation, or, more recently, combinations thereof (Lynd et al. 2002). Consolidated bioprocessing (CBP), however, which features simultaneous cellulose production, cellulose hydrolysis, and fermentation of hexose and pentose sugars, is an attractive processing strategy that lowers biofuel production costs (Lynd et al. 2005).

Although CBP offers tremendous potential for cellulosic biofuel production, microorganisms capable of efficient cellulose degradation and utilization combined with high production yields of desired products are currently not available for CBP. Successful development of such organisms requires a good understanding of the processes involved in cellulose degradation and the metabolic pathways that dictate product yields. Current metabolic engineering strategies aim to express heterologous cellulase systems in organisms with good biofuel yields or to improve end-product yields in cellulolytic bacteria (Lynd et al. 2005). The focus of this research is the latter.

The transition from fossil fuel-based transportation fuels to biofuels via consolidated bioprocessing is currently hindered by biofuels production costs. The economic viability of these biofuels may, in part, be improved by improving production yields. Improvement of biofuel yields of cellulolytic bacteria, through either the manipulation of fermentation conditions or genetic engineering, requires a detailed understanding of how gene content, gene and gene-

product expression, intracellular metabolic levels, and thermodynamics of product formation dictate end-product yields and formation rates. The impact of these key factors on ethanol and H_2 yields will be discussed, with particular emphasis on *Clostridium thermocellum*, the model organism used in this research.

1.2 Rationale for biofuels

1.2.1 Economic and political pressures

Fossil fuels, including petroleum, natural gas, and coal, currently provide 87% of the world's energy (The World Bank 2012; Rodrigue and Comtois 2013). Projected energy consumption is expected to grow at an average rate of 2% per year globally (Energy Information Administration 2006b; Energy Information Administration 2006a; International Energy Agency 2006) and is predicted to increase by 44% from 2006 to 2030 (Energy Information Administration 2009). Economic growth in various non-OECD (Organization for Economic Co-operation and Development) countries, particularly China and India, will drive energy consumption in these regions at rate of 3% annually (Energy Information Administration 2006b; Energy Information Administration 2006a). The transportation sector, 95% of which is reliant on petroleum products, accounts for approximately 25% of the world's energy demands, and accounts for about 61% of oil consumption annually (Rodrigue and Comtois 2013).

While there are ample reserves of energy to meet the demands of future populations, there is increasing difficulty in excavating these reserves. Easily attainable light crude oil, which has a low sulfur content and is thus easily refined, is more or less depleted (Levin et al. 2011b). Excavation and refinement of low quality off-shore or oil-sand reserves, is more expensive and

energy intensive, reducing its productivity and economic feasibility (Kleykamp 2008; Rhodes 2008). Furthermore, these various reserves are not distributed evenly around the world, which leads to political tensions, economic distress, and market volatility. In July 2008, WTI (West Texas Intermediate) light sweet crude oil prices peaked at $145 US per barrel, a stark comparison to $40 US per barrel in the 1980s. Although the Goldman Sachs Group's projections of $200 US per barrel (predicted for 2012) did not come into fruition, likely due to an economic down-turn, oil prices still remain high, at roughly $100 US per barrel. Classic supply and demand pressures will always heavily influence oil prices, and thus increasing demand, in conjunction with shortages in stockpiles, refining capacities, or political/civil unrest, particularly in oil-wealthy countries, will drive oil prices higher. Hence, there is an urgent need to develop sustainable and economically viable energy systems that can displace dependence on fossil fuels and promote energy security.

Current policies promote the production and use of bioethanol to reduce oil consumption, presumably because it can be used as a transportation fuel with little change to current technologies. In 2012, 107 billion liters of ethanol were produced globally, with the U.S. and Brazil accounting for 87% of the market share (Renewable Fuels Association 2012). In 2010, about 49 billion liters of ethanol were blended into the U.S. gasoline supply, accounting for about 10% of gasoline consumption (Levin et al. 2011b). In 2007, the U.S. Energy Independency and Security Act (EISA) proposed to increase its annual production of ethanol from 49 billion liters in 2011 to 136 billion liters by 2022 in order to reduce oil consumption by 25% (U.S. Department of Energy 2007). While current bioethanol production heavily relies on the use of corn or sugarcane, in the U.S. and Brazil respectively, 2022 U.S. EISA targets for

ethanol production from cellulosic feedstocks are 61 billion liters per year, accounting for 44% of total bioethanol targets (U.S. Department of Energy 2007). In order to promote cellulosic bioethanol production in 2011, U.S. ethanol producers received a tax credit of $1.01 US per gallon of cellulosic bioethanol produced compared to a $0.55 US per gallon credit for bioethanol produced from other feedstocks (Congressional Budget Office 2010).

1.2.2 Environmental concerns

Extraction, production, refining, processing, and combustion of fossil fuels contributes to pollution of air, water, and soil, and is thus hazardous to the environment and public health (National Research Council 1999). In fact, according to the University of Miami's Clean Energy Research Institute, worldwide environmental damage and associated health damage to living organisms in 1997 alone was estimated to be $2700 billion, 14% of the Gross World Product (Veziroğlu 1997). Water is used in many ways in every step of oil, natural gas, or coal extraction and processing and thus may become contaminated with a variety of pollutants, including hydrocarbons, heavy metals, arsenic, selenium, acids, synthetic chemicals, and other hazardous materials (Allen et al. 2011). These contaminated waters, which are held in retaining ponds created from toxic tailings piles, can contaminate surface and groundwater and leach into surrounding soil, affecting ecosystems and land productivity. In some cases, spills and other disasters, such as the recent Deepwater Horizon drilling rig spill (a.k.a. BP oil spill) which leaked over 4.9 million barrels of crude oil into the Gulf of Mexico, can lead to the contamination of natural environments and lead to devastating financial burdens (Allen et al. 2011).

In addition to soil and water contamination, the combustion of fossil fuels contributes to greenhouse gas (GHG) emissions. Within the transportation sector, which accounts for approximately 25% of the world's energy demands and 61% of oil consumption, gasoline use is predominant, particularly for light duty vehicles (Rodrigue and Comtois 2013). Gasoline *can* be completely combusted to water and CO_2 ($2\ C_8H_{18} + 25\ O_2 \rightarrow 16\ CO_2 + 18\ H_2O + 5461$ kJ/mol energy), producing 2.3 kg CO_2 per liter of gasoline combusted. In reality, the combustion of gasoline is imperfect and incomplete and therefore also generates nitrogen oxides (NO_x), sulfur oxides (SO_2), carbon monoxide (CO), hydrocarbons (benzene, formaldehyde, butandiene, acetaldehyde), and volatile organic compounds (Rodrigue and Comtois 2013). In addition to the dangers of these gasses to human health, GHGs, namely CO_2, NO_x, methane (CH_4) and chlorofluorocarbons (CFCs), prevent heat reflected from the earth's surface from escaping back into space, which in turn can lead to global climate change (Brown et al. 1998). The Intergovernmental Panel on Climate Change (IPCC) has reported that the mean global surface temperature has increased about 0.2°C per decade since 1975, more than twice that of the average temperature increase over the entire 100 year span from 1906 to 2005 (Brown et al. 1998; Intergovernmental Panel on Climate Change 2007). This increase in warming trend correlates well with the 70% increase in global GHG emissions caused by human activities after pre-industrial times (between 1970 and 2004) (Intergovernmental Panel on Climate Change 2007). In recent years, atmospheric concentration of CO_2, which is the major contributor to greenhouse warming, has been increasing at about 0.5% annually (Brown et al. 1998), 57% of which is caused by anthropogenic fossil fuel use (Intergovernmental Panel on Climate Change 2007).

These anthropogenic climate change drivers have not only increased mean global temperatures, but have also promoted (i) widespread melting of snow and ice, (ii) increased sea levels, (iii) regional changes in precipitation patterns, (vi), and increased frequency of extreme weather conditions (Intergovernmental Panel on Climate Change 2007). These changes in turn affect natural ecosystems, water resources, food security, coastal regions, industry, settlements and society, and human health (Intergovernmental Panel on Climate Change 2007). In order to prevent these potentially catastrophic events, reliance on fossil fuels and GHG emissions must be decreased. Development and implication of sustainable, clean burning energy sources is of paramount importance.

1.3 Bioethanol and biohydrogen

With the transportation sector accounting for approximately 25% of the world's energy demands and 61% of oil consumption, clean burning transportation fuels that can reduce GHG emissions are of paramount importance. Given that biofuels are derived from plant material, the life cycle of biofuel combustion and production is carbon neutral because of the ability to fix CO_2 within the same time frame as these gases are liberated through combustion. In addition to the ability to reduce GHG emissions, novel transportation fuels should ideally be (i) easily stored, (ii) easily transported, (iii) compatible with current technologies, (iv) have a high energy densities, and (v) be produced at a cost comparable to fossil fuels. Although all these criteria are difficult to meet, bioethanol and biohydrogen are promising preliminary candidates.

1.3.1 Bioethanol

Ethanol (C_2H_5OH) is the most widely used biofuel in the world. It is a flammable, clear and colorless liquid, with a melting point of -114.1°C and a boiling point of 78.5°C. It's enthalpy of combustion is -1370.7 kJ/mol. While it is easily stored, it is hygroscopic and corrosive, making it difficult to transport through modern pipelines over long distances (Horn and Krupp 2009). It has a higher octane rating than gasoline (109 vs 91), and thus can withstand higher compression ratios before detonating, allowing for improved engine performance. However, the energy content of ethanol (per liter) is lower than that of gasoline (26.8 vs 44.4 MJ/kg), and thus larger volumes are required to travel the same distance.

Ethanol is widely used as a gasoline additive providing a cleaner-burning fuel with a higher octane rating. Emissions from combustion of a 10% ethanol-blended gasoline reduce CO, CO_2, and NO_x/SO_x by 30, 10, and 7%, respectively, when compared to unblended gasoline (Levin et al. 2011a). Currently, the most popular ethanol-gasoline formulations in North America are E10 (Gashol) and E85, which contain 10% and 85% (%v/v) ethanol, respectively (Wheals et al. 1999). While Gashol is compatible with standard vehicles on the road today, more recently developed flex-fuel engines can run on 0% to 85% ethanol. In Brazil, gasoline contains 24% ethanol, and more than 20% of vehicles, which include ethanol-only and flex-fuel vehicles, can run on hydrous (95%) ethanol (Wheals et al. 1999).

Ethanol can be produced as either a petrochemical or via fermentation. Synthetic ethanol made from petroleum-based feedstocks, is made via the acid-catalyzed hydration of ethylene ($C_2H_4 + H_2O \rightarrow CH_3CH_2OH$) using phosphoric acid adsorbed onto a porous support such as silica gel and high pressure steam (at 300°C) (Kosaric et al. 2005). While this process was used

in large scale ethanol production by Shell in 1947, and still used by LyondellBasell Industries, the majority of ethanol is produced using biological processes, namely fermentation. During fermentation, glucose (or sucrose) is converted to ethanol and CO_2 ($C_6H_{12}O_6 \rightarrow 2\ CH_3CH_2OH + 2\ CO_2$ or $C_{12}H_{22}O_{11} + H_2O \rightarrow 4\ CH_3CH_2OH + 4\ CO_2$) under oxygen-limited conditions. The primary enzymes responsible for ethanol production are acetaldehyde dehydrogenase and alcohol dehydrogenase, which utilize NADH produced during glycolysis to reduce pyruvate to acetaldehyde, and acetaldehyde to ethanol, respectively (discussed below). Currently, the most commonly used species in industrial ethanol fermentation are baker's yeast (*Saccharomyces cerevisiae*) or *Zymomonas mobilis*, a Gram-negative proteobacteria. While these organisms make near-maximal ethanol yields (~90%) of 2 mol ethanol per mol glucose, they cannot utilize cellulosic biomass directly and require more refined sugars or starches. Furthermore, while these organisms have high ethanol tolerance (~15% v/v) compared to most other organisms (Scopes and Griffiths-Smith 1986; Morais et al. 1996), the resulting ethanol must be further concentrated to remove water for industrial uses using distillation, molecular sieves and desiccants, or reverse osmosis using membranes.

Currently, North America and Brazil, the largest producers of bioethanol, produce ethanol from corn (maize) and sugarcane respectively. One of the major arguments against the use of corn for bioethanol production is that this process has poor energy balances and high carbon intensities. When compared to sugarcane, corn yields are only about 15% of that of sugarcane (based on tonnes per hectare), and therefore average ethanol yields are roughly 35% of that of sugarcane (based on liters per hectare). Furthermore, while ethanol yields are comparable for corn and sugarcane (89 and 92%, respectively) typical fermentation times of ethanol

production from corn are more than 4-fold of that of sugar cane. Fermentation, along with more rigorous processing methods, requires a considerable amount of power, typically in the form of electricity generated from coal plants. This in turn offsets the decrease of GHG emissions associated with combustion of ethanol. Nevertheless, the U.S. Department of Energy has reported that, on a life cycle basis, ethanol production from corn results in a 20% reduction of GHG emissions relative to that of gasoline, and can reach as high as 52% with increased efficiency and renewable energy use (Wang et al. 2007). While this is still lower than the 78% reduction in GHG emissions resulting from sugarcane ethanol, it can reach as high as 86% with the use of cellulosic biomass for ethanol production (Wang et al. 2007).

1.3.2 Biohydrogen

Hydrogen (H_2) is a versatile, clean burning, and renewable energy currency that can potentially displace the use of petroleum-based fuels in the transportation sector. It produces only heat and water upon combustion, releasing 286 kJ/mol of energy, and does not generate greenhouse gas emissions. It may be converted into electricity via proton exchange membrane fuel cells (PEMFC). During this process a platinum catalyst causes H_2 to dissociate at the anode to electrons, used to generate a current, and protons, which permeate through a polymer electrolyte membrane and combine with oxygen to form water at the cathode.

Despites its unarguable appeal in terms of energy content, environmental benefits, and utilization technology, H_2 currently constitutes only about 3% of the world's total energy consumption (Das and Veziroğlu 2001). Although it has the highest energy density of any known fuel by mass, three times that of gasoline (142.0 MJ/kg versus 44.4 MJ/kg), it has a very low volumetric energy density at ambient conditions due to its volatility, and must thus be

compressed into a liquid via energy intensive processes leading to economic and energetic drawbacks with packaging, storage, and distribution. More importantly, current methods of H_2 production are expensive, energetically intensive, and potentially environmentally harmful, and thus production of H_2 may be one of the major technological barriers to overcome.

1.3.2.1 *Non-biological hydrogen production*

H_2 fuel can be produced from a diverse array of potential feedstocks including fossil fuels, water, and organic matter using various chemical and electrochemical methods (Table 1.1) (Das and Veziroğlu 2001; Levin et al. 2004; Nath and Das 2004; Demirbas 2009). Currently, over 80% of H_2 production occurs via steam reformation of natural gas (Demirbas 2009). During this process methane, the primary constituent of natural gas, is combined with high temperature steam (700-1000°C) in the presence of a nickel catalyst, breaking it apart into H_2 and CO. The CO produced then reacts with water at high temperatures to further produce H_2 and CO_2 via a process known as the gas shift reaction. Although this process is well understood and less expensive than other methods, it is dependent on a potentially limited supply of natural gas, produces moderate levels of CO_2, and is very energy intensive. Alternatively, H_2 may be produced via coal gasification, during which coal is converted into H_2 and CO_2 via a similar mechanism as natural gas reformation. Regardless of the abundance of coal, however, this process is more expensive and results in much higher CO_2 emissions than natural gas reformation since the production ratio of H_2 to CO_2 is roughly 1:1, when compared to the 4:1 ratio produced via natural gas reformation (Demirbas 2009).

Table 1.1: Comparison of current methods employed for H_2 production

Method	Advantages	Disadvantages
Reformation of Natural Gas ($CH_4 + H_2O \rightarrow CO + 3H_2$) ($CO + H_2O \rightarrow CO_2 + H_2$)	• Most common (80% of H_2 made) • Most inexpensive • Well understood	• Dependant on limited natural gas supply • Not carbon neutral
Gasification of Coal ($C + H_2O \rightarrow CO + H_2$) ($CO + H_2O \rightarrow CO_2 + H_2$)	• Coal is abundant	• Low yields • High CO_2 emissions • Sulfur & CO emissions
Electrolysis of Water ($2H_2O \rightarrow 2H_2 + O_2$)	• 2nd most common method used • Well understood • Potentially emission free if generated from clean energy sources	• Very energy intensive • May result in high GHG if fossil fuel based methods are used
Biomass Reformation ($C_4H_7O_2N \rightarrow H_2 + CO + CH_4$) ($CH_4 + H_2O \rightarrow CO + 3H_2$) ($CO + H_2O \rightarrow CO_2 + H_2$)	• Potentially carbon neutral • Inexpensive • Can utilize organic waste	• Poorly understood
Biological H_2 Production ($C_6H_{12}O_6 + H_2O \rightarrow H_2 + CO_2 +$ organics) ($2H_2O + light \rightarrow 2H_2 + O_2$)	• Carbon neutral • May use light or waste biomass • Not as energy intensive	• May have poor yields • May require constant light • Poorly understood

A number of alternative processes can be used for the production of H_2 that do not result in a net increase in greenhouse gas emissions. The electrolysis of water, during which an electric current is passed through water, splitting it up into H_2 and O_2, is the second most commonly used method of H_2 production. Although this process can be potentially emission free providing that clean, renewable sources such as wind, solar, hydro, or geothermal energy are used for the generation of electricity, it is the most energy intensive method of H_2 production (Demirbas 2009). Biomass gasification is an alternative method of H_2 production that is potentially carbon neutral. In this process, biomass is broken down into H_2, CH_4, and CO, which can in turn be used for steam reformation and the gas shift reaction. Although these subsequent processes result in net CO_2 production, the overall process of biomass gasification is carbon neutral since the CO_2 produced is fixed via photosynthesis during the generation of new biomass. Furthermore, this process can utilize municipal and agricultural wastes, turning them into valuable commodities.

More recent emphasis has been placed on the development of H_2 via biological H_2 production, which employs the use of H_2-producing microorganisms via light dependant or fermentative processes. Biological H_2 production is an attractive alternative to chemical and electrochemical methods discussed above. It is a potentially carbon neutral process that is carried out at lower temperatures and pressures, and is therefore less energy-intensive than chemical and electrochemical processes. Furthermore, unlike chemical methods, which involve the conversion of non-renewable fossil fuels into H_2, fermentative H_2 production can utilize renewable carbohydrate-based substrates such as waste biomass from municipal, agricultural, and forestry sectors, while light driven

processes utilize light energy, water, and/or CO_2 for the production of H_2 (Hallenbeck 2002).

1.3.2.2 *Methods of biological hydrogen production*

A number of different light-dependent and fermentative biological processes may be employed in the production of H_2 including direct biophotolysis (Green algae), indirect biophotolysis (Cyanobacteria), photo-fermentation (photosynthetic bacteria), and dark fermentation (fermentative bacteria) (Hallenbeck 2002; Nath and Das 2004). In all these biological processes, H_2 production ($2H^+ + 2e^- \leftrightarrow H_2$) is catalyzed by hydrogenases (H_2ases) and/or nitrogenases. These enzymes are capable of shuttling electrons between various electron carriers (including NAD(P)H, ferredoxin (Fd), flavodoxin, coenzyme F_{420}, or cytochromes) and H_2 (Discussed below).

When compared to dark fermentation, biological H_2 productions via light-dependent processes have limited economic potential (Levin et al. 2004). This is, in part due, to the diffuse nature of solar energy (Levin et al. 2004). Given that solar irradiation is as high as 6.6 GJ/m^2/year, Hallenbeck and Benemann estimate that at 10% solar conversion efficiency, and a price of H_2 of $15/GJ, light dependant processes translate to only $10 worth of H_2/m^2/year (Hallenbeck 2002). Note that this does not account for various operational costs and energy inputs, which include mixing, gas sparging, cooling, and gas separation. Furthermore, although solar conversion efficiencies are roughly 10% under low-light conditions, they fall to below 1% under high-light conditions since 90% of photons captured by photosynthetic pigments are decayed as heat or fluorescence due to the rate limiting step of electron transfer between photosystem II and photosystem I

during dark reactions when compared to light capture during light reactions (Kok 1973). Thus, under saturating light conditions the value of light-dependant H_2 production decreases to only \$1/m^2/year.

A number of methods of increasing photon use efficiency have been proposed. One solution involves the rapid mixing of dense cultures to create a 'flashing-light effect', during which each cell is exposed to millisecond flashes of high light intensity, followed by a 5 to 10-fold longer dark period, allowing the dark reactions to catch up with the light reactions (Kok 1973). The use of optical fiber photoreactors, in which light harvested at a particular point is dispersed through the culture using fiber optics is another approach to increasing photon use efficiency, although this has been deemed technically and economically impractical (Hallenbeck 2002). An alternative approach is to develop photosynthetic cells with reduced pigment content in their light harvesting complexes that would allow fewer photons to be absorbed at high light intensities, and therefore less photons would be dissipated while waiting for the dark reactions to take place (Benemann 1990; Nakajima and Ueda 1999; Nakajima and Ueda 2000).

Dark fermentation, in which carbohydrates are converted to H_2, CO_2, and organic acids, provides a promising alternative to light dependent processes, particularly when waste biomass is used as a feedstock for the generation of H_2. Since fermentation does not require a constant light supply, it can be run continuously using inexpensive and commercially used systems. Furthermore, H_2 production rates are much higher using fermentations when compared to photosynthesis-based systems (Levin et al. 2004) reducing bioreactor operation costs. A number of genera, including *Bacillus, Escherichia, Enterobacter, Ruminococcoi* and *Clostridium,* are capable of producing H_2 via

fermentation (Nandi and Sengupta 1998; Claassen et al. 1999; Levin et al. 2004; Nath and Das 2004). Some of these organisms not only operate at thermophillic (40-60°C) conditions, at which the solubility of H_2 is lower preventing product inhibition, but are also capable of degrading lignocellulose, a primary component of plant biomass.

1.4 Lignocellulose for biofuels production

1.4.1 First generation versus second generation biofuels

While biofuels production offers great potential to decrease GHG emissions and promote energy security, the transition from the use of food/feed crops to cellulosic waste streams is crucial for their viability (Searchinger et al. 2008). The primary argument against first (1^{st}) generation biofuels (i.e. those derived from food crops) is the food/feed vs fuel trade-off. The production of food-based biofuels has in fact contributed to higher food prices, food shortages, and decreased food security. For example, prices of maize and wheat increased by 131% and 177%, respectively, from 2005 to 2008 (Sims et al. 2008). While a number of factors, including poor crop yields, rising energy costs, rising agricultural production costs, and declining dollar, have all contributed to increasing food prices, the contribution of 1^{st} generation biofuels to these increased prices have been reported to be anywhere from 3% to 75%, averaging at around 15-25% in most of the literature (Sims et al. 2008). Although the majority of 1^{st} generation biofuel food crops are typically used for animal feed rather than human consumption, increasing costs of animal feed have in turn increased the prices of beef, chicken, pork, eggs, breads, cereals, and milk by 10% to 20% (Brown et al. 1998).

In addition to escalating food prices, a number of other arguments have been made against 1st generation biofuels including: (1) shift in land use for the production of agricultural starch and sugar crops can lead to accelerated deforestation, destruction of wetlands, soil erosion, and destruction of habitats; (2) irrigation of energy crops will lead to increased use of potentially scarce fresh water supplies; (3) increased energy crop cultivation will increased fertilizer, insecticide, and herbicide use and runoff; (4) many food crops are seasonal and thus their availability and costs may fluctuate, which can in turn increase overhead costs due to storage for year-round operations; and (5) biofuel feedstocks are limited and thus cannot provide all of the world's energy (Sims et al. 2008). For example, biofuel production from all of the U.S. corn and soybean production would only meet 12% of gasoline demand and 6% of diesel demand (Srinivasan 2009). Furthermore, if all of the global cropland devoted to production of wheat, rice, corn, sorghum, sugarcane, cassava, and sugar beet, (which accounts for 42% of total cropland) was used for biofuel production, only half of global gasoline consumption would be met (Levin et al. 2011a). Many of the problems associated with 1st generation biofuels can be solved through the production of second (2nd) generation lignocellulosic biofuels.

1.4.2 Lignocellulose composition and pre-treatment

The amount of energy received at the earth's surface in the form of solar energy (2.5×10^{21} Btu/year) stored as carbon via photosynthesis far exceeds that of human consumption by more than ten times (Demain et al. 2005). Roughly two-thirds of the annual carbon dioxide (CO_2) fixed through photosynthesis occurs in terrestrial ecosystems (Breznak and Brune 1994). An estimated 200 billion metric tons of

lignocelluloses, a major component of plant cell walls, is produced yearly, making it the most abundant bio-polymer in the world (da Silva et al. 2012). It is renewable, inexpensive, and constitutes a large fraction of waste biomass from municipal, agricultural, and forestry sectors.

Lignocellulose is composed of (i) cellulose, a linear, insoluble polymer consisting of up to 25,000 repeating β-1,4 linked β-D-glucopyranose (cellobiose) units, (ii) hemicellulose, which is a random, amorphous heteropolysaccharide composed of typically β-1,3 links of xylans, arabinoxylan, gluco-mannan, and galactomannan, (iii) lignin, which is a complex hydrophobic network of phenylpropaniod units, and (iv) pectin, consisting of α-(1-4)-linked D-galacturonic acid (Aristidou and Penttila 2000; Desvaux 2005). The approximate composition of lignocellulose found in most biomass feedstock is roughly 45-60% cellulose, 20-40% hemicellulose, 25% lignin, and 1-5% pectin (Lynd et al. 2002; Demain et al. 2005; Desvaux 2006).

Given that cellulose, the main component of lignocellulose, is composed of hexose sugars, it is the primary substrate for 2nd generation biofuels production. In nature, cellulose consists of tightly packed crystalline regions and highly amorphous regions. The crystalline arrangement of 15-45 chains of cellulose forms a microfibril, which in turn associates to form a cellulose fibril (Desvaux 2005). Hemicellulose, which constitutes 5-20% of the carbohydrate fraction of lignocellulose, is composed of potentially fermentable pentoses (xylose and arabinose) and hexoses (mannose and glucose) (Demain et al. 2005). While both cellulose and hemicelluloses can be fermented, lignin cannot. Furthermore, cellulose and hemicelluloses can be broken down by numerous species of cellulolytic bacteria via cellulases (glycolytic hydrolases),

whereas enzyme-mediated lignin-degradation by bacteria is rear (Aristidou and Penttila 2000; Demain et al. 2005). Although cellulolytic bacteria can breakdown cellulose and hemicellulose components of lignocelluloses without pre-treatment (Islam et al. 2009), delignification of lignocellulose by chemical treatments can help bacterial enzymes gain access to cellulose and hemicellulose.

A number of physical and thermochemical pre-treatment strategies can be employed to improve the hydrolysis of lignocelluloses (reviewed by (Kumar et al. 2009)). These pre-treatment strategies aim to increase the surface area of cellulose by removing the lignin seal, solubilizing hemicelluloses, disrupting the crystallinity of cellulose, and increasing pore size to allow cellulase access. This may be achieved either (i) mechanically by chipping, milling, or grinding, (ii) via pyrolysis, during which lignocelluloses is treated at high temperatures (>300°C), which allows decomposition of cellulose to gaseous products, or (iii) via steam or ammonia fiber explosion, during which biomass is treated with high-pressure steam (or ammonia), and then dropping the pressure to allow explosive decomposition of biomass (Kumar et al. 2009). Alternatively, chemical pre-treatments can be employed including (i) acid hydrolysis (using H_2SO_4 or HCl), which breaks down hemicelluloses to xylose, and then converts xylose to furfural, (ii) alkaline hydrolysis (using NaOH, KOH, $Ca(OH)_2$, NH_4OH), which causes less sugar degradation than acid hydrolysis, or (iii) ozonolysis, which primarily degrades lignin keeping cellulose, and most of the hemicelluloses, intact (Carroll and Somerville 2009; Kumar et al. 2009). Biological pre-treatment strategies using various white-rot fungi (ex. *Pleutotus ostreatus*, *Phanerochaete sordida*, *Pycnoporus cinnabarinus*, *Sporotrichum pulverulentum*) can be used to degrade lignin via

peroxidases and laccase (Kumar et al. 2009). Unfortunately these biological systems also attack cellulose, and thus cellulase-free mutants must be developed to prevent loss of cellulose.

Many of these technologies have various limitations including high energy requirements (mechanical processing, pyrolysis), high costs (acid hydrolysis, ozonolysis), long hydrolysis times (alkaline hydrolysis, biological hydrolysis), inefficient lignin disruption (steam and ammonium fiber explosion), and production of toxic/inhibitory compounds (steam explosion, acid hydrolysis) (Kumar et al. 2009). Therefore, pre-treatment strategies must be selected accordingly depending on downstream processing requirements. The use of cellulolytic and hemicellulolytic enzyme cocktails and/or bacteria capable of cellulose and hemicellulose breakdown negates or limits pre-treatment of hemicelluloses.

1.4.3 Enzymatic cellulose and hemicellulose hydrolysis

Over 70 different species of bacteria have been isolated that are capable of hydrolyzing crystalline cellulose (for a complete list see http://www.wzw.tum.de/mbiotec/cellmo.htm (Schwarz 2003)). They belong to 6 of the 23 phyla of the domain bacteria (Thermotogae, Proteobacteria, Firmicutes, Actinobacteria, Fibrobacteres and Bacteroidetes) with the majority belonging to the phylum Firmicutes. These cellulolytic microorganisms have been predominantly isolated from gastrointestinal tracts of ruminant animals, manure, soil, sewage, manure, compost, and hot springs. A number of yeasts, including soft, white, and brown rot fungi, capable

of cellulose degradation have also been isolated (Dennis 1972; Stevens and Payne 1977; Nakase et al. 1994; Kuhad et al. 2011).

Cellulolytic microorganisms posses a number of different proteins that allow for efficient degradation of cellulose. Major components of any cellulase system include (i) secreted endo-β-glucanases, which cleave internal amorphous regions of the cellulose chain into shorter soluble oligosaccharides, (ii) secreted exo-β-glucanases, including cellodextrinases, cellobiohydrolases, and glucanohydrolases, which act in a processive manner on reducing or non-reducing ends of the cellulose chain liberating shorter cellodextrins, D-cellobiose, or D-glucose, respectively, (iii) intracellular cellobiose phosphorylases, which break down cellobiose into glucose & glucose-1-P, (iv) cellodextrin phosphorylases, which phosphorylyze β-1,4-oligoglucans, and (v) β-glucosidases, which hydrolyze cellobiose into glucose (Lynd et al. 2002). Other glycosyl hydrolases that allow hydrolysis of hemicellulose include xylanases, lichenases, laminarinases, β-xylosidases, β-galactosidases, and β-mannosidases, while pectin processing is accomplished via pectin lyase, polygalacturonate hydrolase, and pectin methylesterase (Spinnler et al. 1986; Demain et al. 2005). Few, if any, of these cellulolytic microorganisms are capable of breaking down lignin, and therefore delignification of lignocellulose increases accessibility of cellulose and hemicelluloses to hydrolytic enzymes.

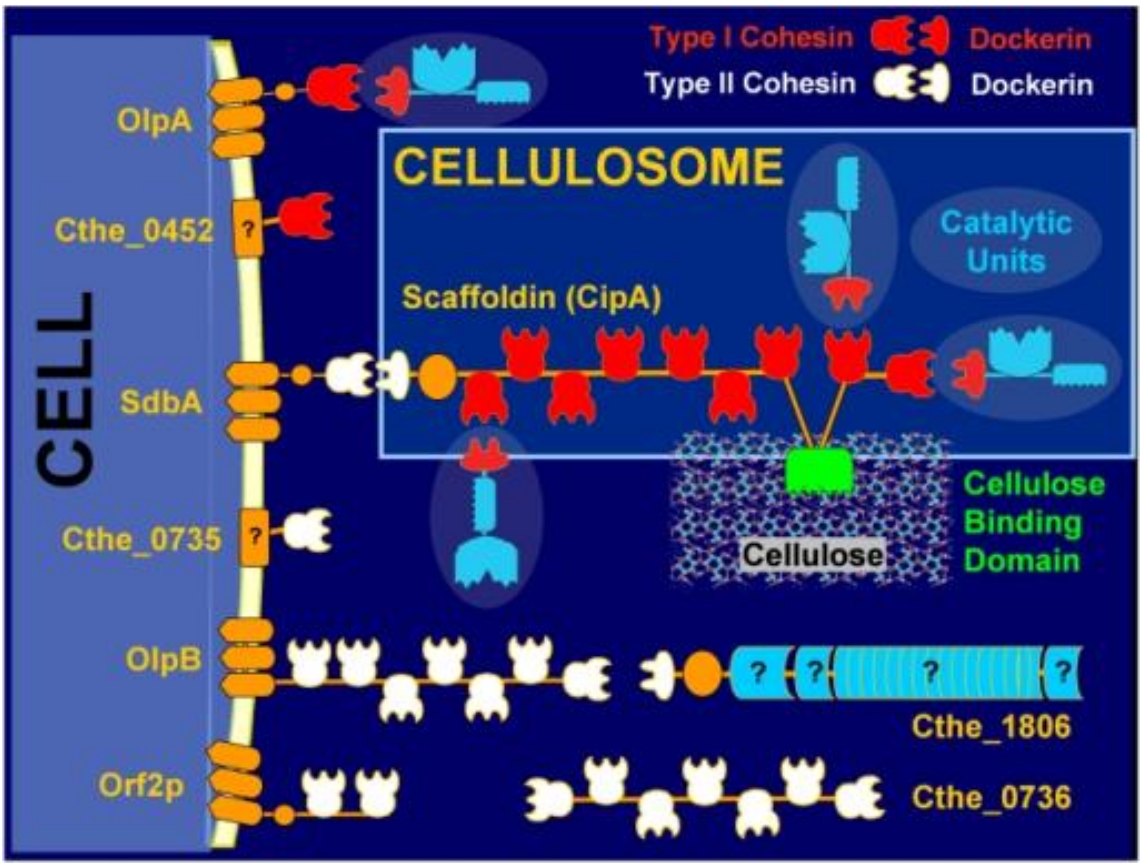
Cellulase systems may exist as non-complexed enzyme systems, during which a microorganism secretes cell-free cellulases into the surrounding environment, or as cell bound multi-enzyme complexes known as ‘cellulosomes’ in which a number of cellulases are attached to the cell surface via a scaffoldin-like protein. Ecologically, complexed

cellulase systems are advantageous for cellulolytic bacteria, which coexist in consortia with other cellulolytic and non-cellulolytic organisms, since direct attachment of the cell-bound cellulosome to cellulose fibers minimizes diffusion of soluble cello-oligosaccharides and allows for efficient uptake of these soluble sugars into the cell, thus allowing for efficient competition with surrounding organisms. Furthermore, since cellulase systems are synergistic, meaning that the collective activity of cellulases is higher than the sum of their individual activities, close proximity of different types of cellulases mediated by attachment of cellulases to the scaffoldin allows for concerted enzyme activity and optimal synergism (Lynd et al. 2002).

Perhaps the best studied cellulosome is that of *C. thermocellum*, which has been used as a model to study the cellulosomes in other cellulolytic bacterial (Figure 1.1). *C. thermocellum* has the highest known rates of cellulose degradation due to expression of an extensive suite of cellulases and hemicellulases (Schwarz 2001). Its cellulosomes can be seen microscopically as protuberances produced on the cell wall when grown in the presence of cellulose. The non-catalytic portion of *C. thermocellum*'s cellulosome, which complexes glycosyl hydrolases to the cell exterior, consists of a 1850 amino acid long, glycosylated, cellulosome-integrating protein (CipA) known as the 'scaffoldin' (Demain et al. 2005). This scaffoldin contains (i) a cellulose binding motifs (CBM) allowing for the binding of the scaffoldin to the cellulose fiber, (ii) nine type I cohesion domains with a 60-100% identity that mediate binding of various glycosyl hydrolases via their type I dockerin domains, and (iii) a type II dockerin domain that mediates binding to the type II cohesion domain found on cell-surface anchoring proteins. The cell-surface anchoring proteins (are in turn non-covalently bound to the peptidoglycan cell wall via C-

terminal surface-layer homology (SLH) repeats (Demain et al. 2005). Four different proteins containing SLH domains and type II cohesion domains (SdbA, OlpB, Orf2p, Cthe__0736) necessary for binding of cipA have been identified in *C. thermocellum* (Raman et al. 2009). Furthermore, two proteins (OlpA, Cthe__0452) that contain SLH domains and type I cohesion domains, which allow for direct binding of type I dockerin-containing glycosyl hydrolases, have also been identified (Raman et al. 2009). Interestingly, Cthe__0736 contains type II cohesion domains but no SLH domain, and thus may coordinate glycosyl hydrolases in the supernatant.

Figure 1.1: Cellulosome components of *C. thermocellum* The scaffoldin (CipA) can attach to the cell surface anchoring proteins (SdbA, Cthe_0735, OlpB, Orf2p) via a type II dockerin. CipA also contains a cellulose binding domain and 9 type I cohesions that can bind various glycosidic catalytic units via their type I dockerins. Several other proteins, including 2 surface anchoring proteins (OlpA, Cthe_0452) containing type I cohesions, a protein containing type II cohesions but no SLH domain (Cthe_0736), and a protein containing a type II dockerin, have also been identified. Source: (Raman et al. 2009).



Several other cellulolytic bacteria, including *Ruminococcus* (*flavefaciens* and *albus*), *Clostridium* (*acetobutylicum, cellulovorans, cellobioparum, papyrosolvens, josui, cellulolyticum*), *Bacteroides* (P-1, *cellulosolvens*), *Acetivibrio* (*cellulolyticus, cellulosolvens*), *Thermobifida* (*fucosa*), and *Fibobacter* (*succinogenes*) species have also been shown to poses a cellulosome based on either presence of dockerin and cohesion sequences, biochemical evidence of glycosidase-scaffoldin complexes, reaction with anti-scaffoldin antibodies, or presence of cell protuberances using electron microscopy (Schwarz 2003).

1.4.4 Consolidated bioprocessing for second-generation biofuels

Conventional production of 2nd generation lignocellulosic biofuels requires separate hydrolysis and fermentation (SHF), a muti-step process involving (i) cellulase production, (ii) hydrolysis of cellulose and hemicellulose, (iii) hexose (glucose, mannose, galactose) fermentation, and (iv) pentose (xylose, arabinose) fermentation. More recent strategies for cellulosic biofuel production, which include simultaneous saccharification and fermentation (SSF), combining cellulose and/or hemicellulose hydrolysis with hexose fermentation, and simultaneous saccharification and co-fermentation (SSCF), combining cellulose and/or hemicellulose hydrolysis with hexose and pentose fermentation, still require the exogenous production of cellulolytic enzymes for cellulose hydrolysis (Lynd et al. 2002).

Consolidated bioprocessing (CBP), which features simultaneous cellulase production, cellulose hydrolysis, and fermentation of hexose and pentose sugars is an attractive alternative processing strategy that lowers biofuel production costs. It offers

simpler feedstock processing, lower energy inputs (and therefore better energy balances), higher hydrolysis rates and conversion efficiencies, improved yields, and lower capital investment costs compared to SCF and SSCF based processes (Lynd et al. 2005). Lynd *et al*. project a more than fourfold higher ethanol production cost (18.9 vs. 4.2 cents/gallon) for biological processes associated with combined cellulase production and SSCF than for CBP (Lynd et al. 2005).

Although CBP offers tremendous potential for cellulosic biofuel production, microorganisms capable of efficient cellulose degradation and utilization combined with high production yields of desired products are currently not available for CBP. Successful development of such organisms requires a good understanding of the processes involved in cellulose degradation and the metabolic pathways that dictate product yields. Current metabolic engineering strategies aim to express heterologous cellulase systems in organisms with good biofuel yields or to improve end-product yields in cellulolytic bacteria (Lynd et al. 2005). While over 70 species of cellulolytic bacterial have been isolated, a smaller portion of these bacteria are capable of fermentative biofuels production, and thus CBP. Examples of cellulolytic anaerobic bacteria capable of ethanol and/or H_2 production include *Acetivibrio cellulolyticus*, *Butyrivibrio fibrisolvens*, and various *Ruminococcus, Caldicellulosiruptor* and *Clostridium* species (Iannotti et al. 1973; Patel et al. 1982; Diez-Gonzalez et al. 1999; Julliand et al. 1999). *Clostridium thermocellum* has perhaps been one of the best studied biofuels-producing, cellulolytic organisms.

1.5 Fermentative pathways of biofuel-producing cellulolytic bacteria

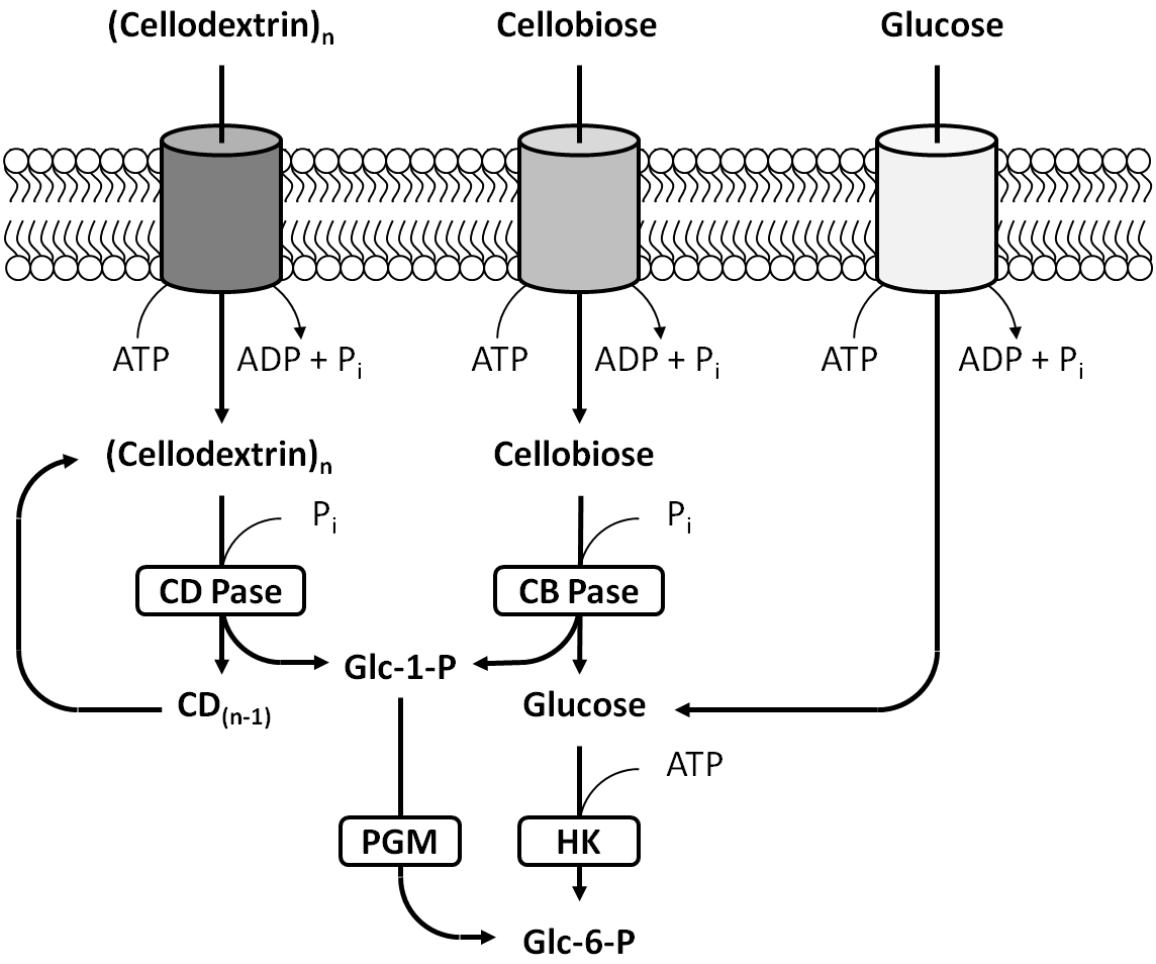
Cellulolytic/hemicellulolytic organisms provide one of the best potentials for biofuels production via consolidated bioprocessing. While environmental benefits and process economics are favorable over SSF and SSCF, low biofuels yields are a key issue that must be addressed. Understanding how metabolic pathways dictate carbon and electron flux is critical in developing educated metabolic engineering strategies and manipulating fermentation conditions. Combustion, by definition, is the oxidation of a fuel (or reduced substance) using an oxidizing element (usually oxygen). Therefore, electron transfer reactions are an underlying theme in production of biofuels. Here, a broad spectrum of key enzymes involved in conversion of cellobiose/cellodextrins to end products in select cellulolytic, biofuels producing bacteria is reviewed. Special emphasis is placed on energy conserving reactions (primarily substrate level phosphorylation) required for bacterial growth and electron transfer reactions that may dictate biofuels yields.

1.5.1 Oligosaccharide entry and breakdown

With the exception of *Clostridium saccharolyticus* and *C. cellulovorans*, which encode bacterial phosphotransferase systems (PTS), most *Caldicellulosiruptor* and *Clostridium* species import sugars via ATP-binding cassette (ABC) transport systems. Following the cellulosome-mediated breakdown of cellulose, glucose, cellobiose, and cellodextrins (with a degree of polymerization ≤ 5) are taken up by *C. thermocellum* via sugar-specific ATP-dependent ABC transport systems. Upon entry, glucose may be directly converted to glucose-6-phosphate via hexokinase, whereas cellobiose and

cellodextrins are cleaved via intracellular cellobiose and cellodextrin phosphorylase, respectively, producing either glucose or shorter cellodextrins and glucose-1-phosphate. Glucose-1-phosphate can subsequently be converted into glucose-6-phosphate via phosphoglucomutase, while glucose is again converted to glucose-6-phosphate via hexokinase (Figure 1.2). It is important to note that phosphorolytic cleavage using cellodextrin and/or cellobiose phosphorylases occurs at a rate 20 times faster than that of hydrolytic cleavage via β-glucosidases (Demain et al. 2005), and is also more energetically beneficial as it decreases the amount of ATP required for transport and glucose-6-phosphate production per glucose equivalent (Fig 1.2), (Strobel 1995; Strobel et al. 1995; Zhang and Lynd 2005a). This supports *C. thermocellum* studies reporting higher growth yields and lower maintenance energy coefficients during growth on cellobiose when compared to glucose (Ng and Zeikus 1982), and even higher on cellotetraose (Strobel et al. 1995). Zhang and Lynd reported that the mean cellodextrin degree of polymerization assimilated of cellobiose grown cells was 4.2 (Zhang and Lynd 2005a). The advantage of reduced ATP requirements for import and metabolism of longer chained cellodextrins offsets energy requirements of cellulase and cellulosome production.

Figure 1.2: Schematic model of carbohydrate uptake and phosphorylation by *C. thermocellum*. Note that uptake of cellobiose and cellodextrins uses a different transport system than is used for glucose uptake and that glucose uptake is energetically more costly. G-1-P, glucose-1-phosphate; G-6-P, glucose-6-phosphate; Pi, inorganic phosphate; CB Pase, cellobiose phosphorylase; CD Pase, cellodextrin phosphorylase; HK, hexokinase; PGM, phosphoglucomutase. (Adapted from Strobel *et al*. (Strobel et al. 1995)).



1.5.2 Conversion of glucose to phosphoenolpyruvate

Both ethanol and H_2 are reduced end products, with a redox value (calculated as the number of oxygen atoms less one-half the number of hydrogens in each compound) of -2 and -1, respectively (Moat *et al*. 2002). Therefore their production requires reducing equivalents in the form of NAD(P)H and/or reduced ferredoxin. Reducing equivalents are, in part, produced during the conversion of glucose to pyruvate via the Embden-Meyerhof-Parnas (EMP) pathway, which produces a net yield of 2 ATP and 2 NADH per glucose, or via the Entner-Doudoroff pathway, which produces a net yield of 1 ATP, 1 NADH, and 1 NADPH per glucose. Several minor variations of the EMP pathway exist in certain fermentative organisms. For example, genomic analysis reveals that certain *Thermotoga*, *Caldicellulosiruptor*, *Thermoanaerobacter*, and *Clostridium* species can carry out glycolysis using both an ATP-dependent or a pyrophosphate (PP_i)-dependent phosphofructokinase (Carere et al. 2012). Alternatively, several H_2-producing Archaea, including *Pyrococcus furiousus*, *Thermococcus celer*, and *T. kodakaraensis*, use glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPFOR) rather than glyceraldehydes-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK) to convert glyceraldehyde-3-phosphate to 3-phosphoglycerate. GAPFOR reduces ferredoxin rather than NAD^+ and does not produce ATP. Given that reduced ferredoxin is more reduced than NAD(P)H, the presence of GAPFOR facilitates improved thermodynamic efficiencies of H_2 production (Sparling et al. 2012).

1.5.3 Inter-conversion of phosphoenolpyruvate and pyruvate

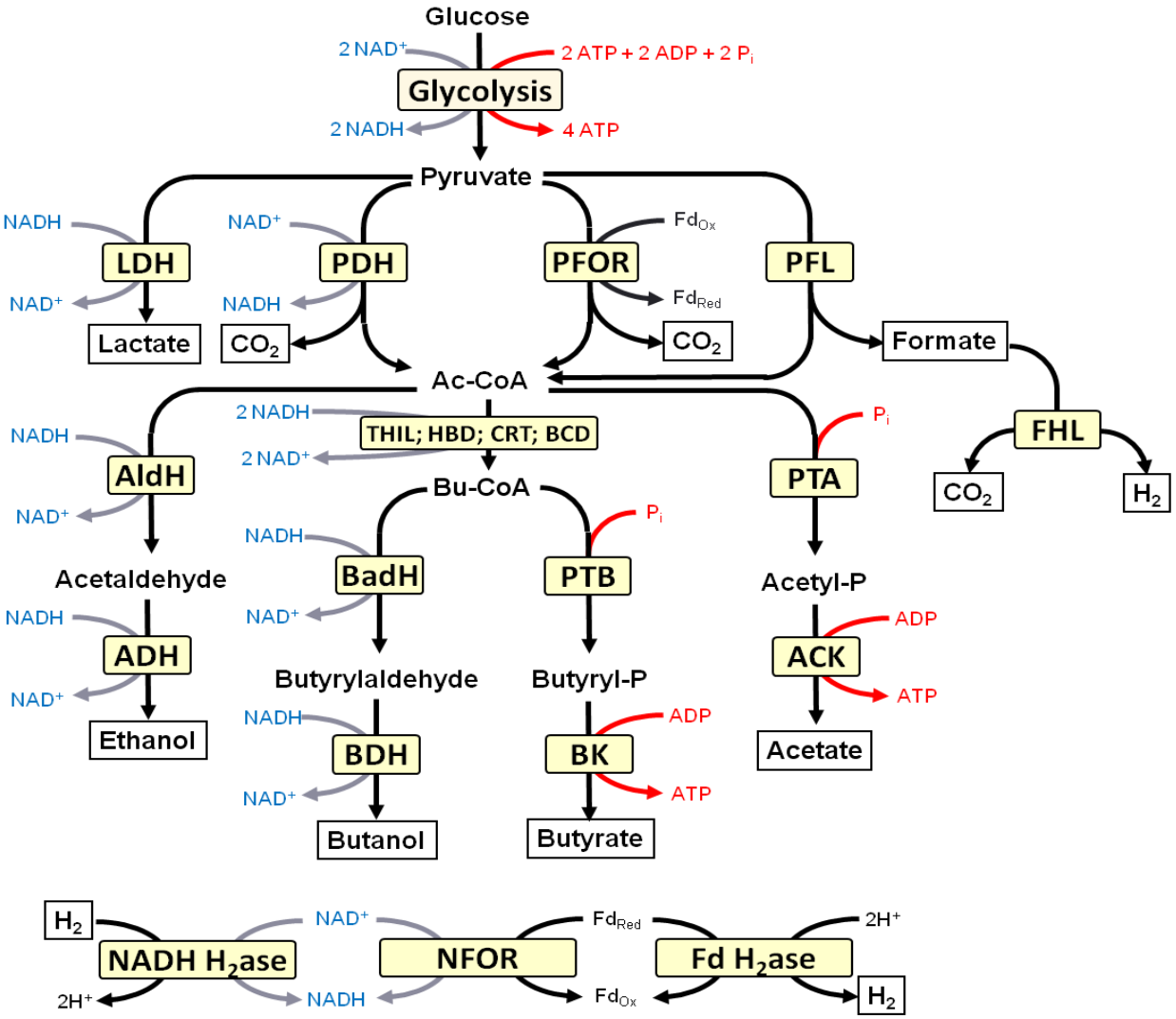
Typically, conversion of phosphoenolpyruvate (PEP) to pyruvate, which yields an ATP from ADP in the process, is catalyzed by pyruvate kinase. However, some organisms such as *C. thermocellum*, do not encode a pyruvate kinase and thus this reaction must be carried out by other means. Phosphoenolpyruvate can potentially be reversibly converted to pyruvate via pyruvate phosphate dikinase (PPDK), producing ATP and Pi from AMP, and PPi, or using PEP synthase (PEPS) which produces ATP and H_2O from AMP, and Pi. Alternatively, PEP carboxykinase (PEPCK), can convert PEP to oxaloacetate while generating ATP. Oxaloacetate can subsequently be converted either directly to pyruvate via oxaloacetate decarboxylase (OAADC), or indirectly through malate via malate dehydrogenase (MDH) and malic enzyme (MalE), the latter of which has been isolated and characterized (Lamed and Zeikus 1981). It has been proposed that MDH and MalE used in this 'malate shunt' can facilitate transhydrogenation from NADH to $NADP^+$ (Lamed and Zeikus 1980a). Given the absence of glucose-6-P dehydrogenase, gluconolactonase, and 6-Pgluconate dehydrogenase, which compose the NADPH-producing oxidative branch of the pentose phosphate pathway, the malate shunt may be a key pathway required for generation of NADPH for biosynthetic pathways. Interestingly, a malate dehydrogenase deletion in a *C. thermocellum* strain expressing an exogenous pyruvate kinase had little impact on growth and ethanol yields when grown on cellobiose (Deng et al. 2013) suggesting that additional pathways can be used to produce NADPH. Furthermore, while a lone malate dehydrogenase or lone lactate dehydrogenase deletion in *C. cellulolyticum* increased ethanol production only marginally (~1.3-fold and ~1.5-fold, respectively), the increase in ethanol production resulting from a double deletion of lactate dehydrogenase and malate dehydrogenase increased ethanol production on

cellobiose by 2.4-fold (Li et al. 2012b). Furthermore, while the lactate dehydrogenase mutant had little effect of on final growth yields on cellobiose, the malate dehydrogenase mutant decreased final growth yields (Li et al. 2012b).

1.5.4 Fermentative metabolism of pyruvate to end-products

The catabolism of pyruvate into end-products can be represented by several major branch points (Figure 1.3). The first is the *Pyruvate/Lactate/Acetyl-CoA* branch point where pyruvate can be catalyzed into either (i) lactate via lactate dehydrogenase (LDH), (ii) NADH, CO_2, and acetyl-CoA via pyruvate dehydrogenase (PDH), (iii) reduced ferredoxin (Fd), CO_2, and acetyl-CoA via pyruvate:ferredoxin oxidoreductase (PFOR), or (iv) formate and acetyl-CoA via pyruvate:formate lyase (PFL). The enzymes used to convert pyruvate into acetyl-CoA may have significant implications on biofuels yields. For example, the production of lactate via LDH requires the oxidation of NADH, which therefore reduces the availability of NADH for production of reduced end-products such as ethanol, butanol, butyrate and/or H_2. While conversion of pyruvate to acetyl-CoA via PFL does not generate additional reducing equivalents and thus has a negligible effect on reduced end-product yields at this branchpoint, formate can potentially be degraded to H_2 and CO_2 via formate:hydrogen lyase (FHL). Alternatively, PDH and PFOR produce additional NADH and reduced ferredoxin, respectively, allowing for the production of additional reduced end-products. The impact of these enzymes is discussed more thoroughly throughout this thesis.

Figure 1.3: Typical fermentation pathways. LDH, lactate dehydrogenase; PDH, pyruvate dehydrogenase; PFOR, pyruvate:ferredoxin oxidoreductase; PFL, pyruvate:formate lyase; AldH, aldehyde dehydrogenase; ADH, alcohol dehydrogenase; THIL, acetyl-CoA acetyltransferase; HBD, β-hydroxybutyryl-CoA dehydrogenase; CRT, 3-hydroxybutyryl-CoA dehydratase; BDC, butyryl-CoA dehydrogenase; BadH, butyrylaldehyde dehydrogenase; BDH, butanol dehydrogenase; PTB, phosphotransbutyrylase; BK butyrate kinase; PTA, phosphotransacetylase; ACK, acetate kinase; PFL, pyruvate:formate lyase; NADH H_2ase, NADH-dependent hydrogenase; NFOR, NADH:ferredoxin oxidoreductase; Fd H_2ase, ferredoxin-dependent hydrogenase.


Glucose
2 NAD+
2 ATP + 2 ADP + 2 Pi
Glycolysis
2 NADH
4 ATP
Pyruvate
NADH
LDH
NAD+
Lactate
NAD+
PDH
NADH
CO2
Fdox
PFOR
FdRed
CO2
PFL
Formate
FHL
CO2
H2
Ac-CoA
2 NADH
THIL; HBD; CRT; BCD
2 NAD+
Bu-CoA
NADH
AldH
NAD+
Acetaldehyde
NADH
ADH
NAD+
Ethanol
NADH
BadH
NAD+
Butyrylaldehyde
NADH
BDH
NAD+
Butanol
Pi
PTB
Butyryl-P
ADP
BK
ATP
Butyrate
Pi
PTA
Acetyl-P
ADP
ACK
ATP
Acetate
H2
NADH H2ase
2H+
NAD+
NADH
NFOR
FdRed
FdOx
Fd H2ase
2H+
H2

At the *Acetyl-CoA/Ethanol/Acetate* branchpoint acetyl-CoA can be catalyzed into (i) acetate using either phosphotransacetylase (PTA) and acetate kinase (ACK), or acetate thiokinase (ATK), or into (ii) ethanol using either acetaldehyde dehydrogenase (AldH) and alcohol dehydrogenase (ADH), or a via bifunctional acetylaldehyde/alcohol dehydrogenase (AdhE). The substrate specificity (aldehyde length and substitution), coenzyme specificity (NADH vs. NADPH), and catalytic directionality (ethanol formation vs. consumption) of ADHs is difficult to determine based on gene sequences alone and is still being actively investigated (Loubiere et al. 1992; Burdette and Zeikus 1994; Liu et al. 2009; Pei et al. 2010). A number of organisms can also convert 2 acetyl-CoA to butyryl-CoA using acetyl-CoA acetyltransferase (THIL), β-hydroxybutyryl-CoA dehydrogenase (HBD), 3-hydroxybutyryl-CoA dehydratase (CRT), butyryl-CoA dehydrogenase (BCD). This requires oxidation of 2 NADH, diverting reducing equivalents from ethanol and/or H_2. At the *Butyryl-CoA/Butanol/Butyrate* branchpoint, an additional 2 NADH can be further used to produce butanol via butyrylaldehyde dehydrogenase (BadH) and butanol dehydrogenase (BDH), whose specificity may also be flexible. Alternatively, butyryl-CoA can be converted to butyrate using phosphotransbutyrylase (PTB) and butyrate kinase (BK), which produces ATP. Organisms such as *C. acetobutylicum* can also produce an array of other end-products such as succinate, acetoin, butanediol, and acetone.

1.5.5 Hydrogenases: the molecular basis for biohydrogen production

Although hydrogenases (H_2ases) are redox metalloenzymes that catalyze one of the simplest reversible chemical reactions ($2H^+ + 2e^- \leftrightarrow H_2$), they are highly complex and

diverse. The core enzymes can contain different types of oxygen liable metal compositions in their diatomic active sites, can have different modular structures and Fe-S cluster accessory domains that carry electrons towards and away from the active site, can be monomeric or associate with accessory subunits to form multimeric enzymes, and require a number of auxiliary enzymes for metallo-cluster insertion and protein maturation (Vignais 2001; Hallenbeck 2002; Calusinska et al. 2010). The incredible diversity of H_2ases has been extensively reviewed by Vignais *et al*., Calusinska *et al*., and Meyer (Vignais 2001; Meyer 2007; Vignais 2008; Calusinska et al. 2010).

H_2ases play a central role in microbial energy metabolism. In fermentative bacteria such as clostridia, H_2ases are used to dispose of excess electrons through H_2 evolution, while in organisms such as methanogens, sulfate reducers, and photosynthetic bacteria H_2ases participate in H_2 uptake and allow for H_2 to be used as a source of electrons (Vignais 2001). Alternatively, in *Bradyrizobium japonicum*, *Rhodobacter capsulatus*, and *Ralstonia eutropha*, H_2ases can act as H_2 sensors whereby they regulate H_2ase gene expression in response to H_2 levels through a two-component regulatory system (Elsen et al. 2000; Kleihues et al. 2000).

H_2ase directionality and redox partners may be dictated by their location, modular structure, and accessory domains/subunits. They may be either cytoplasmic, periplasmic, or membrane bound. All periplasmic/membrane-localized H_2ases contain a 30-70 amino acid signal peptide containing a conserved RRxFxK sequence which is recognized by the twin arginine translocation (Tat) or membrane targeting and translocating (Mtt) pathways (Vignais *et al*., 2001). While H_2ase function was initially believed to be dependent on cellular localization (i.e. cystolic H_2ases were involved in H_2 evolution while

periplasmic/membrane-bound H_2ases were involved in H_2 uptake) (Vignais 2001), this is not necessarily true. For example, the membrane bound H_2ase (Mbh) of *Pyrococcus furiosus* or the membrane bound energy conserving H_2ase (Ech) hydrogenase of *C. thermocellum* are believed to be involved in H_2 production.

H_2ases can shuttle electrons between H_2 and various electron carriers including NAD(P)H, ferredoxin (Fd), flavodoxin, coenzyme F_{420}, or cytochromes (Vignais 2001). The type of electron carrier they use as an electron donor/acceptor can have significant implications on fermentative metabolism, and in turn product yields. For example, although Fd-dependent H_2ases are not conducive for H_2 production under standard conditions ($\triangle G°' = +3$ kJ mol^{-1}), high intracellular Fd_{red}/Fd_{ox} ratios generated by PFOR drive H_2 produciton. Alternatively, NADH-dependent H_2ases can either produce H_2 from NADH, thus reducing available NADH for ethanol synthesis, or generate NADH from H_2 allowing for increased ethanol production. While the standard free energy of H_2 production from NADH is unfavorable ($\triangle G°' = +15$ kJ mol^{-1}), intracellular levels of NADH can have a significant impact on *in vivo* thermodynamics and thus H_2ase directionality. Elucidation of the electron donors/acceptors used by H_2ases based on their modular structure and accessory subunits is important for understanding overall fermentative metabolism.

H_2ases were initially classified based on the electron donors/acceptors they used, namely NADH (EC1.12.1.12), ferredoxin (EC 1.18.99.1), coenzyme F_{420} (EC 1.12.99.1), or cytochromes (EC 1.12.2.1). Today, H_2ases are classified based on active site metal composition and include [NiFe]-H_2ases, [FeFe]-H_2ases, and [Fe]- H_2ases (Vignais 2001). The majority of H_2ases fall into [NiFe]- or [FeFe]-H_2ases, whereas [Fe]-H_2ases have only

been described in methanogenic archaea (Pilak et al, 2006). These metalloclusters are bound to the polypeptide chain via cysteine residues and contain a unique arrangement of CO and CN ligands (Evans and Pickett 2003). H_2ase clusters differ not only by their metal composition but also by how they are attached to the cysteine-containing protein moieties. The conserved sequence signatures encompassing the cysteine residues that bind the metal groups allow for distinguishing between modular structures of different H_2ases (Vignais 2001; Meyer 2007; Calusinska et al. 2010).

1.5.5.1 *[NiFe] hydrogenases*

The core [NiFe]-H_2ase is a heterodimeric protein consisting of a large (α) and small (β) subunit. The large subunit contains the active [NiFe] center bound by 4 cysteine residues and is coordinated to three ligands, one CO and two cyanide (CN^-) molecules (Pierik et al. 1999; Meyer 2007). The conserved N- and C-terminus sequences containing the two pairs of cysteine residues of the NiFe site (L1 and L2) can be used to classify [NiFe]-H_2ases. The small subunit contains two [4Fe-4S] and one [3Fe-4S] iron-sulfur clusters, and has been shown to be homologous to soluble ferredoxins and parts of NADH-ubiquinone oxidoreductase (Complex I) of the respiratory chain (Vignais 2001). It functions to transfer electrons from the electron carrier to the large catalytic subunit. The synthesis of this metallocenter requires a number of auxiliary proteins for metal capture, ligand synthesis, and cluster insertion, genes of which are often clustered with [NiFe]-H_2ase structural genes.

[NiFe]-H_2ases can be classified into four main groups based on sequence analysis of the L1 and L2 cysteine moieties (Vignais 2001; Vignais 2008). These include: (1)

bacterial and archaeal membrane-bound respiratory-uptake H_2ases and periplasmic soluble H_2ases of sulfate reducers; (2) cyanobacterial uptake H_2ases and cytoplasmic H_2 sensory H_2ases; (3) bidirectional cytoplasmic H_2ases that bind soluble cofactors (F_{420}, NADH, NADPH, methyl viologen); and (4) membrane bound, H_2 evolving energy converting (or conserving) H_2ases (Ech). While many members of the genus *Clostridium* cluster into group 1, *C. thermocellum* is not one of them (Calusinska et al. 2010). Instead, like several other clostridia, it encodes a six-subunit, membrane associated group 4 Ech-type H_2ase, similar to that of the methanogenic archaea *Methanosarcina barkeri*. This H_2ase is encoded by the *echABCDEF* operon, where EchE and EchF are the large (α) and small (β) subunits of the core [NiFe] H_2ase (putatively cytoplasmically located), EchA and EchB are internal membrane proteins, and EchC and EchD are small hydrophobic proteins (Calusinska et al. 2010). While there are no physiological data describing the physiological function or directionality of Ech-type H_2ases in clostridia, it is predicted to use ferredoxin as an electron partner based on physiological studies of *M. barkeri* (Kurkin et al. 2002).

1.5.5.2 ***[FeFe] hydrogenases***

Fe-only H_2ases are highly diverse in their modular domains, and may have different amounts of FeS clusters. The core of the [FeFe] H_2ase consists of an H domain containing the H-cluster, which consists of a FeFe active site and a cysteine-bridged [4Fe4S] cluster surrounded by non-protein ligands (Peters et al. 1998; Calusinska et al. 2010). The three segments encompassing the four cysteine ligands of the metal site (L1Fe, L2Fe, L3Fe) are highly conserved among [FeFe]-H_2ases (Vignais 2001). Most

clostridial [FeFe]-H_2ases also have an additional N-terminal bacterial ferredoxin-like 2[4Fe4S] F-cluster accessory domain (Peters et al. 1998). The catalytic subunit may also contain numerous additional conserved accessory domains containing FeS-clusters (reviewed by Vignais *et al*., Calusinska *et al*., and Meyer) (Vignais 2001; Meyer 2007; Vignais 2008; Calusinska et al. 2010). Furthermore, the catalytic subunit can associate with one to four accessory subunits. The modular structures of [FeFe]-H_2ases are designated by a specific lettering and numbering system based on subunit composition (M, monomeric; D, dimeric, TR, trimeric; TE, tetrameric) and the number and types of accessory domains found in the catalytic subunit (indicated by an Arabic numeral and lower cased Latin letter) as outlined by Meyer and Calusinska *et al*. (Meyer 2007; Calusinska et al. 2010). Although little is known about [FeFe] H_2ase maturation and assembly factors, recent mutational studies of *Chlamydomonas reindhardtii* revealed 3 genes (*hydE, hydF, and hydG*) that may be involved in H_2ase maturation (Posewitz et al. 2004). Unlike [NiFe] maturation factors, these genes do not cluster with their corresponding structural [FeFe] H_2ase genes.

Clostridia often encode multiple [FeFe]-H_2ases, which are thought to help them respond to different environmental conditions. They can be monomeric or multimeric. The number and types of catalytic and accessory subunits can influence redox partner specificity. Calusinska *et al*. compared the sequences of over 200 putative [FeFe] H_2ase and classified them into four main groups (A-D) (Calusinska et al. 2010). Each of these groups also contain multiple sub-clusters. Group A contains the most diverse group of [FeFe]-H_2ases, with four different modular structures (M3, M2, M2d, and M2c) and different subunit compositions (M, D, TR). Monomeric [FeFe]-H_2ases of group 1

typically produce H_2 from ferredoxin (M3), as in *C. pasteruianum*. However, some of these H_2ases (M2d) contain a rubrerythrin domain, and are believed to reduce H_2O_2. Alternatively, some of these H_2ases (M2c) contain a soluble-ligand-binding β-grasp fold, which may allow for binding of FeS clusters involved in electron transport. Dimeric H_2ases (D(M2)) of group A contain accessory subunits containing a NuoF-like domain and an adjacent NADPH-dependent glutamate synthase domain, suggesting that it may use H_2 to produce NADPH for ammonium assimilation via catalysis of glutamate from glutamine and 2-oxogluterate. Heterotrimeric H_2ases of group A form either (i) bifurcating H_2ases (TR(M3)), whose accessory subunits contain NADH-binding domains suggesting that they use both NADH and reduced ferredoxin simultaneously to produce 2 H_2, or (ii) formate:hydrogen lyase-like complexes (TR(M2)), whose accessory subunit can connect to a formate-dehydrogenase-like protein allowing oxidation of formate to H_2 and CO_2 (Axley 1990). Recent characterization of a heterotrimeric "bifurcating" H_2ase from *Thermotoga maritma* has in fact demonstrated that it can simultaneously oxidize reduced Fd and NADH to 2 H_2 ($\triangle G°' = +10.6$ kJ mol^{-1} H_2 produced), which drives the highly endergonic production of H_2 from NADH by coupling it to the less endergonic oxidation of reduced Fd (Schut and Adams 2009). While this reaction appears to be unfavorable under standard conditions, high intracellular Fd_{red}/Fd_{ox} and $NADH/NAD^+$ ratios can help drive this reaction.

Monomeric H_2ases (M3a, M3b) of group B show very similar structure to M3 modular H_2ase structure, but the cystein motif that typically binds a plant ferredoxin-like domain is replaced by a series of six or eight cysteine residues. Alternatively, monomeric H_2ase (M2a) contains an additional cysteine residue a conserved motif of the H-cluster

and a modified [4Fe4S] binding sequence at the N-terminus similar to the *nif*-associated ferredoxin in *Rhizobium meliloti* required for symbiotic nitrogen fixation. Dimeric H_2ases of group C contain C-terminal PAS/PAC sensory domains and are often encoded next to trimeric H_2ases (TR(M3)). They are suggested to regulate these H_2ases in response to light or cellular redox through a signal transduction pathway. Monomeric H_2ases of group D have that have less conserved FeS cluster cysteine motifs that, are also found upstream of timeric H_2ases and are next to a histidine kinase.

1.5.6 Electron transfer between NADH and ferredoxin

While the reducing equivalents used for the production of alcohols (ethanol, butanol), acetate, and butyrate come strictly from NADH and/or NADPH, reducing equivalents, in the form of NADH, NADPH, as well as ferredoxin can be used for H_2 metabolism. Therefore, the transfer of electrons between NADH, NADPH, and ferredoxin can redirect electron (and carbon) fluxes among different end-products.

1.5.6.1 *NADH:ferredoxin oxidoreductase*

Electrons may be transferred directly between Fd and NAD(P)H via an Rnf-like (*Rhodobacter n*itrogen *f*ixation) NADH:ferredoxin oxidoreductase (NFOR), a membrane-bound enzyme complex capable of generating a sodium (or proton) motive force driven by the exergonic reduction of NAD^+ via reduced ferredoxin ($\triangle G°' = -21$ kJ mol^{-1}). The sodium (or proton) motive force generated can, in turn, be used to generate ATP via ATP synthases. The Rnf complex, first discovered in the purple non-sulfur bacterium *Rhodobacter capsulatus,* consists of seven gene products (*rnfABCDEF*). Mutational analysis demonstrated it was required for nitrogen fixation (Schmehl et al. 1993). Bioinformatic analysis revealed subunit similarity to the six-subunit (Barquera et al. 2002) Na^+-translocating NADH:quinone-oxidoreductase (Nrq) (Schmidt et al. 2009) first identified in *Vibrio alginolyticus* (Tokuda et al. 1981; Tokuda and Unemoto 1981) which reduces ubiquinone via NADH (Pfenninger-Li et al. 1996). A six-subunit Rnf-type NFOR has been identified in a number of fermentative bacteria, including *C. thermocellum* (See Chapter 2).

1.5.6.2 *NADH-dependent reduced ferredoxin:$NADP^+$ oxidoreductase*

Recently (in October, 2010) a heterodimeric, cytoplasmic *N*ADH-dependent reduced *f*erredoxin:*N*ADP^+ oxidoreductase (NfnAB) complex was identified in *Clostridium kluyveri* (Wang et al. 2010). This enzyme simultaneously uses a reduced Fd and an NADH to reduce 2 $NADP^+$ ($\triangle G°' = -20$ kJ mol^{-1}), which is driven by the exergonic reduction of $NADP^+$ via reduced ferredoxin. Given that typical *in vivo* $NADH/NAD^+$ ratios are around 0.3, and $NADPH/NADP^+$ ratios are generally above 1

(Pfitzer and Decker 1972; Thauer et al. 1977), the free energy of this coupled reaction may allow for production of NADPH from NADH, providing subsequent reduced ferredoxin levels.

1.6 Fermentation thermodynamics

Most fermentative microorganisms have branched product pathways, resulting in a diverse set of end products. Fluxes through these pathways, and in turn product yields, may, in part, be dictated by thermodynamics. The Gibbs free energies ($\triangle G°$) of common fermentative reactions, at different pHs and temperatures, are provided in table 1.2 (Sparling et al. 2012). Free energy values in Table 1.2 do not take into account the different amounts of ATP produced via associated pathways. The phosphorylation of each mole of ADP to ATP requires 31 kJ mol^{-1}.

Table 1.2: Calculated free energy changes (kJ mol^{-1}) per balanced reaction showing possible combinations of end products at different temperatures and pH. (Source:(Sparling et al. 2012))

Reaction (ATP produced)	**pH**	**Temperature (^oC)**						
		25	**37**	**45**	**60**	**70**	**80**	**100**
Glucose to H_2	7	-5	-28	-43	-72	-91	-110	-148
$C_6H_{12}O_6 + 12H_2O \rightarrow 12H_2 + 6HCO_3^- + 6H^+$	6	29	8	-6	-33	-51	-69	-105
	5	64	44	30	5	-12	-29	-62
Glucose to H_2/acetate (+4ATP):	7	-213	-224	-232	-246	-255	-265	-283
$C_6H_{12}O_6 + 4H_2O \rightarrow 4H_2 + 2CH_3COO^- + 2HCO_3^- + 4H^+$	6	-190	-201	-208	-220	-229	-238	-255
	5	-168	-177	-183	-195	-203	-211	-226
Glucose to H_2/butyrate (+3ATP):	7	-263	-271	-276	-285	-292	-298	-311
$C_6H_{12}O_6 + 2H_2O \rightarrow 2H_2 + CH_3CH_2CH_2COO^- + 2HCO_3^- + 3H^+$	6	-246	-253	-257	-266	-272	-278	-290
	5	-229	-235	-239	-247	-252	-258	-268
Glucose to H_2/ethanol/acetate (+3 ATP):	7	-216	-224	-229	-240	-247	-253	-267
$C_6H_{12}O_6 + 3H_2O \rightarrow 2H_2 + CH_3CH_2OH + CH_3COO^- + 2HCO_3^- + 3H^+$	6	-199	-206	-211	-221	-227	-233	-246
	5	-182	-188	-193	-201	-207	-213	-224
Glucose to H_2/ethanol/formate/acetate (+3 ATP):	7	-217	-224	-229	-239	-245	-251	-263
$C_6H_{12}O_6 + 2H_2O \rightarrow H_2 + CH_3CH_2OH + HCOO^- + CH_3COO^- + HCO_3^- + 3H^+$	6	-200	-207	-211	-219	-225	-231	-242
	5	-183	-189	-193	-200	-205	-210	-220
Glucose to H_2/acetate/formate (+3 ATP):	7	-216	-225	-232	-244	-252	-260	-276
$C_6H_{12}O_6 + 2H_2O \rightarrow 2H_2 + 2CH_3COO^- + 2HCOO^- + 4H^+$	6	-193	-201	-207	-218	-225	-233	-247
	5	-170	-178	-183	-193	-199	-205	-218
Glucose to ethanol (+2 ATP):	7	-218	-224	-227	-234	-238	-242	-251
$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3CH_2OH + 2HCO_3^- + 2H^+$	6	-207	-212	-215	-221	-225	-229	-237
	5	-196	-200	-203	-208	-212	-215	-222

Glucose to lactate (+2 ATP):	7	-195	-199	-201	-206	-209	-212	-219
$C_6H_{12}O_6 \rightarrow 2CH_3CH(OH)COO^- + 2H^+$	6	-183	-187	-189	-193	-196	-199	-205
	5	-172	-175	-177	-181	-183	-185	-190
Glucose to ethanol/formate/acetate (+3 ATP):	7	-218	-225	-229	-237	-243	-248	-259
$C_6H_{12}O_6 + H_2O \rightarrow CH_3CH_2OH + 2HCOO^- +$	6	-201	-207	-211	-218	-223	-228	-238
$CH_3COO^- + 3H^+$	5	-184	-189	-193	-199	-203	-208	-217
Acetate to H_2	7	104	98	94	87	82	77	68
$CH_3COO^- + 4H_2O \rightarrow 4H_2 + 2HCO_3 + H^+$	6	110	104	101	93	89	84	75
	5	116	110	107	100	95	91	82

One of the main constraints reducing the economic viability of biohydrogen production is low fermentative H_2 yields. While complete oxidation of glucose can theoretically produce 12 moles of H_2 (and 6 moles of CO_2) per mole of glucose, the standard free energy of this reaction ($\Delta G^{o\prime}$) is near zero, and thus ATP production required for cell growth and maintenance cannot be sustained. In fermentative systems, a maximum of 4 moles H_2 can be produced when acetate and CO_2 are concomitantly produced ($\Delta G^{o\prime}$= -213 kJ mol^{-1}). Commonly, however, H_2 yields are 2 moles per mole glucose, or less, depending on what other end-products are concomitantly produced. The $\Delta G^{o\prime}$ of each of these reactions ranges from -200 to -220 kJ mol^{-1} (Table 1.2). Therefore, there appears to be very little thermodynamic advantage of producing, for example, 2 mol ethanol + 2 mol CO_2 ($\Delta G^{o\prime}$= -218 kJ mol^{-1}) over 4 mol H_2 + 2 mol CO_2 + 2 mol acetate. However, from an energy conservation standpoint, production of H_2 and acetate is advantageous over production of ethanol as it allows for the formation of 2 additional ATPs. Because of the relationship between ΔG^{o} and temperature, thermophiles and

especially hyperthermophiles are more likely to approach yields of 4 mol H_2 per mol glucose.

Most fermentative organisms cannot overcome the thermodynamic barrier of acetate oxidation to 4 mol H_2 + 2 CO_2 ($\Delta G^{o'}$= 104 kJ mol^{-1}) in monocultures. However, some syntrophic acetate-oxidizing bacteria can oxidize acetate to H_2 and CO_2 in the presence of hydrogenotrophic methanogens, which couple acetate oxidation to H_2-consuming methanogenesis ($4H_2 + HCO_3^{-1} + H^+ \rightarrow CH_4 + 3H_2O$; $\Delta G^{o'}$= -136 kJ mol^{-1}), thus maintaining extremely low H_2 partial pressures (20-80 Pa) (Luo et al. 2002; Hattori 2008). N_2-fixing photosynthetic bacteria (i.e. *Rhodospirillum*, *Rhodobacter*, and *Rhodopseudomonas* sp.) can also overcome this energetic barrier of acetate oxidations by harvesting light energy for the production of ATP to drive this reaction via photofermentation. In these organisms, the reaction itself is performed via a nitrogenase, during which H_2 is produced as a by-product of N_2 fixation ($N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$) (Rupprecht et al. 2006). Attempts to improve H_2 yields have therefore focused on two phase fermentation systems in which purple non-sulfur bacteria produce H_2 from effluents of dark fermentations (Lynd et al. 2002; Lee et al. 2010). Alternatively, microbial fuel cells can be used to produce H_2 from acetate. During this process, an external current supplied to a microorganism (ex. *Geobacter* species) to overcome the thermodynamic barriers of H_2 production from acetate. This process is capable of yielding 8-9 mol H_2 per mol glucose (Liu et al. 2005).

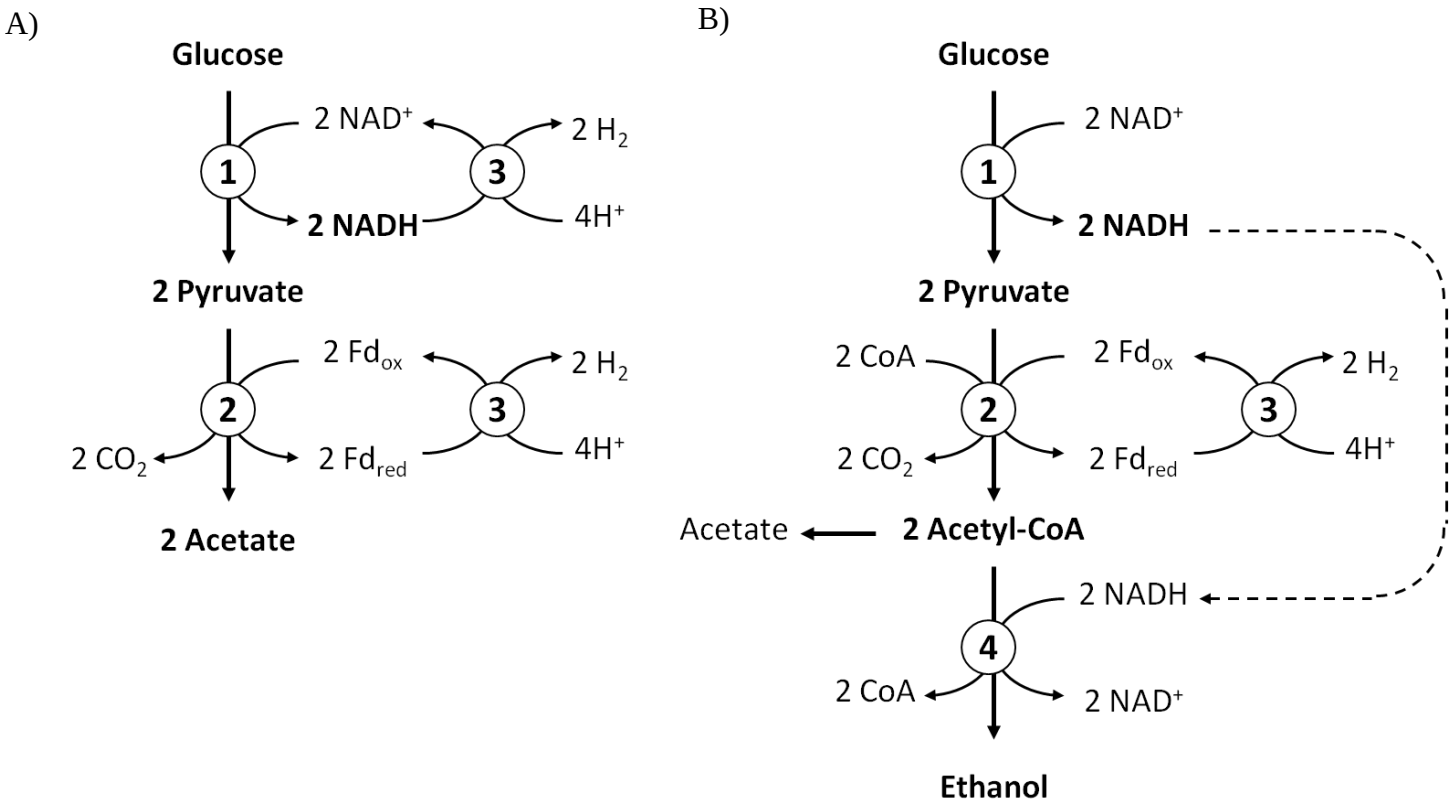
The thermodynamic efficiency of fermentative H_2 production is highly dependent on H_2 partial pressures (Angenent et al. 2004). Assuming that the intracellular

concentrations of oxidized and reduced forms of ferredoxin and NADH are roughly equal, H_2 production becomes unfavorable at H_2 partial pressures greater than:

$$P_{H_2,max} \leq \exp\left\{\frac{2F(E^{\circ\prime}_{H_2} - E^{\circ\prime}_{x})}{RT}\right\}$$

where $E^{\circ\prime}_{x}$ and $E^{\circ\prime}_{H_2}$ are the redox potentials of the electron donor (Fd or NAD(P)H) and acceptor (H^+) respectively, *F* is Faraday's constant, *R* is the ideal gas constant, and *T* is the absolute temperature (Angenent et al. 2004). Given that the redox potentials ($E^{o\prime}$) of H_2, ferredoxin, and NADH are -414 mV, -398 mV, and -320 mV respectively (Thauer et al. 1977), H_2 production can continue from ferredoxin at H_2 partial pressures less than ~0.3 atm, and from NADH at partial pressures less than ~6×10^{-4} atm (Figure 1.4). Thus, by decreasing H_2 partial pressures, increasing temperatures, or increasing the ratios of $Fd_{(red)}/Fd_{(ox)}$ or $NAD(P)H/NAD(P)^+$, H_2 production can be made more thermodynamically favorable, in turn increasing H_2 yields. H_2 partial pressures can be maintained at low levels by rapid sparging with nitrogen (Soboh et al. 2004), H_2 consumption by a hydrogenotroph (Thauer et al. 1977; Jackson et al. 1999), or removal by chemical means (Mountfort and Kaspar 1986) (discussed below).

Figure 1.4: Effect of hydrogen partial pressure on biological H_2 production. (a) H_2 production via NADH oxidation is thermodynamically favorable at H_2 partial pressures < 60 Pa whereas (b) at higher partial pressures, NADH is reoxidized during ethanol production. 1, glycolysis; 2 PFOR; 3, H_2ase; 4, AldH/ADH. Adapted from Angenent *et al.* (Angenent et al. 2004).



In a living cell, the standard equations do not take into account that steady state concentrations of cofactors are themselves not at equilibrium. Cells regulate their thermodynamic efficiency under varying growth conditions in response to shifts in *in vivo* concentration of cofactors such as NAD^+, $NADP^+$, NADH, NADPH, Acetyl-CoA, HS-CoA, AMP, ADP, ATP, and PP_i. Such shifts impact significantly on the flow of carbon and electrons through the various branches of the fermentation pathways available, and can be manipulated through fermentation conditions or metabolic engineering.

1.7 Manipulation of product yields

Fermentation conditions, including (i) changes in media composition, (ii) pH, (iii) substrate oxidation state, (iv) carbon availability, (v) H_2 and CO_2 partial pressures, and (vi) end-product concentrations present in the culture can alter carbon and electron flux, and thus end-product yields. In many of the studies outlined below, a reoccurring theme is observed; a shift from acidogenesis to solventogenesis is triggered by an increase in NADH concentrations. However, the exact mechanisms and products formed may vary from organism to organism. Thus far, of the studies examining changes in carbon and electron fluxes leading to shifts in end product yields, few have been performed in *C. thermocellum* and merit further investigation.

1.7.1 Media composition

Trace element concentrations (i.e. iron and nickel) have been reported to influence end-product ratios. In heat-shocked mixed cultures, H_2 production increased as iron concentrations increased from 0 to 1600 mg $FeSO_4$ per liter, but decreased at

concentrations exceeding 1600 mg/L (Zhang and Lynd 2005b). Increasing iron concentration (up to 20 mg L^{-1}) enhanced total H_2 production in *Enterobacter cloacae* (Nath et al. 2006). Lee *et al*. reported that H_2, acetate, butyrate, ethanol, and butanol yields in mixed seed cultures obtained from soybean meal were highest at 800, 3, 200, 400, and 200 mg $FeCl_2$ L^{-1}, respectively (Lee et al. 2001). In *C. acetobutylicum*, iron limitation (0.2 versus 10 g L^{-1} $FeSO_4 \bullet SO_4$) decreased methyl viologen-dependent H_2ase and acetoacetate decarboxylase activity by 40% and 25%, respectively, and, in turn, decreased butanol/acetone ratio from 3.7 to 11.8 (Junelles et al. 1988). Specifically, total solvent yields increased by 6%, (ii) acetic and butyric acid yields decreased by 78% and 40%, respectively, and (iii) butanol yields increased by 50% (Junelles et al. 1988). Given the high iron content in H_2ases (roughly 12 Fe per H_2ase), iron limitation may prevent synthesis of H_2ases. This is supported by the fact that addition of iron during H_2ase activity assays did not increase H_2ase activities (Junelles et al. 1988). Furthermore, Dabrock *et al.* reported that iron limitation on glucose-grown *C. pasteurianum* cultures shifted product synthesis from butyrate, butanol, and acetate to lactate, which became the major end-product at iron concentrations of 5.7 μM (Dabrock et al. 1992). Nickel (Ni^{2+}) has also shown to influence product yields in H_2-producing bacteria, likely due to the requirements of Ni^{2+} for [NiFe]-H_2ases. In heat-shocked mixed cultures, H_2 yields increased by roughly 60% when Ni^{+2} concentrations increased from 0 to 0.1 mg/L (Wang and Wan 2008). This increase in H_2 yields was accompanied by a decrease in ethanol and butyric acid yields, and increase in biomass production yields.

Vitamin composition in media can also influence end-product yields. In *C. thermocellum*, substitution of complex 1191 medium with chemically defined medium

(MJ medium), where yeast extract was replaced by pyridoxamine-HCl, biotin, *p*-aminobenzoic acid, and vitamin B_{12}, resulted in good growth yields and similar end-product concentrations with the exception of lactate, which was produced at slightly higher concentrations in the defined medium (Johnson et al. 1981). Growth of *C. cellulolyticum* on defined medium CM3, in which yeast extract was replaced with a vitamin solution and a modified oligoelement solution (containing Na_2SeO_3), resulted in threefold increased cellobiose consumption and a decrease in NADH/NAD^+ ratios (from 57 to 2.08) due to increased rates of NADH reoxidation via lactate, ethanol, and H_2 production (Guedon et al. 1999a). The 10-fold increase in specific production rates of ethanol and lactate on synthetic medium compared to complex medium indicate better control of electron flow (Guedon et al. 1999a).

1.7.2 pH

pH is an important controlling factor of anaerobic fermentation processes. Not only does it affect cell growth and fermentation rate, but also greatly influences end-product synthesis profiles. In fermentative bacteria, organic acid production decreases media pH. Furthermore, these acids (i.e.. acetate and butyrate) are concentrated within the cell under acidic conditions, decreasing the intracellular pH (Huang et al. 1986). Metabolic shifts in response to decreased pH have been observed in a number of organisms, and typically result in a redirection of flux away from organic acid (and H_2) production towards solvent production (Hallenbeck 2005). In *Clostridium actetobutylicum*, a pH shift from 5.6 to 4.3 was accompanied by a shift from acetate/butyrate production to acetone/butanol production (Grupe and Gottschalk 1992).

The shift from acidogenesis to solventogenesis was triggered by changes in ATP/ADP and NADH/NAD^+ ratios. Specifically: (i) intracellular acid concentrations increased, decreasing intracellular pH (ii) ΔpH maintenance via proton-translocating ATPases in response to acid production decreased ATP concentrations, (iii) low ATP/ADP ratios stimulated ATP and NAD(P)H production via glycolysis, and (iv) increased NADH triggered the onset of solventogenesis allowing NAD(P)H reoxidation (Grupe and Gottschalk 1992). In *Sarcina ventriculi*, the shift from acetate/H_2 production to ethanol synthesis in response to decreased pH was accompanied by a decrease *in vitro* H_2ase activity (Goodwin and Zeikus 1987). In *C. cellulolyticum*, a decrease in pH resulted in a decrease in cellulose degradation and biomass formation, and a shift from acetate/H_2 production to ethanol, lactate, and extracellular pyruvate production (Desvaux et al. 2001c). Conversely, a decrease in pH triggered an increase in lactate production accompanied by a decrease in formate, ethanol, and acetate in *Lactobacillus bulgaricus* (Rhee and Pack 1980) and *Streptococcus liquefaciens* (Gunsalus and Wood 1942), whereas no pH-dependent changes in end-product concentrations were observed in *C. beijernicki*, *C. tetanomorphum*, or *C. pasteurianium* (Gottwald and Gottschalk 1985; Dabrock et al. 1992). Hence, the types of pH-dependent end-product shifts are organism-specific.

1.7.3 Substrate oxidation state

Differences in the oxidation state of a substrate lead to shifts in end-product ratios. Hence, given that electron balanced must be maintained, reduced substrates will increase reduced end-product yields, whereas oxidized substrates will increase oxidized end-product yields. However, metabolic networks dictate which specific end products are produced.

Glycerol, produced via transesterification of triglycerides, is a major by-product of biodiesel production. Given that it is more reduced than glucose (redox value of -1), it may serve as a good feedstock for industrial fermentative biofuels production (da Silva et al. 2009). Glycerol can be metabolized into dihydroxyacetone phosphate using either glycerol dehydrogenase and dihydroxyacetone kinase (via a dihydroxyacetone intermediate), or using glycerol kinase and glycerol-3-phosphate dehydrogenase (via a glyceol-3-phosphate intermediate) (Biebl 1991; Luers et al. 1997; Saint-Amans et al. 2001). Both pathways generate an additional NADH, which in turn can be used for alcohol of H_2 production. Alternatively, glycerol can be metabolized to 1,3-propandiol using glycerol dehydratase (DhaB, DhaC, DhaE) and 1,3-propandiol dehydrogenase, during which an NADH is oxidized (Biebl 1991; Luers et al. 1997; Saint-Amans et al. 2001). A number of enterobacter, klebsiella, citrobacter, and saccharolytic clostridia species, including *C. butyricum* and *C. pasteurianum*, can grow anaerobically on glycerol as a sole carbon and energy source (Homann et al. 1990; Biebl 1991; Abbad-Andaloussi et al. 1995; Luers et al. 1997; Biebl 2001; Ito et al. 2005; da Silva et al. 2009).

Addition of glycerol to glucose-grown *C. acetobutylicum* was shown to decrease H_2, acetate, and butyrate production, whereas butanol, and to a lesser extent, ethanol production increased when compared to cells grown on glucose alone (Girbal and Soucaille 1994; Vasconcelos et al. 1994). The 7-fold increase in intracellular NADH levels in glucose-glycerol grown *C. acetobutylicum* cultures may account for the shift to solvent production, during which NADH is reoxidized (Vasconcelos et al. 1994). This is further supported by the fact that the increase in ethanol is less than the increase in butanol production, during which twice as many NADH equivalents are reoxidized. Initially, a 2.5-fold increase in ATP concentration was also observed in these cultures (Vasconcelos et al. 1994), although subsequent studies report decreased levels of intracellular ATP (Girbal and Soucaille 1994), which would correspond to a decrease in flux through energy-conserving acidogenic pathways. Similar effects were seen in *Clostridium butyricum* grown on mixtures of glucose and glycerol, during which 1,3-propanediol production increased considerably, while acetate, butyrate, H_2, and CO_2 production decreased (Saint-Amans et al. 2001).

Pyruvate, which is more oxidized than glucose (redox value of +1) can also be used to study how flux is distributed through end-product synthesis pathways in response to substrate oxidation states. A number of organisms, including *Thermoanarobacter brocii* (Ben-Bassat et al. 1981), *Clostridium thermosaccharolyticum* (Lee and Ordal 1967), *Corynebacterium glutamicum* (Jolkver et al. 2009), *Pyrococcus furiosus* (Schäfer and Schönheit 1991), *Streptococcus faecalis* (Deibel and Niven 1964), *Oenococcus oeni* and *Leuconostoc mesenteroides* (Wagner et al. 2005), can use pyruvate as a sole carbon and electron source anaerobically, whereas others, including *Clostridium acetobutylicum*

(Junelles et al. 1987), *Oenococcus oeni* and *Leuconostoc mesenteroides* (Zaunmuller et al. 2006) do not grow on pyruvate as a sole carbon source, but do metabolize pyruvate in the presence of hexose sugars.

T. brockii, grown solely on pyruvate, displayed greater acetate yields, and lower ethanol and lactate yields, when compared to growth on glucose (Ben-Bassat et al. 1981). The addition of pyruvate to glucose grown *C. acetobutylicum* has been shown to promote acetogenesis over ethanologenesis, and was accompanied by a decrease in acetoacetate decarboxylase activities, while acetate kinase and butyrate kinase activities were unchanged, (Ballongue et al. 1986; Junelles et al. 1987). Girbal and Soucaille demonstrated that partial replacement of glucose with pyruvate in glucose/glycerol-grown *C. acetobutylicum* cultures increased butyrate, CO_2, and lactate, production rates and decreased ethanol and acetoin production rates when compared to growth on glucose and glycerol alone (Girbal and Soucaille 1994). These changes were accompanied by low ATP/ADP and high NADH/NAD^+ ratios, as well as higher butryaldehyde dehydrogenase, phosphotransacetylase, acetate kinase, phosphotransbutyrylase, butyryl kinase, pyruvate:ferredoxin oxidoreductase, and ferredoxin:NAD^+ reductase (analogous to NADH:ferredoxin oxidoreductase) activities. It was proposed that high ATP levels can decrease expression of phosphotransbutyrylase and butyryl kinase expression, whereas high NADH/NAD^+ ratios can increase expression of ethanol dehydrogenase, butryaldehyde dehydrogenase, butanol dehydrogenase, and ferredoxin:NAD^+ reductase, allowing for a more efficient electron flow from reduced ferredoxin to NADH and subsequent production of butanol and ethanol (Girbal and Soucaille 1994).

1.7.4 Carbon and nitrogen loading

The effect of carbon and nitrogen loading and carbon source on carbon flux in a cellulolytic, fermentative, H_2 and ethanol-producing bacterium has, perhaps, been best studied in *Clostridium cellulolyticum*. Payot *et al*. demonstrated that *C. cellulotyicum* grown on 23.4 mM cellobiose in pH controlled (pH 7.2) batch cultures displayed biphasic growth kinetics (Payot et al. 1999). During the first growth phase, acetate was the major end-product, and an accumulation of intracellular NADH levels was observed. Conversely, in the second growth phase, lactate production was induced and became the major end-product of fermentation. Induction of lactate production was also accompanied by a decrease in NADH concentrations, allowing growth to resume. During both growth phases, H_2/CO_2 ratios were 1.4-1.5 demonstrating that H_2 can be generated from NADH rather than from reduced ferredoxin alone. It was proposed that NADH is used to reduce ferredoxin, using NADH:ferredoxin oxidoreductase, which can in turn be used to produce H_2 via a Fd-dependent H_2ase.

Guedon *et al*. demonstrated that changes in dilution rates (D = 0.016 to 0.138 h^{-1}) of feed medium containing limited cellobiose concentrations (5.84 mM) had a significant impact on *C. cellulolyticum* metabolism when grown on synthetic medium in continuous cultures (Guedon et al. 1999b). At low dilution rates, high H_2/CO_2 and $NADH/NAD^+$ ratios were observed, coinciding with low ethanol and lactate production rates and low *in vitro* ethanol and lactate dehydrogenase activities. Conversely, at high dilution rates, and thus higher specific cellobiose consumption rates, (i) intracellular dihydroxyacetone phosphate, glyceraldehydes-3-P, and fructose-1,6-bisphosphate levels increased, (ii) H_2/CO_2 and $NADH/NAD^+$ ratios decreased, (iii) extracellular pyruvate concentrations

increased, (iv) ethanol and lactate production rates increased, and (v) *in vitro* ethanol and lactate dehydrogenase activities increased. It was proposed that at low carbon fluxes, acetate is preferentially produced in order to maximize ATP formation and redox balance is maintained primarily by NADH:Fd oxidoreductase and H_2ases, whereas at high carbon fluxes, and presumably high ATP/ADP ratios, redox balance is maintained via lactate dehydrogenase and ethanol dehydrogenase, resulting in increased lactate and ethanol production.

Interestingly, manipulation of dilution rates (D = 0.035 to 0.115 h^{-1}) using cellobiose-sufficient (14.62 mM) rather than cellobiose-limited (5.84 mM) feed medium had a different impact on end-product synthesis patterns in *C. cellulolyticum* (again grown in continuous cultures on synthetic medium) (Guedon et al. 2000b). Increasing dilution rates of cellobiose-sufficient feed medium (i) decreased intracellular dihydroxyacetone phosphate levels, (ii) increased H_2/CO_2, $NADH/NAD^+$ and ATP/ADP ratios (iii) eliminated lactate production, which was observed at low dilution rates, (iv) increase acetate and ethanol production rates, (v) increased *in vitro* acetate kinase and ethanol dehydrogenase activities, and (vi) decreased *in vitro* lactate dehydrogenase activities. At high carbon flow, excreted amino acids, extracellular proteins, and exopolysaccharides were also detected, the latter constituting 34% of cellobiose consumed at D = 0.115 h^{-1}. Subsequent studies using *C. cellulolyticum* grown on excess cellobiose (14.62 mM) in continuous cultures with ammonia as the growth limiting nutrient, which was used to decouple catabolism and anabolism, revealed that as much as 21.4, 16.7, and 16.0% of cellobiose consumed was converted into glycogen, exopolysaccharides, and cellotriose (Guedon et al. 2000a). Given that

exopolysaccharides and glycogen are produced from glucose-1-phosphate, it was proposed that under cellobiose-excess conditions, the glucose-1-phosphate/glucose-6-phosphate branchpoint (presumably consisting of cellobiose phosphorylase, glucose kinase, and phosphoglucokinase) plays a critical role in excretion of carbon surplus and energy dissipation (Guedon et al. 2000a).

Parallel *C. cellulolyticum* studies (i.e. continuous culture experiments on synthetic medium using different dilution rates with carbon-limited, carbon excess, and ammonia-limited feed medium) have been performed using cellulose rather than cellobiose (Desvaux et al. 2001b; Desvaux et al. 2001a; Desvaux and Petitdemange 2001). In contrast to studies using cellobiose-limited feed medium (Guedon et al. 1999b), increasing dilution rates (D = 0.014 to 0.083) using cellulose-limited (0.37% wt/vol) feed medium did not increase lactate production or generate extracellular pyruvate, but did increase ethanol/acetate ratios (Desvaux et al. 2001a). Although growth yields were better on cellulose than cellobiose, growth rates were lower (Desvaux et al. 2001a). The transition from cellulose-limited to cellulose sufficient ($\geq$ 7.2 g cellulose L^{-1}) conditions in continuous cultures increased acetate/ethanol ratios and lactate production (Desvaux et al. 2001d). In contrast to cellulose-limited continuous cultures, increasing dilution rates of cellulose-sufficient feed medium (1.81% wt/vol) increased both acetate/ethanol ratios and decreased lactate production (Desvaux et al. 2001d). Furthermore, ATP/ADP and $NADH/NAD^+$ ratios were higher suggesting a more active metabolism. Conversely, maximum growth yields and energetic yields decreased. When compared to cellobiose-sufficient conditions (Guedon et al. 2000b), cellulose-sufficient conditions resulted in (i) a lower rate of influx of carbon, (ii) lower $NADH/NAD^+$ ratios, and (iii) less production

of exopolysaccharides, extracellular protein, and free amino acids (Desvaux et al. 2001b). This may be due to a more efficiently regulated carbon flow on cellulose versus cellobiose, likely stemming from the rate-limiting step of cellulose degradation (Desvaux et al. 2001b; Desvaux and Petitdemange 2001).

The effects of carbon limitation and carbon excess conditions (≤1.1g L^{-1} & 4.5g L^{-1} cellobiose respectively) have also been studied in *C. thermocellum* in batch cultures. At high concentrations of cellobiose, growth ceased prior to exhaustion of carbon source possibly due to a drop in pH (~6.0), end-product inhibition, or nutrient exhaustion (Islam et al. 2006). Growth rates and end-product ratios were generally unchanged by different substrate concentrations with the exception of lactate synthesis, which was stimulated in late exponential phase and became a major end-product (Islam et al. 2006). In carbon-excess cultures, small amounts of extracellular pyruvate were also detected (Islam et al. 2006). Changes in dilution rates, and in turn growth rates, also affected *C. thermocellum* end-product yields and gene expression. As in *C. cellulolyticum,* acetate yields were high in order to maximize ATP production during low dilution rates (using cellobiose or cellulose-limited (0.3% wt/vol) feed) and low growth rates (Strobel 1995; Stevenson and Weimer 2005). However, as dilution rates increased, flux through glycolysis, and in turn ATP production increased, resulting in an increase in ethanol/acetate ratios (Strobel 1995). Although H_2 or CO_2 production was not monitored in these studies, increases in gas production presumably mirrored acetate production. RT-PCR analysis of putative genes encoding for enzymes potentially involved in end-product formation (*hydA*, *ack*, *ldh*, & *adhZ*) displayed modest increases in expression with increasing dilution rate on

both cellulose and cellobiose, whereas expression of *adhY* decreased (Stevenson and Weimer 2005).

1.7.5 End-product inhibition

The thermodynamic efficiency of H_2 has been shown to be dependent on H_2 partial pressures (Angenent et al. 2004). Furthermore, H_2 evolving pathways may be inhibited by H_2, leading to a metabolic shift towards other reduced products such as lactate, ethanol, butanol, butyrate, and acetone (Soboh *et al*. 2004, Levin *et al*. 2004). Growth inhibition by H_2 has also been demonstrated in several organisms (i.e. *Thermotoga maritima* and *Pyrococcus furiosus*) (Schäfer and Schönheit 1991; Soboh et al. 2004). Hence, decreasing H_2 partial pressures by intermittent venting, stirring, N_2 sparging, or H_2 consumption using methanogens may increase growth and H_2 yields.

A number of studies have demonstrated how H_2 partial pressures can affect end-product yields. Vigorous stirring of *C. thermocellum* cultures resulted in a threefold decrease in ethanol/acetate ratio and proportional increase in H_2 production (Lamed et al. 1988). This was attributed to the fact that transfer of H_2 from the liquid to the gas phase is lower in unstirred cultures, which can lead to H_2 supersaturation of the culture medium and thus inhibition of acetate production (Lamed et al. 1988). Similarly, ethanol/acetate ratios were also lower when *C. thermocellum* was grown in open continuous cultures stirred and sparged with N_2 at a low rate rather than in sealed batch cultures (Lamed et al. 1988). Soboh *et al*. also demonstrated that decreased H_2 partial pressures resulting from growth of *Thermoanaerobacter tengcongensis* (a.k.a. *Caldanaerobacter subterraneus* subsp. *tengcongensis*) in open versus closed batch cultures eliminated ethanol production,

increased acetate yields from 0.7 to 2.0 moles per mole glucose, and presumably increased H_2 yields from 0.3 to 4.0 moles per mole glucose (Soboh et al. 2004). These changes were accompanied by a four-fold increase in NADH-dependent H_2ase activity, negligible changes in Fd-dependent H_2ase activity, and a decrease in aldehyde dehydrogenase and alcohol dehydrogenase activities (Soboh et al. 2004). N_2 sparging of a mixed culture primarily composed of *Clostridium* species also increased H_2 yields and production rates from 0.85 to 1.43 moles H_2 per mole glucose and 1.4 to 3.1 ml H_2 min^{-1} g^{-1} biomass, respectively (Mizuno et al. 2000). Seeding *Methanobacterium thermoautotrophicum* with *C. thermocellum* also resulted in decreased ethanol production and increased acetate and H_2 production (calculated from methane formation) (Thauer et al. 1977; Weimer and Zeikus 1977).

Although many of these methods can increase H_2 yields, recovery of dilute H_2 gas is difficult and costly using presently available technologies. An alternative and more practical way reducing fermenter H_2 partial pressures is through the use of various membrane technologies that allow for H_2 specific removal. The use of a silicone rubber membrane to separate biogas from liquid medium was shown to improve H_2 yields and production rates by15% and 10%, respectively (Liang et al. 2002).

Alternatively, ethanol production can be increased by increasing H_2 partial pressures. Exogenous H_2 addition (0.4 to 1.0 atm) increase ethanol/acetate ratios of *T. brocii* and *C. thermocellum* strain AS39 but not LQRI (Lamed and Zeikus 1980a). *T. brockii* growth rate decreased at 0.5 atm H_2 and growth was completely inhibited at 1.0 atm H_2 (Ben-Bassat et al. 1981). A similar increase in ethanol/acetate ratios in response to exogenous H_2 was observed using *C. thermocellum* JW20 (Freier et al. 1988) and YS

(Lamed et al. 1988). Interestingly, Lamed *et al*. also demonstrated that addition of tritium ($^3H_2$) to the gas phase resulted in tritiated water (^3H_2O), suggesting that *C. thermocellum* possesses H_2ases capable of H_2 uptake (Lamed et al. 1988). Bothun *et al*. manipulated dissolved H_2 concentrations by changing hydrostatic pressures rather than by stirring or sparging (Bothun et al. 2004). They demonstrated that elevating hydrostatic pressures to 7.0 or 17.3 MPa, *C. thermocellum* JW20 growth was inhibited by roughly 40 and 60%, respectively, whereas ethanol/acetate ratios increased by 38- and 60-fold, respectively.

While there have been numerous reports discussing the effects of H_2 on end-product yields, fewer studies have looked at the effects of sub-inhibitory concentrations of CO_2, ethanol, or carboxylic acids. Park *et al*. have shown that KOH-mediated scavenging of CO_2 produced by mixed consortia in batch increased H_2 yields by 43% (from 1.4 to 2.0 moles H_2/mol glucose), (Park et al. 2005). Alternatively, addition of exogenous butyrate (47 mM) was shown to decrease H_2 (8%) production, and increase acetate (21%) and ethanol (27%) production in mixed anaerobic cultures (Zheng and Yu 2005). He *et al*. demonstrated that addition of acetate (150 mM) to *Thermoanaerobacter ethanolicus* batch cultures grown in the absence of yeast extract increased ethanol production by as much as 210, 175, and 300%, when cells were grown on xylose, glucose, or cellobiose, respectively (He et al. 2009). Similar results were observed when exogenous acetate was added to cellobiose-grown *C. thermocellum* cells in the absence of yeast extract (He et al. 2009). Alternatively, addition of acetate to cultures grown in the presence of yeast extract increased ethanol production by only ~10% on xylose or glucose, regardless of acetate concentration used (He et al. 2009). Furthermore, addition

of lactate (150 mM) to *T. ethanolicus* cells decreased ethanol production by ~20% in the absence of yeast extract and increased ethanol production by ~35% in the presence of yeast extract (He et al. 2009). Few other studies, to our knowledge, report the effects of addition of exogenous end-products or removal of endogenous end-products on product yields in H_2/ethanol producing bacteria.

1.7.6 Exogenous electron carriers

Shift in metabolism is strongly influenced by the regulation of electron flow. Thus, the use of reducing agents such as neutral red (-530 mV), methyl viologen (-446 mV), or benzyl viologen (-359 mV) can manipulate electron flux (Michaelis and Hill 1933). The addition of these dyes tends to increase solvent production and decrease H_2 and acid production. Girbal *et al*. demonstrated that addition of neutral red to *C. acetobutylicum* cultures increased butanol and ethanol production, and decreased H_2, acetate, and butyrate production (Girbal et al. 1995). This was accompanied by an increase in aldehyde dehydrogenase, alcohol dehydrogenase, and NFOR activities, and a decrease in phosphotransacetylase and phosphotransbutyrylase activities. Rao *et al*. demonstrated that the addition of methyl viologen or benzyl viologen to *C. acetobutylicum* cultures decreased H_2, butyrate, and acetate production, and increased butanol and ethanol production (Rao and Mutharasan 1987; Rao et al. 1987). Similarly, addition of methyl viologen decreased H_2, acetate, and lactate production, and increased ethanol production in *T. ethanolicus* (Rao and Mutharasan 1987).

It was proposed that addition of these dyes can divert electrons from reduced ferredoxin to $NAD(P)^+$, thus increasing NADH production rates used for solventogenesis

at the expense of H_2 production (Rao and Mutharasan 1987; Rao et al. 1987; Girbal et al. 1995). It was proposed that viologen dyes may competitively block binding of reduced ferredoxin to the H_2ase active site, thus decreasing ferredoxin oxidation via H_2 production (Rao and Mutharasan 1987). Given that reoxidation of ferredoxin is still required for pyruvate oxidation via PFOR, oxidation of reduced ferredoxin may instead occur via NFOR, generating NADH for solventogenesis. Alternatively, viologen dyes may be reduced directly via PFOR, which can in turn reduce NAD^+ either directly, or indirectly by first reducing reduced ferredoxin, which can in turn reduce NAD^+ using NFOR. Reduced viologens have in fact been shown to directly donate electrons directly to NAD^+ *in vitro*, likely due to the large difference in redox potential between the two electron carriers (Adams et al. 1981). Interestingly, these mechanisms were proposed before the characterization of bifurcating H_2ases. The presence of viologen dyes may competitively block binding of ferredoxin to the PFOR active site, thus decreasing ferredoxin reduction. Given lower ferredoxin reduction rates, H_2 production rates via bifurcating H_2ases may decrease, thus increasing $NADH/NAD^+$ ratios which can be balanced via solventogenesis. If these viologen dyes are in fact reduced via PFOR, they may again transfer electrons directly to NAD^+. Interestingly, the decrease in lactate production in *T. ethanolicus*, despite higher NAD^+ reduction rates, was thought to be a consequence of increased flux through PFOR resulting from a higher turnover rate of reduced ferredoxin (Rao et al. 1987).

1.7.7 Enzyme inhibitors

Redirection of carbon and electron flux has also been demonstrated through the use of enzyme inhibitors, which can mimic gene knockouts/knockdowns. These protein inhibitors, including hypophosphite (HPP) and carbon monoxide (CO), may be used to mimic gene knockouts/knockdowns of PFL and H_2ases, respectively. HPP, a formate analogue which can inhibit PFL activity by covalently binding its active site via a free radical dependent mechanism, can be used to mimic PFL deletions (Brush et al. 1988; Knappe and Sawers 1990). Preliminary studies by Maness *et al.* demonstrated that HPP increased methyl viologen-dependent H_2ase activity and H_2 production in *C. thermocellum* (Maness and Thammannagowda 2008). Alternatively, CO, which inhibits H_2ases by binding to vacant [FeFe] or [NiFe] coordination sites, can be used to mimic H_2ase deletions (Hallahan et al. 1986; Fauque et al. 1988; De Lacey et al. 2002). Addition of CO to the headspace of *C. acetobutylicum* redirected carbon and electrons away from H_2, CO_2, acetate, and butyrate towards ethanol and butanol (Kim et al. 1984; Datta and Zeikus 1985; Meyer et al. 1985; Meyer et al. 1986). While inhibition of H_2ases using CO in *C. thermocellum* has been shown to increase ethanol/acetate ratios as well as lactate production (Li et al. 2012a), H_2, CO_2, and formate concentrations were not reported.

1.7.8 Mutagenesis

Metabolic engineering can be used to improve cellular function through modulation of enzymatic, transport, or regulatory functions in the cell. While random mutagenesis can alter cellular function, the characterization and screening of different

phenotypes generated by random mutagenesis is often cumbersome. Two different approaches may be used to produce and/or select for a desired phenotype. The first approach, here called '*directed mutagenesis*', employs the use of a selection agent, typically an anti-metabolite that may be metabolized by either specific proteins or fermentation conditions to produce cytotoxic compounds. In the presence of these anti-metabolites, mutants generated via random mutagenesis that cannot metabolize these anti-metabolites into cytotoxic compounds will be selected for. The second approach, here called '*targeted mutagenesis*', utilizes molecular biology techniques to delete, modulate, or (over)express specific genes (selected based on biochemical and genetic characterizations) to potentially improve strain characteristics. Manipulation of product yields using these two approaches are discussed below.

1.7.8.1 ***Directed mutagenesis***

The use of cytotoxic anti-metabolites, specific for proteins of interest, may be used to select for mutations that affect the expression or structure of these targeted proteins, in turn preventing metabolism of an anti-metabolite and thus the toxic effects of its products. The advantage of using anti-metabolites for mutagenesis studies is that it is possible to obtain phenotypes with improved product yields. The use of fluoroacetate, the proton suicide method (PMS), and allyl alcohol are discussed.

In *E. coli* and *Klebsiella aerogenes,* fluoroacetate can be metabolized by acetate kinase and phosphotransacetylase to fluoroacetyl-CoA, which can then be converted to fluorocitrate via citrate synthase (Brown et al. 1972; Brown et al. 1998). Fluorocitrate can then irreversibly bind aconitase, preventing the citric acid cycle, and is thus cytotoxic

to the cell (Glusker 1971). Hence, fluoroacetate can be used to select for ACK and PTA mutants. *Ta. thermosaccharolyticum* fluoroacetate resistant phosphotransacetylase (PTA) and acetate kinase (ACK) mutants, which could not metabolize fluoroacetate to cytotoxic fluorocitrate, produce as much as 70% less acetate and two times as much ethanol (Rothstein 1986).

The proton suicide method (PSM), using sodium bromate ($NaBrO_3$) and sodium bromide (NaBr), is an alternative method to select for mutants unable to produce acids (Winkelman and Clark 1984). Bromate and bromide can react with protons produced by acids to make cytotoxic elemental bromine as shown in the reaction mechanism below:

1) $BrO_3^- + 2Br^- + 2H^+ \rightarrow Br_2 + BrO_2^- + H_2O$
2) $BrO_2^- + 2Br^- + 2H^+ \rightarrow Br_2 + BrO^- + H_2O$
3) $3BrO^- \rightarrow Br_2 + BrO_3^-$

Thus, mutations that lead to decreased acid production may prevent the toxic effects of bromate and bromide. *Klebsiella oxytoca* PSM mutants displayed a 7.8% increase in alcohol (2,3-butanediol) production accompanied by an 88 and 92% decrease in lactic and acetic acid production (Ji et al. 2008). However, while some *C. acetobutylicum* PSM mutants produced more butanol and less acids (butyric, acetic), others produced less butanol and more acids suggesting that these mutants may be defective in acid recycling (Cueto and Mendez 1990).

Alternatively, allyl alcohol can be dehydrogenated by ADHs to acrolein, which then irreversibly alkylates and inhibits enzymes, producing cytotoxic effects. Thus allyl alcohol can be used to select for ADH and/or AldH mutants, in which electron flux

redirected away from ethanol (Rando 1972). *Enterobacter aerogenes* treated with allyl alcohol displayed higher H_2 yields and lower ethanol yields (Rachman et al. 1997), while *E. coli* cell extracts treated with allyl alcohol shown lower ADH and/or AldH activities (Lorowitz and Clark 1982). Interestingly, allyl alcohol treatment of *C. acetobutylicum* had little effect on acetone, ethanol, and in most cases acetate production, but prevented butanol production (Dürre et al. 1986). It was shown that the cell extracts from these cells had lower butryaldehyde activities and higher H_2ase activities, while butanol dehydrogenase activities remained unchanged. *Enterobacter cloacae* mutants that were selected for allyl alcohol resistance and subsequent PSM resistance produced increased H_2 yields from 2.2 to 3.5 mol mol-glucose^{-1} (Kumar et al. 2001), demonstrating the effectiveness of using these techniques for manipulating product yields.

1.7.8.2 ***Targeted mutagenesis***

A number of genetic engineering strategies have been employed to optimize ethanol or H_2 yields in various fermentative bacteria. Typically, genes encoding pathways that produce unwanted products are inactivated or suppressed using knock-out vectors or anti-sense RNA. Alternatively, native genes encoding desirable pathways are either overexpressed or foreign genes can be heterologously expressed to introduce new pathways. These strategies aim to divert carbon and electron flow from unwanted by-products towards desired products such as H_2 and ethanol. While many of these strategies have produced desirable results, others have not. Therefore, in order to optimize genetic and regulatory processes that lead to alleviation of metabolic constraints that may result in higher biofuels yields and production rates, a thorough understanding

of the complex metabolic and regulatory networks that dictate metabolic flux must first be elucidated.

1.7.8.2.1 Increasing ethanol

Improvement of ethanol yields via genetic engineering has been most successful in a number of Gram-negative bacteria including *E. coli*, *K. oxytoca*, and *Zymomonas mobilis* as they are amiable to genetic engineering (reviewed by Dien *et al*. (Dien et al. 2003)). Although these bacteria cannot degrade raw lignocelluosic biomass, *E. coli* and *K. oxytoca* ferment a wide spectrum of sugars, including cellobiose and cellotriose for the latter. However, they encode branched end-product pathways that produce various organic acids and neutral products. *Z. mobilis*, on the other hand, has a homo-ethanol fermentative pathway but cannot ferment xylose and arabinose. While *Z. mobilis* produces ethanol via pyruvate decarboxylase (PDC) and ADH, *E. coli* and *K. oxytoca* produce ethanol via PFL, acetaldehyde dehydrogenase, and ADH. This pathway is unbalanced because only one NADH is generated per each pyruvate made via glycolysis, and 2 NADH are required to produce one ethanol. Therefore a number of acids (i.e. acetic, succinic) are produced to balance the additional pyruvate that must be produced to generate the additional NADH for ethanol production (Dien et al. 2003).

One of the most common methods of increasing ethanol yields is through the introduction of the PET (*P*roduction of *ET*hanol) operon which contains *pdc* and *adhII* from *Z. mobilis* under the control of the native *lac* promoter. Ingram *et al*. was the first to introduce this operon on a plasmid into *E. coli* (Ingram et al. 1987). They found that heterologous expression of *Z. mobilis* ADH, had no effect on ethanol production, whereas

expression of PDC or PDC with ADH increased final ethanol concentrations to 18 and 750 mM (95% of fermentation products), respectively, compared to the wild-type (18 mM lactate; 7 mM acetate; 1.5 mM succinate; trace ethanol). Ohta *et al*. subsequently integrated the PET operon into the PFL gene of the *E. coli* chromosome to eliminate the competing pyruvate oxidation pathway and to stabilize these heterologous genes (Ohta et al. 1991a). However lower amounts of ethanol (~4.0 g L^{-1}) were produced due to reduced gene dosages. This was remedied by selecting for spontaneous mutants resistant to high concentrations of chloramphenicol, which in turn produced high levels of *Z. mobilis* genes (Ohta et al. 1991a). Further disruption of the gene encoding fumarate reductase (*frd*), the terminal gene involved in the succinate production, marginally increased ethanol production as well (Ohta et al. 1991a). Similar strategies were employed for *K. oxytoca* and produced comparable results (Ohta et al. 1991b).

More recent research has been directed at developing molecular biology techniques for genetic engineering of Gram positive bacteria in order to improve ethanol yields, particularly in bacteria capable of degrading lignocellulosic biomass. The first cellulolytic bacteria in which *Z. mobilis* PDC and ADH (i.e. PET operon) were heterologously expressed was *C. cellulolyticum* (Guedon et al. 2002). In the presence of high concentrations of substrate (cellobiose or cellulose), growth of *C. cellulolyticum* is typically inhibited due to high flux through glycolysis that cannot be alleviated by fermentation pathways. This is typically associated with pyruvate excretion, high $NADH/NAD^+$ ratios, and increased flux towards lactate production (See above – section 1.7.4). Heterologous expression of PDC and ADH indeed increased cell dry weight, cellulose consumption, and final H_2, acetate, CO_2, and ethanol production by 180, 150,

~100, 93, ~80, and 53%, respectively, whereas lactate and extracellular pyruvate production decreased by 48 and 60%, respectively (Guedon et al. 2002). In contrast to PET expression in *E. coli*, acetate production did not decrease and ethanol did not approach near maximal (2 mol mol-glc^{-1}) yields.

Alternative metabolic engineering strategies have been employed to increase ethanol yields in *Thermoanaerobacterium saccharolyticum* and *C. cellulolyticum*, in which competing pathways were deleted. Knockouts of LDH had little impact on fermentation end-products in *Ta. saccharolyticum* grown on 4g L^{-1} xylose (Shaw et al. 2008). However, both *pta*/*ack*$^{-}$ and *ldh*/*pta*/*ack*$^{-}$ mutants (ALK1 strain) produced 37% more ethanol and negligible amounts of acetate or H_2 (Shaw 2008). Mutation of *mdh* and *ldh*, which have overlapping substrate specificities in *C. cellulolyticum*, also increased ethanol production (Li et al. 2012b). While lactate production in *ldh*$^{-}$ mutants on both cellobiose (5 g L^{-1}) and cellulose (10 g L^{-1}) was virtually eliminated, an increase in acetate and ethanol production was observed on cellobiose, but not on cellulose. Conversely, *mdh*$^{-}$ mutants marginally decreased acetate and increased ethanol production on both cellobiose and cellulose, but lactate production was increase on cellobiose and decreased on cellulose. This suggests that fluxes may be regulated in response to substrate and/or carbon loading (see section 1.7.4). Double mutants (*mdh*/*ldh*$^{-}$), however, produce less lactate and acetate, and more ethanol (as high as 8.5-fold on cellulose), regardless of growth substrate.

Surprisingly, deletion of *pta* in *C. thermocellum* DSM 1313, which eliminated acetate production, resulted in negligible changes in lactate or ethanol production (Tripathi et al. 2012). Similarly, while deletion of *ldh* and/or *pta* in *C. thermocellum*

DSM 1313 eliminated lactate and acetate production, respectively, there were negligible increases in ethanol production in these mutants (Argyros et al. 2011). Subsequent evolution of the double mutant (△*ldh* △*pta*) on increasing cellulose concentrations, however, did result in a 4.2-fold increase in ethanol yield (Argyros et al. 2011). A recent study by Deng *et al*. demonstrated that heterologous expression of *T. saccharolyticum* pyruvate kinase (inserted at the *ldh* locus) accompanied by decreased expression of phosphoenolpyruvate carboxykinase (obtained by modifying its start codon from ATG to GTG) in *C. thermocellum* DSM 1313 increased ethanol yields by 3.25 (Deng et al. 2013). Similarly expression of pyruvate kinase accompanied by deletion of MalE and partial deletion of MDH increased ethanol yields by more than 3-fold. The increase in ethanol production in both strains was accompanied by a decrease in acetate and H_2 yields and an increase in formate yield. Redirection of flux away from the malate branch involved in transhydrogenation evidently increased NADH availability for ethanol production. Finally, a recent patent by Guss and Lynd describe a *C. thermocellum* strain in which *pfl* and *hydG* (required for [FeFe]-H_2ase maturation) mutations eliminated formate and H_2 production, and decreased acetate production, and increased ethanol yields (Guss and Lynd 2012)

1.7.8.2.2 Increasing H_2

Various metabolic engineering strategies have also been employed for improvement of H_2 yields in *E. coli*. Unlike most clostridia, which produce H_2 using H_2ases, *E. coli* produces H_2 from formate via a formate:hydrogen lyase complex (FHL) consisting of a formate dehydrogenase F (FDH_F; *fdhF*) and a H_2ase (Hyd3; encoded by

hyc genes), and thus only 2 mol H_2 per mol glucose can be obtained. H_2 yields are further decreased due to presence of genes encoding competing pathways that lead to (i) production of succinate (fumarate reductase, *frdBC*), lactate (LDH, *ldhA*), acetate (pyruvate oxidase, *poxB*), and acetate/ethanol (PHD, *aceE*), (ii) H_2 uptake (Hyd1, *hya* genes; Hyd2, *hyb* genes), (iii) oxidation of formate not linked to H_2 production catalyzed by membrane bound formate dehydrogenases (FDH_N, *fdnGHI*; FDH_O, *fdoGHI*), and (iv) formate export (*focAB*). Deletions of these various genes (Reviewed by Vardar-Schara *et al*. (Vardar-Schara et al. 2008)) have in fact led to improved H_2 production rates and H_2 yields, albeit yields were still below 2 mol H_2 mol-glc^{-1} (Vardar-Schara et al. 2008). Alternatively, heterologous overexpression of *Enterobacter cloacae* H_2-evolving [FeFe]-H_2ase (*hydA*) in *E. coli* increased H_2 yields to 3.1 mol mol-glc^{-1} (Chittibabu et al. 2006). Other engineering strategies used to increase H_2 production in *E. coli* include (i) deletion of FHL repressor (*hycA*) (Penfold et al. 2006; Maeda et al. 2007; Yoshida et al. 2007) (ii) overexpression of FHL activator (*fhlA*) (Yoshida et al. 2005), and (iii) mutation of the twin-arginine translocation (TAT) system to prevent transport of Hyd1, Hyd2, FDH_N, FDH_O to the membrane, and in turn H_2 uptake and non-H_2 coupled formate dehydrogenase activity (Penfold et al. 2006).

Fewer engineering attempts to increase H_2 production have been made in Gram positive bacteria, and specifically in clostridia. One example of successful improvement of H_2 production via genetic engineering in clostridia was demonstrated using *C. paraputrificum*. Overexpression of a native a [Fe]-H_2ase (*hydA*) in *C. paraputrificum*, which also encodes a FHL, increased H_2 production by 1.7-fold (Morimoto et al. 2005). The increase in H_2 production was accompanied by abolishment of lactate production,

increase in acetate production, and negligible changes in butyrate and formate production. More recent studies demonstrated that heterologous overexpression of a *Candida boidinii* NAD^+-dependent formate dehydrogenase FDH1 gene (*fdh1*) in *C. paraputrificum* generated more NADH from formate, which led to a 59% increase in H_2 yield (Lu et al. 2012). An alternative strategy was used to improve H_2 (and butyrate) production in *C. tyrobutyricum*. Deletion of acetate kinase (*ack*) in *C. tyrobutyricum* resulted in improved tolerance to butyrate, increased H_2 and butyrate production (50 and 23.5%, respectively), and higher H_2 and butyrate yields (Liu et al. 2006). Surprisingly, acetate production was not affected in the mutant. Furthermore, while both PTA and H_2ase activities increase by 40%, ACK activity decreased by only 50% in the mutant strain. It was suggested that other enzymes (ex. phosphotransbutyrylase, CoA-transferase, ACK isoenzymes) may be responsible for observed acetate kinase activity in the mutant.

1.8 *Clostridium thermocellum*

Clostridium thermocellum ATCC 27405 (DSM 1237) is an anaerobic, strictly fermentative gram positive bacterium, possessing all of the qualities of a good candidate for consolidated bioprocessing. It exhibits one of the highest growth rates on cellulose (Lynd and Grethlein 1987; Lynd et al. 1989; Lynd et al. 2002) and is capable of fermenting cellulose and other β-glucans directly to H_2, ethanol, CO_2, acetate, formate, and lactate (Islam et al. 2006; Sparling et al. 2006). Its thermophilic nature gives it improved H_2 production rates when compared to mesophiles (Hallenbeck 2005), may allow for the utilization of non-sterile substrates (Demain et al. 2005), and provides

enhanced H_2 recovery since high temperatures decrease the partial pressure of dissolved gasses in the medium (Angenent et al. 2004). Despite these appealing characteristics, *C. thermocellum* normally produces ethanol and H_2 yields (~0.6 and 1.3 mol per mol hexose, respectively) well below the 'Thauer limit' of either 2 moles of ethanol or 4 moles of H_2 per mol hexose, respectively (Thauer et al. 1977; Islam et al. 2006; Sparling et al. 2006; Islam et al. 2009). This is due to branched fermentative pathways that lead to the production of both ethanol and H_2 (with concomitant production of CO_2 and acetate), as well as branches leading to formic acid and lactic acid that compete for carbon and/or electrons required for the production of either ethanol or H_2 (Thauer et al. 1977; Lynd et al. 2005).

To date, four strains of *C. thermocellum*, including DSM 1237 (ATCC 27405; VPI7372)[1], DSM 2360 (ATCC 35609; LQRI), DSM 4150 (ATCC 31549; JW20) and DSM 1313 (LMG 6656; LQ8) have been sequenced. This has allowed for deduction of metabolic pathways *in silico*, PCR, microarray, and proteomic analysis, and the potential for genetic engineering strategies. Genomic analyses of these strains reveal that they encode different proteins with redundant functions. While activities of select proteins have been confirmed by enzyme activity assays (Lamed and Zeikus 1980a) in strain DSM 2360 and AS39, characterization of isolated proteins (lactate dehydrogenase and acetate kinase) (Lin et al. 1998a; Özkan 2004) and verification of gene transcription of genes involved in pyruvate catabolism and the branched end-product pathway using RT-

[1] Note that strain names will be used interchangeably throughout this thesis

PCR (Carere et al. 2008a) has been verified in strain DSM 1237. More recent microarray (Raman et al. 2011) and proteomic analysis (Raman et al. 2009) assessing gene and protein expression patterns, have also been performed on strain DSM 1237. The great wealth of research on *C. thermocellum* has allowed for cross-comparison and validation of results.

1.9 Thesis objectives

As the literature review has suggested, biofuel production yields of fermentative bacteria, particularly those capable of consolidated bioprocessing (i.e. *C. thermocellum*), are less than optimal. Production yields, however, may be improved by manipulation of fermentation conditions or through targeted metabolic engineering strategies which require thorough understanding of the complex metabolic and regulatory networks that dictate metabolic flux. Therefore, the broad objective of this thesis is to determine how gene content, gene product expression, and enzyme cofactor specificity can influence metabolic flux and end-product yields in response to growth conditions, using *C. thermocellum* as a platform organism. It was hypothesized that:

1. Fermentation conditions can influence end-product yields
2. Changes in end-product yields are dictated by enzyme expression levels
3. Genome content dictates end-product synthesis patterns
4. Distinct genes that encode proteins with analogous functions are differentially expressed
5. Protein expression patterns dictate carbon and electron flux

The specific objectives are:

1. To elucidate enzymes activities involved in pyruvate catabolism and end-product synthesis, with particular emphasis on alcohol dehydrogenases and hydrogenases (Chapters 2, 3, and 4)
2. To demonstrate that end-product yields are affected by fermentation conditions including:
    a. Addition of exogenous end products prior to fermentation (Chapter 3)
    b. Fermenter sparging with H_2 or N_2 at different rates (Chapter 4)
    c. Manipulation of substrate oxidation state (addition of pyruvate) (Chapter 7)
    d. Use of protein inhibitors including hypophosphite (pyruvate:formate lyase inhibitor) of carbon monoxide (H_2ase inhibitor) (Chapter 7).
3. To assess the impact of genome content on end-product patterns (Chapter 5)
4. To elucidate absolute protein expression profiles of proteins involved in core metabolism using:
    a. Two dimensional high pressure liquid chromatography tandem mass spectrometry (2D-HPLC-MS/MS) proteomics using peptides labeled with isobaric tags for relative and absolute quantitation (iTRAQ) (Chapter 6)
    b. Multiple reaction monitoring (MRM) mass spectrometry (Chapter 7)
5. To determine if gene or gene-product expression profiles are responsible for changes in end-products using:
    a. Cell extract enzyme activity assays (Chapters 3 and 4)

b. Reverse transcription quantitative PCR (RT-qPCR) (Chapter 4)

c. 2D-HPLC-MS/MS proteomics using iTRAQ labeled peptides (Chapter 6)

d. MRM mass spectrometry (Chapter 7)

1.10 Thesis outline

- **Chapter 1: Literature review**
- **Chapter 2: Growth phase-dependent enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405**
    - **Published:** *Journal of Biotechnology* 140, no. 3 (2009): 169-175.
    - **Highlights:** End-product synthesis and enzyme activities involved in pyruvate catabolism, H_2 synthesis, and ethanol production were assayed during growth.
- **Chapter 3: End-product induced metabolic shifts in *Clostridium thermocellum* ATCC 27405**
    - **Published:** *Applied Microbiology and Biotechnology* 92, no 1 (2011): 199-209.
    - **Highlights:** Growth, final end-product concentrations, and late-exponential phase enzyme activities involved in pyruvate catabolism and end-product formation were studied in response to addition of individual end-products to growth medium.

- **Chapter 4: Role of Transcription and Enzyme Activities in Redistribution of Carbon and Electron Flux in Response to N_2 and H_2 Sparging in Open-Batch Cultures of *Clostridium thermocellum* ATCC 27405**
    - **Submitted:** *Applied Microbiology and Biotechnology*
    - **Highlights:** Effect of high N_2, low N_2, and high H_2 gas sparging growth, end-product synthesis, enzyme activities, and transcription of select genes associated with the "malate shunt", pyruvate catabolism, H_2 synthesis, and ethanol production was studied
- **Chapter 5: Linking genome content to biofuel production yields: a meta-analysis of major catabolic pathways among select H_2 and ethanol-producing bacteria.**
    - **Published:** *BMC Microbiology* 12, no. 1 (2012): 295
    - **Highlights:** A meta-analysis of genes encoding pyruvate metabolism and end-product synthesis pathways was performed on select H_2 and/or ethanol producing Firmicutes, Euryarchaeota, and Thermotoga to demonstrate how gene composition can influence end-product yields.
- **Chapter 6: Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression**
    - **Published:** *BMC Microbiology* 12, no 1 (2012):214
    - **Highlights:** Shotgun 2D-HPLC-MS/MS was performed on exponential phase cultures to determine relative protein expression profiles of core metabolic proteins involved in carbohydrate utilization, energy conservation, and end-product synthesis. iTRAQ based protein quantitation was used to determine

changes in protein expression profiles in response to growth phase (exponential vs. stationary).

- **Chapter 7: Redirection of pyruvate flux in *Clostridium thermocellum* through manipulation of substrate oxidation states and use of enzyme inhibitors**
    - **Submitted:** *Applied Microbiology and Biotechnology*
    - **Highlights:** Use of pyruvate as a co-carbon source and enzyme inhibitors (hypophosphite and carbon monoxide) were used to manipulate end-product yields. Potential electron fluxes involved in fermentative pathways were proposed based on MRM protein expression profiles.

Chapter 2: Growth Phase Dependant Enzyme Profile of Pyruvate Catabolism and End-Product Formation in *Clostridium thermocellum* ATCC 27405[2]

2.1 Abstract

End-product synthesis and enzyme activities involved in pyruvate catabolism, H_2 synthesis, and ethanol production in mid-log (OD_{600} ~ 0.25), early stationary (OD_{600} ~ 0.5), and stationary phase (OD_{600} ~ 0.7) cell extracts were determined in *Clostridium thermocellum* ATCC 27405 grown in batch cultures on cellobiose. CO_2, H_2, ethanol, acetate and formate were major end-products and their production paralleled growth and cellobiose consumption. Lactate dehydrogenase, pyruvate:formate lyase, pyruvate:ferredoxin oxidoreductase, methyl viologen-dependant hydrogenase, ferredoxin-dependant hydrogenase, NADH-dependant hydrogenase, NADPH-dependant hydrogenase, NADH-dependant acetaldehyde dehydrogenase, NADH-dependant alcohol dehydrogenase, and NADPH-dependant alcohol dehydrogenase activities were detected in all extracts, while pyruvate dehydrogenase and formate dehydrogenase activities were not detected. All hydrogenase activities decreased (2 to 3-fold) as growth progressed from early exponential to stationary phase. Alcohol dehydrogenase activities fluctuated only marginally (< 45%), while lactate dehydrogenase, pyruvate:formate lyase, and

[2] Contributing authors: Thomas Rydzak, David B. Levin, Nazim Cicek, and Richard Sparling. *Journal of Biotechnology* 140, no. 3 (2009): 169-175.

pyruvate:ferredoxin oxidoreductase remained constant in all cell extracts. We have proposed a pathway involved in pyruvate catabolism and end-product formation based on enzyme activity profiles in conjunction with bioinformatic analysis.

2.2 Introduction

Clostridium thermocellum carries out mixed-product fermentation, synthesizing H_2 and ethanol, as well as CO_2 and a number of organic acids including acetate, formate, and lactate (Islam et al. 2006; Sparling et al. 2006). In order to develop strategies to enhance H_2 or ethanol production by *C. thermocellum*, a greater understanding of the metabolic and genetic mechanisms by which H_2 or ethanol is synthesized is required. Bioinformatic analysis has identified genes encoding key enzymes involved in pyruvate catabolism and H_2, ethanol, and acetate synthesis (Markowitz et al. 2006; Oak Ridge National Laboratory 2007). They include 2 putative lactate dehydrogenases (*ldh*; Cthe_1035 and Cthe_0345[3]), 1 putative pyruvate:formate lyase (*pfl*; Cthe_505), at least 4 putative PFL-activating enzymes (*fnr*; Cthe_0506, Cthe_0647, Cthe_1167, Cthe_1578), 3 putative pyruvate:ferredoxin oxidoreductase (PFOR)-like enzymes[4] (*pfor*; Cthe_2390-2393, Cthe_2794-2797, Cthe_3120), 2 putative NADH-dependant HydABC-type hydrogenases (*hyd*; Cthe_0340-0342 and Cthe_0428-0430), a putative 2 subunit hydrogenase, that may be NADPH-dependent (Cthe_3003-3004), and 1 putative ferredoxin-dependant Ech-type hydrogenase (*ech*; Cthe_3019-3024). The membrane

[3] Later identified as a malate dehydrogenase (Chapters 4 and 5)

[4] Two additional putative PFOR-like protein encofing genes (Cthe_0866, Cthe_0614) were later identified (Chapters 4 and 5)

bound Ech-type hydrogenase can pump out protons to generate an electrochemical membrane potential when coupled to ferredoxin-dependant H_2 production (Soboh et al. 2004). A putative *Rhodobacter* nitrogen fixation (*rnf*)-type NADH:ferredoxin oxidoreductase (*nfor*; Cthe_2430-2435) can potentially shuttle electrons between ferredoxin and NADH. An aldehyde dehydrogenase (*aldH*; Cthe_2238) and a number of putative Fe-dependant alcohol dehydrogenases (*adh*; Cthe_0101, Cthe_0394, Cthe_0423[5], Cthe_2579) can be used for ethanol synthesis, while a putative phosphotransacetylase (*pta*; Cthe_1029), acetate kinase (*ack*; Cthe_1028), and acetate thiokinase (*atk*; Cthe_0551) can be involved in acetate synthesis and ATP generation (Carere et al. 2008a). Although transcription of many of these genes has been confirmed by reverse transcriptase polymerase chain reaction (RT-PCR) throughout growth on α-cellulose (Carere et al. 2008a), little is known about protein expression levels, enzyme activities, and how enzyme expression and activities can influence metabolic flux or respond to end-product accumulation during growth in strain ATCC 27405. Hence, we set out to identify activity profiles of enzymes involved in pyruvate catabolism, H_2 production, and alcohol production in relation to end-product synthesis and growth phase under cellobiose-limited conditions.

2.3 Materials and methods

2.3.1 Organism, media, and growth conditions

[5] Later characterized as *adhE* (a bifunctional acetaldehyde/alcohol dehydrogenase)

Clostridium thermocellum ATCC 27405 (a.k.a. DSM 1237) purchased from the American Type Culture Collection was employed for all growth experiments. Fresh cultures were maintained by routinely transferring 10% (v/v) mid-exponential phase inoculum into complex 1191 medium (Islam et al. 2006) containing 1.1 g L^{-1} (5.9 mM) cellobiose. This complex medium contained (per liter of distilled deionized water): KH_2PO4, 1.5 g; $Na_2HPO_4•12H_2O$, 4.2 g; NH_4Cl, 0.5 g; $MgCl_2•6H2O$, 0.18 g; yeast extract (BD 212750), 2.0 g; resazurin (0.25 mg/ml), 1.0 ml; vitamin solution, 0.50 ml; and mineral solution, 1.00 ml. Vitamin solution contained the following (per 1,000 ml): biotin, 20 mg; p-aminobenzoic acid, 50 mg; folic acid, 20 mg; nicotinic acid, 50 mg; thiamine, 50 mg; riboflavin, 50 mg; lipoic acid (thioctic acid), and 50 mg; cyanocobalamin, 10 mg. Mineral solution contained (grams per liter): trisodium nitrilotriacetate 20.2; $FeCl_3•6H_2O$, 2.1; $CoCl_2•6H_2O$, 2.0; $MnCl_2•4H_2O$, 1.0; $ZnCl_2$, 1.0; $NiCl_2•6H_2O$, 1.0; $CaCl_2•2H_2O$, 0.5; $CuSO_4•2H_2O$, 0.5; and $Na_2MoO_4•2H_2O$, 0.5. Cultures were grown at 60ºC and stored anaerobically at 4ºC. All chemicals were reagent grade and were obtained from Sigma Chemical Co (St. Louis, MO). N_2 gas was purchased from Welder's Supply (Winnipeg, MB).

Batch experiments were carried out anaerobically in sealed Corning bottles (1.1 L) and Balch tubes (26 ml; Bellco Glass Inc., Vineland, NJ), containing 500 ml and 10 ml of media respectively adjusted to pH 7.2. Media preparation and inoculation protocols were followed as described by Islam *et al.* (Islam et al. 2006). All tubes or bottles containing media were air-sealed with butyl rubber stoppers and sealed with aluminum seals or copper wire, respectively, and degassed:gassed (4:1 min) four times with N_2 (100%). Media were then reduced with Na_2S (to a final concentration of 2 mM) and autoclaved.

Tubes were then injected with filter sterilized, anaerobic cellobiose for a final concentration of 1.1 g L^{-1}. All tubes/bottles were inoculated with 10% of freshly growing, early stationary phase (OD_{600} ~ 0.500) culture grown on 1.1g L^{-1} cellobiose. All batch cultures were incubated at 60°C until stationary phase. End-products, cell biomass, and pH were measured throughout growth. Mid-exponential (OD_{600} ~ 0.25), early stationary (OD_{600} ~ 0.5), and stationary (OD_{600} ~ 0.7) phase cell extracts were used for enzyme activity analysis.

2.3.2 Analytical procedures

2.3.2.1 *Growth, cell biomass, and pH measurements*

Cell growth was measured spectrophotometrically (Biochrom, Novaspec II) at 600 nm. Samples (1.5 ml) were centrifuged (14000 x g, 10 min) and supernatant containing soluble end products was removed. Pellets were washed with 0.9% (w/v) NaCl and resuspended in 1 ml of 0.2 M NaOH. Suspensions were incubated for 10 min in a boiling water bath. After cooling, samples were re-centrifuged (14000 x g, 10 min) and the supernatants were collected for protein assays. Protein content of samples was determined using a scaled down version of the Bradford method (Bradford 1976) for use in a microtiter plate reader. Protein standards were prepared using bovine serum albumin and all absorbance readings were performed at 595 nm using a PowerWave-XS single channel spectrophotometer using KCjunior software (BIO-TEK Instruments Inc., Winooski, VT). The elemental biomass composition ($C_4H_7O_2N$) based on a stoichiometric conversion of cellobiose into cell material was calculated from protein content (Guedon et al. 1999b). A molecular weight of 101 g mol^{-1}, corresponding to the

average composition of cell material ($C_4H_7O_2N$), was used to calculate cell biomass in moles. The pH of each sample was measured directly from the culture supernatant with a model AP62 pH/mV meter (Fisher Scientific, Ottawa, ON) equipped with a needle probe.

2.3.2.2 *Gas measurements*

Product gas composition (H_2 and CO_2) was measured using a Multiple Gas Analyzer #1 Gas Chromatograph System Model 8610-0070 (SRI Instruments, Torrance, CA) with a thermal conductivity detector capable of detecting concentrations between 200-500 ppm, and using argon as the carrier gas. H_2 measurements were conducted with a stainless steel (3.2 mm x 1.8 m) column packed with a 13x molecular sieve, whereas CO_2 was analyzed using a stainless steel (3.2 mm x 1.8 m) column packed with a silica gel. Gas concentrations were quantified by comparing peak areas to those of a standard curve prepared on the same day from known concentrations of H_2 and CO_2. Gas production was then calculated taking temperature and atmospheric pressure into account. All gas measurements were corrected by calculating their solubility in water (Sander 1999) and the bicarbonate equilibrium was taken into account for CO_2 quantification (Darrett and Drisham 1995).

2.3.2.3 *Sugar and organic end-product analysis*

Samples (1 ml) of culture supernatant were collected and stored at -20°C until assayed. Cellobiose and glucose were measured by high-pressure liquid chromatography (Dionex, ICS 3000) with an anion-exchange CarboPac-PA1 analytical column (4×250 mm). Lactate, acetate, formate, and extracellular pyruvate were measured using an IonPac AS11-HC anion-exchange column (Dionex Corporation, Sunnyvale, CA).

Ethanol concentrations were determined using the Enzymatic BioAnalysis UV-Test kit (R-Biopharm AG, Darmstadt, Germany) based on NAD^+ produced by alcohol dehydrogenase, which was measured spectrophotometrically at 340 nm. Standards for acids, sugars, and ethanol were prepared on 1191 medium to correct for the background.

2.3.3 Preparation of cell extracts

Anaerobic conditions were maintained throughout the entire procedure, and all manipulations were performed at 4°C under N_2 atmosphere in a Coy anaerobic glove chamber (DiaMed Laboratory Supplies, Mississauga, ON). Cells were grown in 1 litre Corning bottles (500 ml medium) and anaerobically harvested by centrifuging (8000 x g, 20 min) in sealed centrifuge bottles (Nalgene, Rochester, NY). Supernatant was discarded and pellets were resuspended into Balch tubes containing 10 ml cell suspension buffer (HEPES, 50 mM, pH 7.0; DTT, 10mM; glycerol, 10%) or reactivation buffer (potassium phosphate buffer, 50 mM, pH 7.0; S-adenosyl-L-methionine, 0.6 mM; pyruvate, 10 mM, DTT, 2.85 mM; $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$, 2 mM; methyl viologen, 1 mM) (Yamamoto et al. 2000) and were sealed in the glove bag, then re-gassed with N_2 to ensure anoxic conditions. Cells resuspended in cell suspension buffer were used immediately or stored at -70°C, while cells resuspended in reactivation buffer were used immediately. Cells were lysed by passing the suspension three times through a 2 mM Na_2S rinsed French pressure cell at 20,000 kg m^{-2}. Suspensions were centrifuged (8000 x g, 20 min) to remove unlysed cells and supernatants were collected in Balch tubes, sealed with a rubber stopper, and re-gassed with N_2. Protein content of extracts was measured by the Bradford method as described above.

2.3.4 Enzyme assays

All assays were performed at 25°C under anaerobic conditions in stoppered pyrex tubes (Daniels and Wessels 1984), despite evidence that LDH, ADH, and AldH are oxygen tolerant (data not shown). Extracts suspended in cell suspension buffer were used for all assays with the exception of pyruvate:formate lyase, for which extracts in reactivation buffer were used after a 90 minute incubation period at 37°C (Sparling et al. 2006). Specific activities were determined from the linear part of the reaction curve and were reported in nmol min^{-1} mg $protein^{-1}$ representing conversion of substrate into specific products. One unit of enzyme activity is defined as the amount of enzyme that catalyzes the conversion of 1 nmol of substrate per minute. Formation of reduced methyl viologen, NAD(P)H, and metronidazole were measured spectrophotometrically at 560 (ε = 8.0 mM^{-1} cm^{-1}), 340 (ε = 6.22 mM^{-1} cm^{-1}), and 340 nm (ε = 9.3 mM^{-1} cm^{-1}), respectively. Each assay was done in triplicate on a minimum of two different extracts. Standard deviations were low (<5% deviation) for activities obtained within the same extract but were higher between different extracts (shown in brackets in Table 2.3). All activities were measured by modifications to standard assay methods outlined by Sridhar *et al*. and Soboh *et al*. and were determined in the thermodynamically favorable direction. The reaction mixtures (2.3-2.7 ml total volume) used for different assays contained 0.3-2 mg cell free extract protein (Sridhar et al. 2000; Soboh et al. 2004). Assay conditions are outlined in Table 2.1.

Table 2.1: Enzyme assay conditions

Enzyme Name (enzyme classification number)	Assay Condition
Pyruvate catabolism	
Lactate dehydrogenase (FDP[1]-activated) (EC 1.1.2.3)…….	0.1 M Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 0.25 mM NADH, 10 mM pyruvate, 1 mM fructose-1,6-bisphosphate
Pyruvate:formate lyase (EC 2.3.1.54)…….	0.1 M Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 5 mM lactate, 50 mM pyruvate, 0.2 mM CoA, 5 mM malate, 1 mM NAD^+, 4U citrate synthase, 20U malate dehydrogenase
Pyruvate:ferredoxin oxidoreductase (EC 1.2.7.1)……...	50 mM KPO_4, pH 7.0, 20 mM dithiothreitol, 0.1 CoA, 10 mM methyl viologen, 5 mM pyruvate
Ethanol production	
Acetaldehyde dehydrogenase (EC 1.2.1.10)……...	0.1 M Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 0.3 mM $NAD(P)^+$, 0.1 mM CoA, 10 mM acetaldehyde, 10 mM sodium arsenate, 0.5 U phosphotransacetylase
Alcohol dehydrogenase (EC 1.1.1.12)……	0.1 M Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 0.3 mM NAD(P)H, 10 mM acetaldehyde
Hydrogen production	
Hydrogenase (MV-dependant) (EC 1.12.7.1)…….....	0.1 Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 100% H_2, 20 mM methyl viologen
Hydrogenase ($NAD(P)^+$-dependant) (EC 1.12.1.2; EC 1.12.1.3)…..	0.1 Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 100% H_2, 75 mM $NAD(P)^+$
Hydrogenase (Fd-dependant) (EC 1.12.7.2)……	0.1 Tris-hydrochloride, pH 6.5, 10 mM dithiothreitol, 0.2 mM metronidazole, 0.02 mM ferredoxin

[1]FDP; fructose-1,6-bishosphate

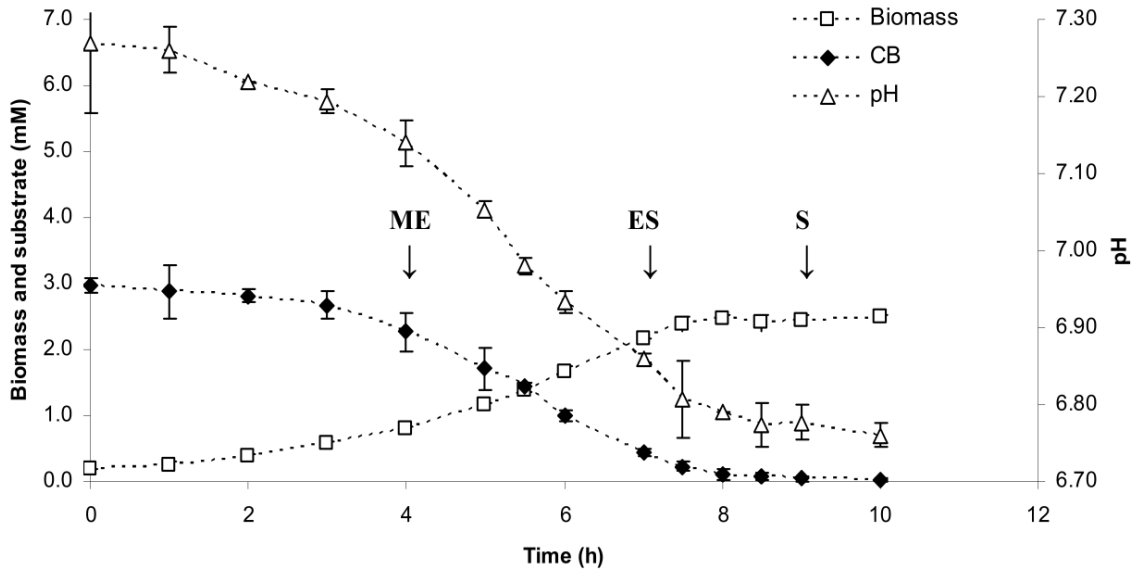
2.4 Results and discussion

2.4.1 Growth, pH, and cellobiose consumption

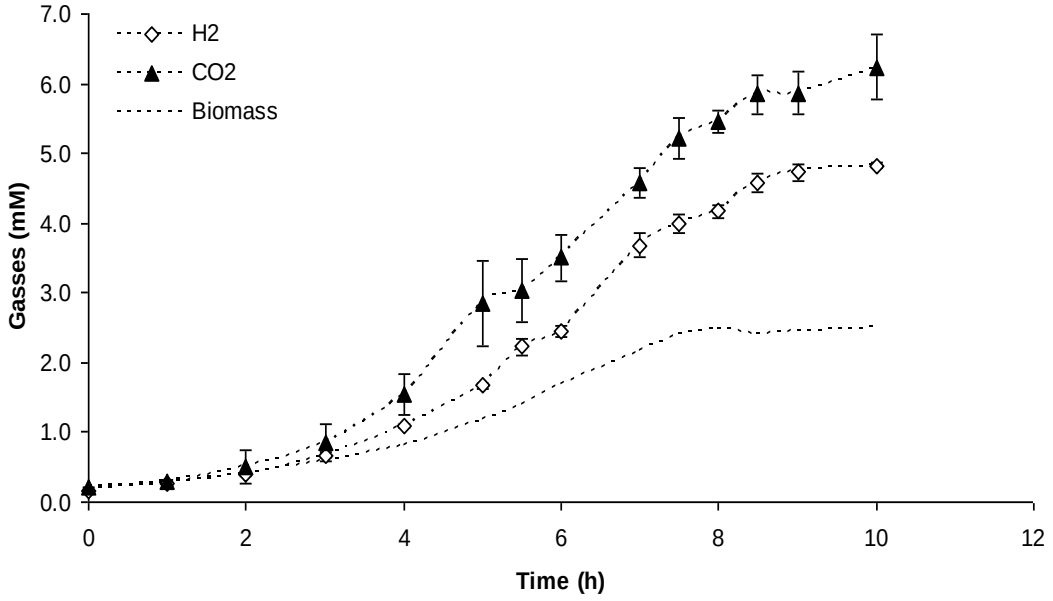
Growth of *C. thermocellum* ATCC 27405 was monitored in batch cultures containing 1.1 g L^{-1} (5.9 mM) cellobiose in 1191 medium with an initial pH of 7.2 (Figure 2.1a). No significant lag was observed when fresh, mid-exponential cells were used as inoculum, although lag of 2-3 h was observed with older and stationary phase cultures (data not shown). Cell densities measured at 600 nm (not shown) mirrored biomass production as measured by protein content. Final cell biomass concentrations reached 2.5 mM after 8 h of growth and corresponded to an OD_{600} of 0.69 (±0.06). All detectable cellobiose initially present in the medium was utilized upon entry into stationary phase (Figure 2.1a) indicating that growth was carbon limited. Generation time during mid exponential growth was calculated to be 2.4 h per generation. Culture pH decreased to a final pH of 6.7 reflecting production of acidic end products during growth (Figure 2.1a). Growth and pH data are in accordance with growth experiments previously performed by Islam *et al*. (Islam et al. 2006).

Figure 2.1: Fermentation kinetics of *Clostridium thermocellum* grown at 60º C on 1.1 g L^{-1} cellobiose in 1191 medium at an initial pH of 7.2. (a) Biomass, cellobiose (CB) consumption, and associated pH change, (b) gas, and (c) organic end-products are given in mM. Standard deviations are indicated with error bars. Time points selected for determination of enzyme activities corresponding to mid-exponential (ME), early stationary (ES), and stationary (S) phase are indicated by arrows (a). Carbon recoveries were 96, 84, and 92% and oxidation/reduction ratios of single carbon (CO_2 and formate) versus two carbon (acetate and ethanol) end-products (O/R) were 1.12, 1.02 and 1.05 in mid-exponential, early stationary, and stationary phase respectively.

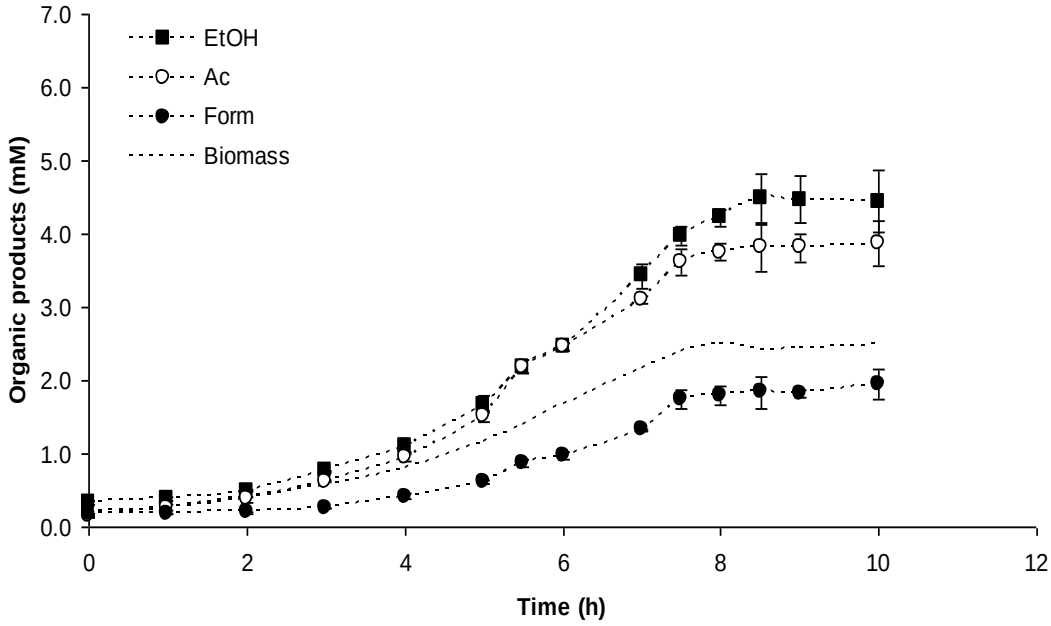
(a)



(b)


H2
CO2
Biomass
Gasses (mM)
Time (h)
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
0
2
4
6
8
10
12

(c)


EtOH
Ac
Form
Biomass
Organic products (mM)
Time (h)
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
0
2
4
6
8
10
12

2.4.2 End-product synthesis

Typical end-product formation kinetics during cellobiose fermentation are shown in Figures 2.1 b and c. CO_2, H_2, ethanol, acetate and formate were major end products and paralleled growth and cellobiose consumption. Only a small decrease in culture pH was observed (0.54) and thus no apparent metabolic shift from acid to alcohol synthesis, as found in *Clostridium acetobutylicum*, or from acetate/H_2 production to ethanol, lactate, and extracellular pyruvate production, as found in *Clostridium cellulolyticum*, was observed in response to decreasing culture pH (Grupe and Gottschalk 1992; Desvaux et al. 2001c). Final end-product yields (per mole glucose equivalent) and mid-exponential catabolite production rates are given in Table 2.2. Lactate production was below detectable limits (~ 0.2 mM). Other products such as extracellular pyruvate, butyrate, succinate, or other alcohols were not detected. Average carbon recovery, oxidation/reduction (O/R) value of oxidized to reduced end-products, and ratio of single carbon (CO_2 and formate) versus two carbon (acetate and ethanol) end-products (C_1/C_2) among different time points was 91%, 1.05, and 0.95 respectively confirming that all major end-products were accounted for.

Table 2.2: Final product yields, and mid-exponential phase catabolite production rates and cellobiose consumption rates.

	Final product yields (mol mol-glc^{-1})	Production/consumption rate (µmol µmol-cells^{-1} h^{-1})
Biomass	0.42	—
H_2	0.82	1.86
CO_2	1.06	2.16
Ethanol	0.76	1.78
Acetate	0.66	1.62
Formate	0.33	0.78
Lactate[1]	0.00	—
Cellobiose	—	1.29

[1] Lactate production was below detectable limits (~0.2 mM)

The fermentation pattern reported in our study resembled end-product concentrations of *C. thermocellum* in other studies, where ethanol and acetate were major organic end-products respectively, and lactate was only observed in trace amounts (Lynd et al. 1989; Strobel 1995; Stevenson and Weimer 2005). H_2 and CO_2 production was not reported in these studies, as cells were cultured in an open system with gas sparging. Batch cultures of *C. thermocellum* LQRI (a.k.a DSM 2360) and AS39, a mutant strain of ATCC 27405 displaying increased cellulase activity, reported substantial H_2 and CO_2 production accompanying ethanol, acetate, and trace amounts of lactate (Lamed and Zeikus 1980a). More recently, formate was also reported as a major end-product (Islam et al. 2006; Sparling et al. 2006). Although end-product concentrations reported in this study were comparable to those reported by Islam *et al.* and Sparling *et al.*, our study indicated greater CO_2 and ethanol production and lower H_2 production (Islam et al. 2006; Sparling et al. 2006). This may have resulted from differences in gas partial pressures during gassing of tubes.

Final end-product ratios and bioinformatic analysis of the genome sequence (http://genome.ornl.gov/microbial/cthe/) (Oak Ridge National Laboratory 2007), provided considerable insight on possible enzymes involved in pyruvate metabolism and end-product synthesis. Assuming that H_2 and CO_2 were formed via PFOR and ferredoxin-dependant hydrogenase alone, we would expect a 1:1 CO_2 to H_2 ratio. Our observed ratio of 1/0.77 suggested the presence of either a NAD(P)H-dependant hydrogenase or a NAD(P)H:ferredoxin oxidoreductase capable of transferring electrons from ferredoxin to $NAD(P)^+$ for ethanol production. In the absence of these enzymes we would also expect a 1:1 ethanol to acetate ratio (in the absence of lactate production)

given that each glucose molecule generates 2 acetyl-CoA regardless of pathway (via PFL or PFOR), and 2 NADH via the Embden-Meyerhof-Parnas pathway (Patni and Alexander 1971a; Patni and Alexander 1971b) which must be subsequently reoxidized. Thus, one acetyl-CoA may be converted into an ethanol molecule, during which the 2 NADH are reoxidized, while the other may be converted into acetate producing an ATP. Experimentally, however, our ethanol to acetate ratio is 1.0/0.87 further supporting the presence of an enzyme capable of utilizing H_2 for the production of NADH used for ethanol production. This is thermodynamically feasible given that the standard reduction potentials ($E^{o''}$) of H_2 and NADH are -414 mV and -320 mV respectively, and therefore production of NADH from H_2 is more favorable under standard conditions than H_2 production from NADH (Thauer et al. 1977; Angenent et al. 2004). In fact, growth of *C. thermocellum* under elevated hydrostatic pressures during which media were supersaturated with dissolved H_2 and CO_2 resulted in the channeling of electrons to NADH and thus increased ethanol production and decreased acetate production known to accompany H_2 production (Bothun et al. 2004).

2.4.3 Enzyme activity involved in pyruvate metabolism and end-product synthesis

Enzyme activities involved in the catabolism of pyruvate and formation of H_2 and ethanol were assayed in cell extracts obtained from cells harvested in three different growth phases; mid-exponential (OD_{600} = 0.25), early stationary (OD_{600} = 0.5), and stationary (OD_{600} ~ 0.7). Activities (nmol min^{-1} mg-$protein^{-1}$) and assay conditions are reported in Table 2.3 and represent averages obtained from extracts obtained from a particular growth phase on at least two separate occasions.

Table 2.3: Enzymes activities of *C. thermocellum* involved in pyruvate catabolism and hydrogen production

Enzyme	Substrate	Specific Activity (nmol min^{-1} mg $protein^{-1}$)[1]		
		Mid-exponential	**Early Stationary**	**Stationary**
Lactate dehydrogenase	Pyruvate: NADH: FDP	31 (11)	22 (6)	28 (3)
Pyruvate:formate lyase[2]	Pyruvate: CoA	ND[3]	19 (1)	19 (1)
Pyruvate:fd oxidoreductase	Pyruvate: CoA :MV	444 (203)	355 (25)	388 (7)
Hydrogenase	H_2: MV	2157 (445)	ND[2]	702 (18)
	H_2: Fd: metronidazole	3 (1)	2 (1)	2 (<1)
	H_2: NAD^+	4 (1)	2 (1)	1 (1)
	H_2: $NADP^+$	33 (11)	9 (<1)	11 (1)
Alcohol dehydrogenase	Acetaldehyde: NADH	16 (2)	11 (3)	9 (2)
	Acetaldehyde: NADPH	13 (1)	9 (1)	15 (1)
Acetylaldehyde dehydrogenase	Acetaldehyde: CoA: NAD^+	22 (2)	23 (3)	21 (3)

[1]Detection limits are 1.3, 0.4, and 0.3 nmol min^{-1} mg $protein^{-1}$ for assays spectrophotometrically monitoring redox changes in methyl viologen (MV), NAD(P)H, and metronidazole respectively.

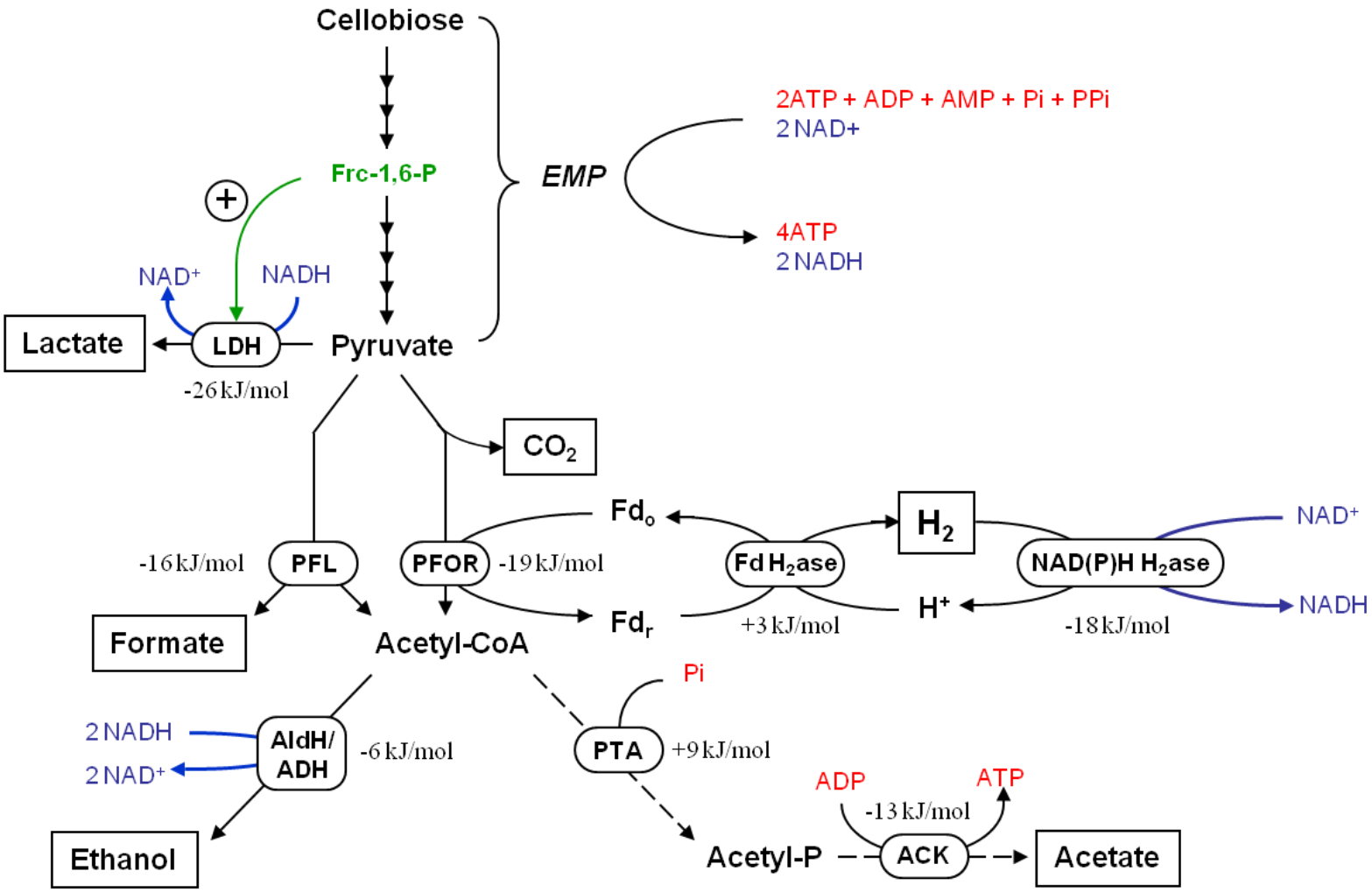
[2]Specific activity was not determined for cell extracts obtained from early exponential phase

[3]Not determined

2.4.3.1 *Pyruvate/acetyl-CoA/lactate branchpoint*

LDH, PFL, and PFOR activities were detected in all extracts (Table 2.3) thus suggesting that pyruvate can be metabolized into (i) lactate, (ii) formate and acetyl-CoA, and (iii) CO_2, reduced ferredoxin, and acetyl-CoA respectively (Figure 2.2). Despite the presence of NADH-dependant LDH activity, no lactate was produced by *C. thermocellum* during growth. Of the two putative LDHs, only the one encoded by *ldh* Cthe_1053 has been cloned and characterized. It was shown to be allosterically activated by fructose-1,6-bisphosphate (FDP), and thus low intracellular FDP levels may explain the lack of lactate production (Özkan 2004). In our study LDH activity was detected in the absence of exogenous fructose-1,6-bisphosphate, but increased four-fold in its presence (data not shown). Quantitative polymerase chain reaction (qPCR) has also shown that genes encoding putative LDHs Cthe_0345 and Cthe_1053 are transcribed (Carere et al. 2008a). Interestingly, the L-lactate dehydrogenase family includes both LDH and malate dehydrogenase (Carere et al. 2008a). Furthermore, a single amino acid substitution in malate dehydrogenase can change substrate binding specificity from malate to lactate (Yin and Kirsch 2007). Thus distinguishing between these enzymes based on sequence homology is difficult. However, given that Cthe_0345 is located next to a malic enzyme suggests that it may be a malate dehydrogenase involved in malate metabolism.

Figure 2.2: Metabolic pathways involved in pyruvate catabolism and end-product synthesis in *Clostridium thermocellum* ATCC 27405 based on end-product synthesis, enzyme activities, and preliminary genome annotation. Thermodynamic efficiencies are given for each reaction. —— pathways verified by enzyme activity analysis; ------ genetically deduced pathway not verified by enzyme activity analysis. LDH, lactate dehydrogenase (fructose-1,6-bisphosphate activated); PFL, pyruvate:formate lyase; PFOR, pyruvate:fd oxidoreductase; AldH/ADH, acetaldehyde dehydrogenase/alcohol dehydrogenase; PTA, phosphotransacetylase; ACK, acetate kinase; Fd H_2ase, ferredoxin hydrogenase; NAD(P)H H_2ase, NAD(P)H hydrogenase.


Cellobiose
Frc-1,6-P
EMP
2ATP + ADP + AMP + Pi + PPi
2 NAD+
4ATP
2 NADH
NAD+
NADH
Lactate
LDH
Pyruvate
-26 kJ/mol
CO2
Fdo
H2
NAD+
-16 kJ/mol
PFL
PFOR
-19 kJ/mol
Fd H2ase
NAD(P)H H2ase
NADH
Fdr
H+
+3 kJ/mol
-18 kJ/mol
Formate
Acetyl-CoA
Pi
2 NADH
2 NAD+
AldH/
ADH
-6 kJ/mol
PTA
+9 kJ/mol
ADP
ATP
-13 kJ/mol
Ethanol
Acetyl-P
ACK
Acetate

Transcription of *pfor* Cthe_2392, *por* Cthe_2796, *por* Cthe_2392, and *pfl* Cthe_0505 has been previously confirmed by Carere *et al*. using RT-PCR (Carere et al. 2008a). Unfortunately, enzyme activity analysis cannot distinguish between different PFOR enzymes. In agreement with previous reports (Sparling et al. 2006), we found that the activity of PFOR was much greater than that of PFL (Table 2.2). While *in vitro* enzyme assays are not always indicative of *in vivo* activities, our PFOR and PFL activities were consistent with specific production rates of formate and CO_2 during exponential phase (Table 2.2), which are likely formed exclusively by PFOR and PFL respectively. Pyruvate dehydrogenase activity was not detected in cell extracts and thus NAD^+-dependent oxidative decarboxylation of pyruvate to acetyl-CoA is improbable.

2.4.3.2 ***Acetyl-CoA/ethanol/acetate branchpoint***

Acetyl-CoA catabolism can result in NADH reoxidation or ATP generation during production of ethanol or acetate, respectively. Bioinformatic analyses have confirmed the presence of genes encoding an NADH-dependant aldehyde dehydrogenase (*aldh* Cthe_2238), numerous Fe^{2+} and Zn^{2+} dependent alcohol dehydrogenases, and a fused bifunctional acetaldehyde/alcohol dehydrogenase (*adhE*, gene Cthe_0423) proposed to be the key enzyme involved in ethanol synthesis (Carere et al. 2008a). Transcription of *adhY*, *adhZ*, and *adhE* has been confirmed by RT-PCR (Stevenson and Weimer 2005; Carere et al. 2008a). We detected NADH-dependant AldH, as well as NADH and NADPH-dependant ADH activities in cell extracts from all three growth phases. NADH and NADPH-dependant ADH activities were comparable, and thus did not agree with results published by Lamed & Zeikus showing ≥200-fold higher NADH-

dependant ADH activities in comparison to NADPH-dependant ADH activities in *C. thermocellum* strains LQRI and AS39 (Lamed and Zeikus 1980a).

Acetate may be produced from acetyl-CoA either indirectly through an acetyl phosphate intermediate using phosphotransacetylase and acetate kinase, or directly via acetate thiokinase. Although both reactions produce ATP, the former uses ADP and Pi whereas the latter uses AMP and inorganic pyrophosphate (PP_i) as substrates for ATP synthesis. Genes encoding all three enzymes have been annotated, although only the expression of acetate kinase has been confirmed through RT-PCR and enzyme activity analyses (Lamed and Zeikus 1980a; Stevenson and Weimer 2005). Acetate kinase has also been purified and characterized by Lin *et al*. (Lin et al. 1998a). Bioinformatic analysis has also confirmed a gene encoding a putative vacuolar-type H^+-pyrophosphatase (gene Cthe_1425) capable of producing PP_i from inorganic phosphate (P_i) for acetate thiokinase. Thermodynamically, however, production of acetate via phosphotransacetylase and acetate kinase ($\Delta G^{o'} = -4$ kJ mol^{-1}) is more favorable than that of acetate thiokinase ($\Delta G^{o'} = +9$ kJ mol^{-1}).

2.4.3.3 ***Hydrogen generating pathways***

In addition to ethanol and lactate production, disposal of reducing equivalents can be accomplished via production of molecular H_2. In *C. thermocellum*, annotated hydrogenases include one ferredoxin-dependent Ech-type (Energy Converting Hydrogenase) [NiFe]-hydrogenase, two putative NADH-dependent Fe-only HydABC-type hydrogenase, and one putative NADPH-dependant HydABC-type Fe-only hydrogenase. Transcription of genes encoding select subunits of each hydrogenase has

been confirmed using RT-PCR by Carere *et al*. (Carere et al. 2008a). Activities of all three hydrogenases (ferredoxin, NADH, and NADPH-dependant) as well as high methyl viologen-dependant hydrogenase activity, representing the sum of all hydrogenase activities, were detected in cell extracts from all three growth phases (Table 3). Activities of ferredoxin-dependant and NADH-dependant activities were comparable but minimal, while NADPH-dependant hydrogenase activities were roughly 5 to 10-fold greater. Activity of formate:hydrogen lyase (FHL), which oxidizes formate to H_2 and CO_2 in enteric bacteria such as *E. coli* (Sawers 1994; Bagramyan and Trchounian 2003) was not detected in cell extracts. The gene previously thought to encode the fhlγ subunit of FHL (Sparling et al. 2012) has been re-annotated as part of a 2 subunit NADPH-dependent hydrogenase.

2.4.4 Effects of growth phase on enzyme activities

Of the enzymes involved in pyruvate catabolism, H_2 synthesis, and ethanol production, specific activities of enzymes involved in H_2 and ethanol synthesis were most influenced by growth phase. Methyl viologen-dependant hydrogenase specific activity was 3-fold higher in mid-exponential phase cell extracts than in stationary phase extracts. NADH and NADPH-dependant hydrogenase specific activity was also 2 to 4-fold higher in mid-exponential phase than in early stationary and stationary phase. A more subtle decrease in ferredoxin-dependant hydrogenase specific activity was observed as growth progressed.

Genes coding for both Ech and HydABC-type hydrogenases in *C. thermocellum* have high sequence similarity to those of *Thermoanaerobacter tengcongensis*, a non-

cellulolytic thermophile which also produces CO_2, ethanol, lactate, acetate and H_2. *T. tengcongensis* NADH-dependant Fe-only hydrogenase specific activity was 4-fold lower in the presence of increased H_2 partial pressures, while ferredoxin-dependant [NiFe] hydrogenase activity remained unchanged (Soboh et al. 2004). Thus, a similar mode of regulation of NADH-hydrogenase expression in response to increased H_2 partial pressures resulting from growth of sealed batch cultures may be involved in both *T. tengcongensis* and *C. thermocellum*.

Changes in specific activity of enzymes involved in ethanol production were less pronounced (< 2-fold differences), while those of enzymes involved directly in pyruvate catabolism (LDH, PFOR, and PFL) were negligible in response to growth. Early exponential phase NADH-dependant ADH specific activity was 44% higher than in stationary phase, while NADPH-dependant ADH specific activity was slightly lower in early stationary phase extracts than in mid-exponential and stationary phase extracts. NADH-dependant acetaldehyde dehydrogenase activity remained constant in extracts from different growth phases. No NADPH-dependent AldH activity was detected. Unfortunately, *C. thermocellum* encodes numerous AldHs and ADHs which have not been studied and characterized in terms of pyridine nucleotide and substrate specificity. Thus, we could not distinguish between changes in expression levels of the various annotated alcohol dehydrogenases based on enzyme activity data alone.

2.5 Conclusion

In order to develop strategies to enhance H_2 or ethanol production by *C. thermocellum*, a greater understanding of the metabolic and genetic mechanisms by

which H_2 is synthesized is required. We have proposed a metabolic pathway involved in pyruvate catabolism and end-product synthesis based on bioinformatic analysis and the presence of given enzyme activities (Figure 2.2). Significant changes in the specific activities of different hydrogenases were observed in extracts from cells harvested at different growth phases. Changes in ADH activities were more subtle, while all other activities remained constant. The presence of NADPH-dependent hydrogenase and alcohol dehydrogenase combined with an ethanol to acetate ratio greater than 1 suggests possible uptake of H_2 for ethanol production. The accumulation of end-products during growth may influence expression of these enzymes as well as carbon and electron fluxes through these pathways. Although enzyme activity profiles provide some insight into intracellular protein levels, quantitative PCR and proteomics must be employed to fully understand transcriptional, translational, and post-transcriptional regulation of these proteins. Furthermore, *in vitro* enzyme activities are not always indicative of *in vivo* activities, which may be influenced by intracellular concentrations of metabolites involved in allosteric regulation, ratios of $NADH/NAD^+$ and ATP/ADP, and internal pH, and thus measurement of internal metabolite pools may play a pivotal role in fully understanding carbon and electron fluxes during end-product formation.

2.6 Authors' contributions

Thomas Rydzak performed growth experiments, end-product analysis, enzyme activity analysis, and authored the manuscript. Richard Sparling, Nazim Cicek, and David B. Levin conceived of the study, participated in its design, and helped draft the manuscript.

2.7 Acknowledgements

This work was supported by funds provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944-04) and the BIOCAP Canada Foundation.

Chapter 3: End-Product Induced Metabolic Shifts in *Clostridium thermocellum* ATCC 27405[6]

3.1 Abstract

When attempting to increase yields of desirable end-products during fermentation, there is the possibility that increased concentrations of one product redirects metabolism towards the synthesis of less desired products. Changes in growth, final end-product concentrations, and activities of enzymes involved in pyruvate catabolism and fermentative end-product formation were studied in *Clostridium thermocellum* in response to the addition of individual end-products (H_2, acetate, ethanol, formate, and lactate) to the growth medium. These were added to the growth medium at concentrations ten times greater than those found at the end of growth in cultures grown under carbon limited conditions using cellobiose (1.1 g L^{-1}) as a model soluble substrate. Although growth rate and final cell biomass decreased significantly with the addition of all end-products, addition of individual end-products had less pronounced effects on growth. Metabolic shifts, represented by changes in final end-product concentrations, were observed; H_2 and acetate yields increased in the presence of exogenous ethanol and lactate, while ethanol yields increased in the presence of exogenous hydrogen (H_2),

[6] Contributing authors: Thomas Rydzak, David B. Levin, Nazim Cicek, and Richard Sparling. *Applied microbiology and biotechnology* 92, no 1 (2011): 199-209.

acetate, and lactate. Late-exponential phase enzyme activity data of enzymes involved in pyruvate catabolism and end-product formation revealed no changes in enzyme levels greater than 2-fold in response to the presence of any given end-product, with the exception of pyruvate:formate lyase (PFL), ferredoxin-dependent hydrogenase (Fd-H_2ase), and pyruvate:ferredoxin oxidoreductase (PFOR): PFL and Fd-H_2ase activity increased 2-fold in the presence of ethanol, while PFOR activity decreased by 57% in the presence of sodium formate. Changes in enzyme levels did not necessarily correlate with changes in final end-product yields, suggesting that changes in final end-product yields may be governed by thermodynamic considerations rather than levels of enzyme expressed under the conditions tested. We demonstrate that bacterial metabolism may be manipulated in order to selectively improve desired product yields.

3.2 Introduction

In fermentative organisms, 4 moles of H_2 per mole glucose can be expected during strict acetogenic fermentation:

$$C_6H_{12}O_6 + 2\ H_2O \rightarrow 4\ H_2 + 2\ CH_3COOH + 2\ CO_2 \qquad (\Delta G^{o\prime} = -213\ kJ\ mol^{-1})$$

Alternatively, 2 moles of ethanol per mole glucose would be expected during strict ethanologenic fermentation:

$$C_6H_{12}O_6 \rightarrow 2\ CH_3CH_2OH + 2\ CO_2 \qquad (\Delta G^{o\prime} = -218\ kJ\ mol^{-1})$$

C. thermocellum, however, like most other H_2 and ethanol producing fermentative bacteria, typically exhibits H_2 and ethanol yields <1 mol mol-glc^{-1}, respectively

(Angenent et al. 2004; Islam et al. 2006; Rydzak et al. 2009) due to the presence of branched product synthesis pathways[7]. The catabolism of pyruvate into end-products in *C. thermocellum* ATCC 27405 can be represented by two major branch points: (i) the *Pyruvate/Lactate/Acetyl-CoA* branch point where pyruvate can be catabolized into either lactate, formate and acetyl-CoA, or CO_2, acetyl-CoA and reduced ferredoxin used to generate H_2, and (ii) the *Acetyl-CoA/Ethanol/Acetate* branchpoint, in which acetyl-CoA is converted to either ethanol or acetate (Rydzak et al. 2009).

While H_2 can also be produced from NADH generated through glycolysis via the Emdben-Meyerhof pathways, the reoxidation of NADH via lactate or ethanol synthesis is more thermodynamically favorable than the reduction of protons to H_2, and thus decreases H_2 yields (Thauer et al. 1977). In order to increase H_2 production yields, two strategies may be employed. Redirecting carbon and electron flux from pyruvate:formate lyase (PFL) to pyruvate:ferredoxin oxidoreductase (PFOR) may potentially increase H_2 yields by increasing reduced ferredoxin pools. Although H_2 production from reduced ferredoxin is more thermodynamically favorable than from NADH, additional H_2 could also be produced from NADH. Thus, redirecting bacterial metabolism and reducing equivalents away from pathways that produce lactate and ethanol, during which NADH is reoxidized, may increase $NADH/NAD^+$ ratios, making H_2 production from NADH more thermodynamically favorable. However, increased $NADH/NAD^+$ ratios may also inhibit

[7] Refer to Figure 2.2 in section 2.4.3.1 above

carbon and electron flux through glycolysis (Lovitt et al. 1988), in turn reducing cell growth rate and thus H_2 production rates.

While several studies have shown changes in end-product yields in response to changes in H_2 partial pressures in clostridia and related species (Weimer and Zeikus 1977; Lamed and Zeikus 1980a; Ben-Bassat et al. 1981; Freier et al. 1988; Lamed et al. 1988; Bothun et al. 2004; Soboh et al. 2004), few have eluded to the effects of H_2 on the distribution of the entire suit of carbon end-products produced. Soboh *et al*. have proposed that changes in end-product yields in response to increased H_2 partial pressures are dictated by changes in enzyme levels (Soboh et al. 2004). There have been fewer studies reporting the effects of sub-inhibitory concentrations of ethanol or carboxylic acids on end-product yields (Lovitt et al. 1988; Zheng and Yu 2005). However, from a biofuels production perspective, whether in batch or continuous operations, fermentation will be expected to proceed in the presence of the highest possible concentrations of the desired end-products.

Given this precedent, we set out to determine the effect of elevated, but sub-inhibitory concentrations of H_2, ethanol, lactate, acetate, and formate on growth and final end-product concentrations of *C. thermocellum* grown in batch cultures under carbon-limited conditions using cellobiose (1.1 g L^-) as our model soluble carbon source. Activity profiles of enzymes involved in pyruvate catabolism and product formation were determined to deduce if changes in final product ratios are dependent on availability of enzyme activity or thermodynamic considerations only. The selective redirection of reducing equivalents (NAD(P)H, reduced ferredoxin) towards reduced end-products and

applications of exogenous end-product loading during bacterial fermentations to increase product yields are discussed.

3.3 Materials and methods

3.3.1 Organism, media, and experimental design

Clostridium thermocellum ATCC 27405 was employed for all growth experiments. Culture maintenance, media composition and preparation, and growth experiments were performed as previously described[8] by Rydzak *et al*. (Rydzak et al. 2009) with the following exceptions. While growth experiments were carried out on 1.1 g L^{-1} cellobiose, organic end-products (individually or together) were added to the medium prior to adjusting the pH and gassing. Final concentrations of added ethanol, sodium acetate, and sodium formate (36, 55, and 37 mM, respectively) were ten times greater than those found as end-products in late exponential phase on 1.1 g L^{-1} cellobiose as reported by Islam *et al*. (Islam et al. 2006). Final concentrations of added sodium lactate were 55 mM, corresponding to ten times the concentration of lactate found in late exponential phase on 4.5 g L^{-1} cellobiose, as no lactate was produced under carbon limited conditions (1.1 g L^{-1} cellobiose), (Islam et al. 2006). Tubes were gassed and degassed (1:4 min) four times with either N_2 (100%), H_2 (100%) or H_2:CO_2 mixture (80:20%), reduced, and autoclaved according to Daniels *et al*. (Daniels et al. 1986). In experiments using H_2:CO_2 gas mix, the medium was bubbled with the gas mix for about

[8] Refer to section 2.3.1 for details

an hour, and pH was subsequently adjusted to account for carbonate formation prior to further gassing. Samples for end-product analysis were taken at the beginning of stationary phase (OD_{600} ~ 0.7) while samples for enzyme activity analysis were taken in late-exponential phase (OD_{600} ~ 0.45).

3.3.2 Cell growth, pH, and end-product analysis

Cell growth measurements, sample processing, pH measurement, product gas, protein, sugar, and end-product analyses, and biomass and gas calculations were performed as previously described[9] (Rydzak et al. 2009). Data represents the average of a minimum of three biological replicates.

3.3.3 Preparation of cell extracts for enzyme assays

Cell samples for enzyme activity analysis were harvested anaerobically and cell extracts were made as previously described[10] (Rydzak et al. 2009) with the following exceptions. Given that the periplasmic hydrogenase in *Desulfovibrio gigas* was shown to be reactivated by H_2 and lower redox potential (Lissolo et al. 1984), cell extracts re-suspended in reactivation buffer were used to measure both pyruvate:formate lyase and all hydrogenase activities. Cell-free extract enzymes were reactivated in the presence of H_2. This was done by degassing and gassing (4:1 min) four times the cell free extract with H_2 (rather than N_2), incubating at 37°C for thirty minutes to ensure reactivation, and subsequently degassing and gassing with N_2. All other activities were measured using extracts of cells re-suspended in cell suspension buffer.

3.3.4 Enzyme assays

---

[9] Refer to section 2.3.2 for details

[10] Refer to section 2.3.3 for details

All assay conditions and specific activity calculations were performed as previously described[11] by Rydzak *et al*. (Rydzak et al. 2009). Each assay was done in triplicate on a minimum of two different extracts. Standard deviations were low (<5% deviation) for activities obtained within the same extract but were higher between different extracts (shown in brackets in Table 3.3).

3.4 Results

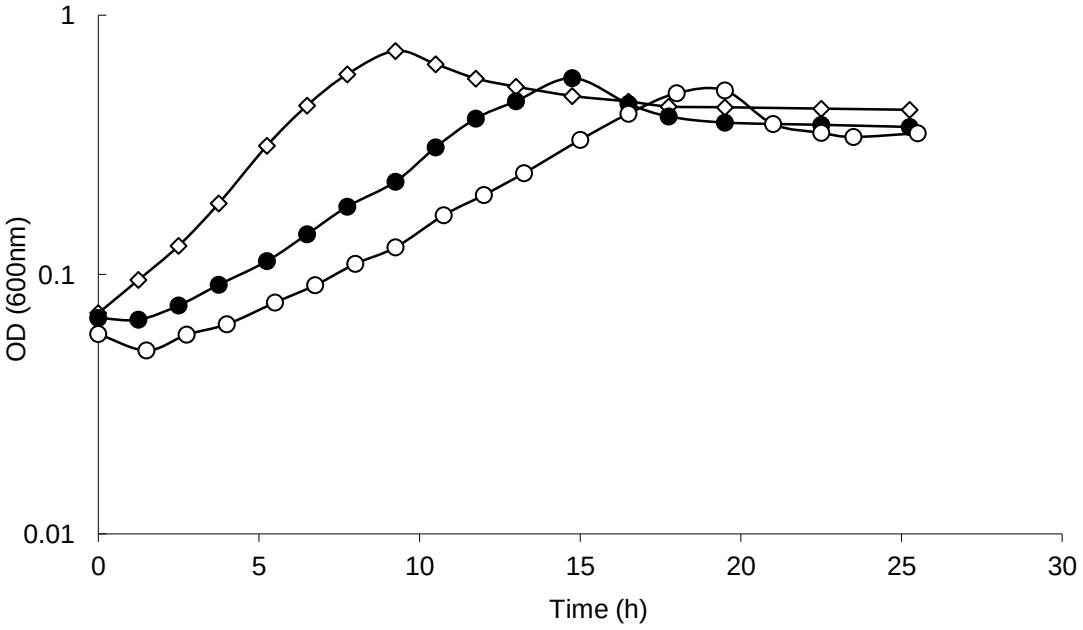
3.4.1 Effect of end-products on growth and final culture pH

Growth of *C. thermocellum* in batch cultures under carbon limited conditions (1.1 g L^{-1} cellobiose) and a N_2 (100 %) or H_2:CO_2 (80:20%) atmosphere was monitored in the presence of elevated concentrations of all organic end-products (ethanol, 36 mM; sodium acetate, 55 mM; sodium formate, 27 mM; and sodium lactate, 55 mM; Figure 3.1a). The presence of all organic end-products in cultures grown under N_2 atmosphere increased the generation time during exponential phase from 2.2 h to 4.2 h, whereas final OD_{600} decreased from 0.73 to 0.57 (Table 3.1). In the presence of a H_2:CO_2 (80:20%) head space the generation time further increased to 5.1 h and final OD_{600} decreased to 0.51 (Table 3.1).

[11] Refer to section 2.3.4 for details

Figure 3.1: Growth patterns of *C. thermocellum* cultured under carbon limited conditions in the absence and presence of exogenous end-products. Concentrations of end-products were added at 10x of that normally found in stationary phase cultures grown at 1.1 g L^{-1} cellobiose with the exception of lactate (added at 10x the concentration of cells found in stationary phase when grown on 4.5 g L^{-1} cellobiose) and gasses (added at percent concentrations as indicated). Growth in response to (a) addition of all organic end-products cultured under 100% N_2 or 80:20% H_2:CO_2 atmosphere versus addition of no organic end-products under 100% N_2 atmosphere and (b) addition of individual end-products cultured under N_2 atmosphere. No addition (◇), addition of all organic end-products under N_2 atmosphere (●), addition of all organic end-products under 80:20% H_2:CO_2 atmosphere (○), acetate addition (▲), formate addition (△), lactate addition (■), ethanol addition (□), 100% H_2 addition (◆). Arrow indicates sampling time for enzyme activity analysis (OD_{600} ~ 0.45).

a)


1
0.1
0.01
OD (600nm)
0
5
10
15
20
25
30
Time (h)

b)

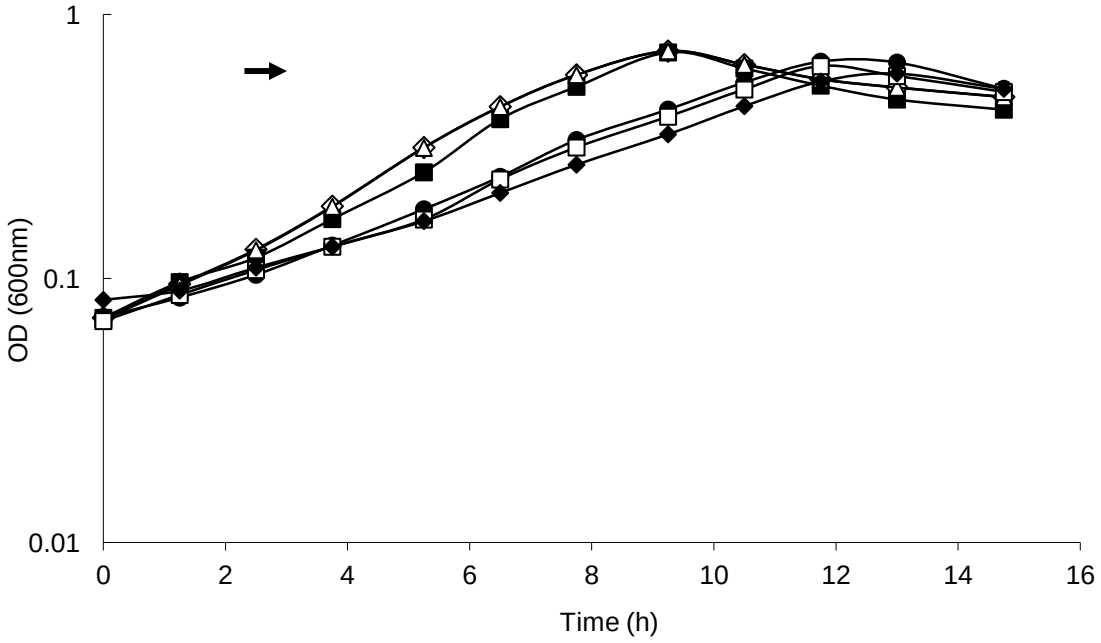

1
0.1
0.01
OD (600nm)
0
2
4
6
8
10
12
14
16
Time (h)

Table 3.1: Effect of added end-products on mid-exponential phase generation time, maximum OD_{600} reached, and final pH of *C. thermocellum* cultured in batch cultures under carbon limited conditions. Concentrations of additions are indicated in the table. Data represent three biological replicates. Standard deviations are ≤ 0.3, 0.02, and 0.05 for generation time, maximum OD_{600}, and final pH respectively.

Initial Media	Addition to	t/n (h/generation)	Max OD (600nm)	Final pH
No Addition		2.2	0.73	6.65
Organic End-Products		4.2	0.57	ND[1]
All End-Products		5.1	0.51	ND[1]
Acetate (55 mM)		2.4	0.72	6.63
Formate (27 mM)		2.2	0.73	6.60
Lactate (55 mM)		3.5	0.66	6.64
Ethanol (36 mM)		3.6	0.64	6.54
H_2 (100%)		3.7	0.60	6.55

[1]ND; not determined

Figure 3.1b illustrates the effect of individual organic end-products on growth of *C. thermocellum*. The presence of acetate or formate had no impact on final culture OD_{600} or generation time. In contrast, presence of lactate, ethanol, or H_2 increased generation time during exponential phase (3.5, 3.6, and 3.7 h, respectively) and slightly decreased final culture OD_{600} (0.66, 0.64, and 0.60, respectively). Despite the changes in final OD_{600}, the changes in final biomass concentration (Table 3.2) were of little statistical significance. The presence of end-products had no significant impact on final culture pH, which decreased from 7.2 to roughly 6.6 under all conditions as a result of acid production during growth (Table 3.1).

Table 3.2: Effect of exogenous end-products on final end-product concentrations of *C. thermocellum* grown under carbon limited conditions (1.1 g L^- cellobiose). Standard deviations of three biological replicates are indicated in brackets. Values indicated in bold represent calculated end-product concentrations based on an expected carbon recovery of ~90%, and an O/R index and C_1/C_2 ratio of 0.9-1.1 (shown in bold).

End-Product (mM)	**End-Product Added**					
	None	**H_2**	**Acetate**	**Ethanol**	**Formate**	**Lactate**
Biomass	2.4 (0.2)	2.4 (0.2)	2.4 (0.1)	2.1 (0.1)	2.6 (0.2)	2.2 (0.2)
H_2	4.1 (0.1)	**2.7**	3.8 (0.1)	6.3 (0.3)	4.7 (0.3)	5.7 (0.1)
CO_2	6.2 (0.2)	4.6 (0.1)	6.4 (0.6)	5.8 (0.2)	6.2 (0.2)	7.0 (0.1)
Acetate	3.2 (0.4)	2.9 (0.2)	**2.0**	3.9 (0.1)	2.8 (0.2)	4.6 (1.1)
Ethanol	4.4 (0.3)	4.8 (0.1)	5.5 (0.2)	**4.3**	4.7 (0.9)	5.1 (0.1)
Formate	2.2 (0.4)	3.1 (0.3)	1.7 (0.6)	1.9 (0.1)	**1.1**	0.5 (0.1)
Lactate	0.1 (0.1)	0.1 (0.0)	0.1 (0.1)	0.2 (0.1)	0.2 (0.1)	**-1.0**
Carbon recovery (%)	92.4	90.5	**90.0**	**90.2**	**90.6**	**90.2**
O/R index	1.1	**1.0**	1.0	**0.9**	**1.0**	1.1
C_1/C_2	1.0	1.0	**1.1**	**0.9**	**1.0**	0.93

* Final cellobiose was below limits of detection (0.1 mM)

3.4.2 Effect of end-product loading on final product concentrations

Final end-product concentrations of *C. thermocellum* grown under carbon limited conditions (1.1 g L^{-1} cellobiose) in the presence of added H_2, sodium acetate, ethanol, sodium formate, and sodium lactate are shown in Table 3.2. Under all conditions tested, final cellobiose concentrations were below detectable limits and thus growth ceased due to carbon limitation. Changes in the production of added end-product could not be measured accurately and thus were calculated based on an expected carbon recovery of ~90%, O/R index of 0.9-1.1, and C_1/C_2 ratio of 0.9-1.1 (shown in bold, Table 3.2). In the absence of end-products H_2 and ethanol yields were 0.68 and 0.71 mol mol-glc^{-1}, respectively. Replacement of the N_2 headspace with H_2 (100%) resulted in a significant decrease in CO_2 and calculated H_2 production accompanied by an increase in ethanol and formate production, while changes in other end-products were not statistically significant. In the presence of sodium acetate, production of ethanol increased by 28% and was accompanied by a decrease in calculated acetate production and marginal decrease in H_2 production. In contrast, production of H_2 and acetate increased by 54% and 22% respectively in the presence of exogenous ethanol, whereas production of other end-products remained unchanged. Statistically, final end-product concentrations were unaffected by sodium formate with the exception of H_2, which increased marginally, despite a 2-fold decrease in calculated formate production. Production of acetate, H_2, CO_2, and ethanol was favoured over formate in the presence of sodium lactate, which was surprising due to the fact that, despite the presence of lactate dehydrogenase (LDH) activity in cell extracts, lactate was not produced under carbon limited conditions and

thus addition of sodium lactate was not expected to influence product formation. However, calculations of lactate based on carbon recovery suggest that lactate may have been consumed. Addition of 40 mM or 80 mM NaCl (equivalent to the amount of sodium added with the addition of the organic acid salts) had no effect on final end-product concentrations (data not shown). Final H_2 yields ranged from 0.62 to 1.04 mol mol-glc^{-1} in the presence of sodium acetate and ethanol, respectively, while ethanol yields ranged from 0.71 to 0.91 mol mol-glc^{-1} in the absence of any end-products and the presence of sodium acetate, respectively.

3.4.3 Effect of organic end-product loading on enzyme activities

Activities (in nmol min^{-1} mg-protein^{-1}) of key enzymes involved in pyruvate catabolism and H_2, ethanol, and lactate production were measured in cell extracts obtained from late-exponential phase cells (OD_{600} ~ 0.45) grown in the presence of sodium acetate, ethanol, sodium lactate, sodium formate, or a H_2 headspace (Table 3.3). Since our aim was to determine the changes in enzyme levels in response to the presence of a given end-product, all activities were tested at 22°C and thus do not represent actual *in vivo* activities at 60°C. The activities of different enzymes were first compared within each growth condition, revealing that PFOR activities were on average 18 and 11-fold higher than that of pyruvate:formate lyase (PFL) and LDH activities, respectively. While the difference between ferredoxin-dependent (Fd-H_2ase) and NADH-dependent hydrogenase (NADH-H_2ase) activity was less than 3-fold, the activity of NADPH-dependent hydrogenase (NADPH-H_2ase) was more than 4-fold higher than that of both

Fd-H_2ase and NADH-H_2ase under most growth condition. NADH and NADPH-dependent alcohol dehydrogenase (ADH) activities were comparable.

Table 3.3: Effect of exogenous end-products on late-exponential phase enzyme activities involved in catabolism of pyruvate into end-products of *C. thermocellum* grown under carbon limited conditions (1.1 g L^- cellobiose). Samples represent a minimum of two biological replicates. Standard deviations are given in parentheses.

Enzyme	Substrate	Specific Activity (nmol min^{-1} mg $protein^{-1}$)[1] in the Presence of Given End-Product					
		None	Hydrogen	Acetate	Ethanol	Lactate	Formate
Pyruvate:formate lyase	Pyruvate: CoA	12 (1)	17 (1)	14 (3)	24 (2)	15 (2)	18 (4)
Pyruvate:fd oxidoreductase	Pyruvate: CoA :MV	330 (42)	225 (109)	345 (128)	282 (42)	303 (183)	141 (18)
Hydrogenase	H_2: MV	1501 (119)	2249 (601)	1466 (90)	1544 (84)	1339 (367)	1476 (58)
	H_2: NAD^+	9 (4)	6 (0)	4 (1)	7 (6)	6 (1)	8 (1)
	H_2: $NADP^+$	58 (14)	77 (16)	50 (10)	65 (12)	72 (20)	54 (8)
	H_2:Fd:Metronidazole	9 (2)	10 (2)	9 (3)	18 (5)	14 (4)	9 (1)
Lactate dehydrogenase	Pyruvate: NADH: FDP	26 (2)	26 (2)	24 (2)	20(4)	30 (2)	28 (3)
Alcohol dehydrogenase	Acetaldehyde: NADH	13 (1)	14 (2)	14 (2)	12 (4)	16 (3)	12 (2)
	Acetaldehyde: NADPH	11 (3)	10 (2)	12 (3)	12 (3)	15 (4)	17 (6)

[1]Detection limits are 1.3, 0.4, and 0.3 nmol min^{-1} mg $protein^{-1}$ for assays spectrophotometrically monitoring redox changes in methyl viologen (MV), NAD(P)H, and metronidazole respectively

In most cases, cell-extract enzyme activity data revealed no changes in enzyme levels greater than 2-fold in response to the presence of any given end-product with the exception of PFL, which increased in the presence of ethanol 2-fold, PFOR, which decreased in the presence of sodium formate by 57%, and Fd-H_2ase, which increased 2-fold in the presence of ethanol with respect to enzyme levels found in the absence of added end-products. Only a 50% and 42% increase in PFL activity in the presence of formate and H_2, respectively, was observed. Although methyl viologen-dependent hydrogenase (MV-H_2ase) activity increased in the presence of H_2 by 50%, standard deviation was high and thus the change may not be statistically relevant. Changes in LDH and NADH and NADPH-dependent ADH activities were of no statistical significance.

3.5 Discussion

3.5.1 Growth inhibition

Islam *et al*. report that growth of *C. thermocellum* ATCC 27405 in batch cultures ceases prior to carbon exhaustion when grown at 4.5 g L^{-1} cellobiose, while growth of cultures on 1.1 g L^{-1} cellobiose is carbon limited (Islam et al. 2006). End-products produced during microbial fermentations have been shown to be potential growth inhibitors of some fermentative microorganisms (Ben-Bassat et al. 1981; Kell et al. 1981; Herrero et al. 1985a). Specifically, carboxylic acids can act as proton motive force uncouplers (Kell et al. 1981; Herrero et al. 1985b), alcohols can denature proteins and cause deleterious alteration in membrane structures (Herrero et al. 1985a), and H_2 can limit the recycling of reduced cofactors (Ben-Bassat et al. 1981). To determine if end-

product accumulation inhibited growth we grew batch cultures under carbon limited conditions (1.1 g L^{-1} cellobiose) in the presence of end-products at concentrations roughly 10-fold and 4-fold higher than those found under carbon limited (1.1 g L^{-1} cellobiose) and carbon excess (4.5 g L^{-1} cellobiose) conditions, respectively. Although generation time increased by 90% and 132% in the presence of all organic end-products under a N_2 and H_2:CO_2 atm, respectively, significant growth was observed, albeit to somewhat lower optical densities (OD_{600} ~ 0.73 versus 0.57 and 0.51, respectively), suggesting that growth under carbon excess conditions is not directly inhibited due to accumulation of end-products, but rather due to nutrient depletion or a drop in pH brought about by acid formation. Given that *C. thermocellum* ATCC 2405 does not grow below pH 6.0 (Kurose et al. 1989), we propose that the latter explanation is more valid as final culture pH was roughly 6.0 as growth ceased under carbon excess conditions (Islam et al. 2006).

The presence of exogenous end-products at the concentrations used in this experiment had more subtle effects on growth when added to culture medium individually, consistent with the fact that concentrations of added sodium acetate, ethanol, and H_2 used (55 mM, 36 mM, and 1 atm, respectively) were below that of previously reported concentrations known to induce a 50% decrease in growth rate in *C. thermocellum* ATCC 27405 (280 mM and 110 mM for acetate and ethanol, respectively), (Herrero et al. 1985a; Herrero et al. 1985b). While He *et al*. (He et al. 2009) report a concentration-dependent increase in growth rate and final cell yield of *Thermoanaerobacter ethanolicus* and *C. thermocellum* ATCC 27405 in the presence of acetate in minimal medium, this increase is less pronounced and independent of the

concentration of acetate added in medium containing 0.1% (w/v) yeast extract (Table 3.4). We, however, did not see any changes in growth in response to addition of sodium acetate (and sodium formate) possibly because our media contained 0.2% (w/v) yeast extract. In contrast, the presence of added ethanol and H_2 did decrease growth rate by 64% and 68%, respectively. The 59% reduction in growth rate in the presence of 55 mM sodium lactate was comparable to the 50% reduction of growth rate of *C. thermocellum* ATCC 27405 in the presence of 30 mM lactate (Herrero et al. 1985a) and coincided with the decrease in growth rate of *T. ethanolicus* (He et al. 2009) (Table 3.4). Interestingly, in our study, each of the end-products that affected growth rate are electron sinks, whose production involves the re-oxidation of NADH. Given that NAD^+ is required for glycolysis and that glyceraldehyde-3-phosphate dehydrogenase has been shown to be inhibited by increased ratios of $NADH/NAD^+$ (Lovitt et al. 1988), it is plausible that reducing carbon and electron flux through pathways that re-oxidize NADH to NAD^+ (i.e. via exogenous addition of products that function as electron sinks) may increase $NADH/NAD^+$ ratios, leading to decreased carbon flux through glycolysis, and thus slower growth rates.

Table 3.4: Comparative effects of exogenous end-product addition on growth and end-product generation in various saccharolytic thermophilic organism grown on different substrates and in different media.

Organism	Strain	Substrate	Media[1]	Growth Condition	Highlights of Exogenous Product Loading Effects	References
Clostridium thermocellum	LQ8	Cellulose	CM3 (C)	Batch	• Co-culturing with *Methanobacterium thermoautotrophicum* increased final H_2 (200%) and acetate (220%) concentrations, and decreased final ethanol (80%) concentrations	Weimer & Zeikus 1977
Thermoanaerobium brockii	HTD4	Glucose	LPBB+YE (C)	Batch	• Added H_2 (1.0 atm) ceased growth • Added H_2 (0.5 atm) prevented H_2 production, and decreased final acetate (32%) and ethanol (10%) concentrations	Ben-Bassat et al. 1981
Clostridium thermocellum	JW20	Cellulose Cellulose Cellobiose Cellobiose	Unnamed (C)	Batch Batch[2]	• 40 mM lactate and 250 mM acetate arrested growth • Agitation (100 strokes/min) increased final H_2 (84%), CO_2 (34%) and acetate (58%), and decreased ethanol (48%) production on cellulose • Agitation (75 strokes/min) increased final H_2 (47%), CO_2 (4%), acetate (44%), and lactate (49%), and decreased ethanol (19%) production on cellobiose • Added H_2 (100%; 45 mM)) increased final ethanol concentration (~215%)	Freier et al. 1988
Clostridium thermocellum	YS; AS39; LQRI	Cellulose Cellobiose	CM3- YE (D)	Batch	• Stirring (150 rpm) increased H_2 (180%; 83%; 84%) and acetate (167%; 78%; 81%) production, and decreased ethanol (14%; 33%; 29%) production for strains YS, AS39, and LQRI, respectively, on cellulose • Stirring (150 rpm) increased H_2 (56%; 63%; 125%) and acetate (58%; 61%; 113%) production, and decreased ethanol (34%; 23%; 23%) production for strains YS, AS39, and LQRI, respectively, on cellobiose	Lamed et al. 1988

		Cellulose			• Added H_2 (2.5 atm) increased ethanol/acetate product ratios by 23%,, 45%, and 9% for strains YS, AS39, and LQRI, respectively, grown on cellulose	
		Cellobiose			• Added H_2 (2.5 atm) increased ethanol/acetate product ratios by 57%,, 30%, and 9% for strains YS, AS39, and LQRI, respectively, grown on cellobiose	
Clostridium thermocellum ***Thermoanaerobium brockii***	AS39; LQRI HTD4	Cellobiose	CC (C)	Batch	• Added H_2 (1 atm) increased ethanol/acetate ratio by 67% and 89% for strains AS39 and HTD4, respectively, but had no impact on strain LQRI	Lamed & Zeikus 1980
Clostridium thermocellum	JW20	Cellobiose	Unnamed (C)	Continuous[2]	• Elevated hydrostatic pressures (0.1, 7.0, and 17.3 MPa) increased ethanol/acetate ratios (0.5, 1.5, and 2.4, respectively, at a dilution rate of 0.05 h^{-1}) • Growth was inhibited by 40% and 60% under hydrostatic pressures of 7.0 and 17.3 MPa compared to growth at 0.1 MPa)	Bothun et al. 2004
Thermoanaerobacter tengcongensis	MB4	Starch	Modified MB (C)	Batch	• Growth in open system (N_2-flushed bioreactor) vs closed system (sealed bottles) decreased ethanol/acetate ratio by 98%	Soboh et al. 2004
Thermoanaerobacter ethanolicus	39E	Glucose	TYE (C)	Batch	• Added H_2 (100%) increased ethanol (13%), lactate (78%), and H_2 (22%) yields, and decreased acetate (30%) yields	Lovitt et al. 1988
		Pyruvate			• Added ethanol (5% [wt/vol]) decreased final growth yield by (27%) and specific growth rate by 62%	
Mixed anaerobic culture		Glucose	Unnamed (D)	Batch	• Added butyrate (47 mM) decreased H_2 (8%) production, and increased acetate (21%) and ethanol (27%) production	Zheng & Yu 2005

Thermoanaerobacter ethanolicus	39E	Xylose; Glucose; Cellobiose	Unnamed (D/C)	Batch	• Enhancement of growth and substrate utilization by added acetate (25, 75, and 150 mM) was concentration-dependent in the absence of yeast extract and concentration-independent in the presence of yeast extract on all substrates • In the absence of yeast extract, added (25, 75, 150 mM) increased ethanol production (~60%, 110%, 210%, respectively, on xylose; ~100%, 190%, 175%, respectively, on glucose; ~160%, 230%, 300%, respectively on cellobiose) • In the presence of yeast extract, added acetate (25, 75, 150 mM) increased ethanol production by ~10% on xylose or glucose regardless of concentration of acetate added • Added lactate (150 mM) decreased ethanol production by ~20% in the absence of yeast extract, and increased ethanol production by ~35% in the presence of yeast extract	He et al. 2009
Clostridium thermocellum	ATCC 27405	Cellobiose			• Enhancement of growth and substrate utilization by added acetate (25, 75, and 150 mM) was concentration-dependent in the absence of yeast extract • In the absence of yeast extract, added acetate (25, 75, 150 mM) increased ethanol production by ~200%, 330%, and 320%, respectively, on cellobiose	
Clostridium thermocellum	ATCC 27405	Cellobiose	CM4 (C)	Batch	• 280 mM acetate and 30 mM lactate inhibit growth by 50%	Herrero et al. 1985b

[1]Defined as complex media (C) containing yeast extract (YE) and defined media (D) without yeast extract

[2]Controlled pH

3.1.1 End-product shifts and enzyme activities

In agreement with previous studies (Lamed and Zeikus 1980a; Lynd et al. 1989; Strobel 1995; Stevenson and Weimer 2005; Islam et al. 2006; Rydzak et al. 2009) ethanol, H_2, CO_2, acetate, and formate (Islam et al. 2006; Rydzak et al. 2009) were major end-products of *C. thermocellum* ATCC 27405, while lactate was produced in trace amounts under carbon limited conditions. The use of mildly-inhibitory concentrations of exogenous end-products, during which significant growth was sustained, allowed us to determine the effects of these end-products on pyruvate catabolism and end-product formation. Stationary phase end-product data of carbon-limited cultures reveals shifts in carbon and electron flux in response to individually added end-products (Table 3.2), despite marginal changes in intracellular levels of key enzymes representing the various branches of the fermentation pathways (Table 3.3). Addition of a given exogenous end-product may impede production of that end-product through altering thermodynamic efficiencies, and thus redirecting carbon and electrons through alternative metabolic pathways.

3.5.2 Response to exogenous hydrogen

Numerous studies have shown changes in end-product yields in response to changes in H_2 partial pressures (Table 3.4). Increasing gas phase and dissolved H_2 partial pressures by either growing cultures in a closed versus an open system (Lamed et al. 1988; Soboh et al. 2004), in the presence of a H_2 headspace (Lamed and Zeikus 1980a; Ben-Bassat et al. 1981; Freier et al. 1988; Lamed et al. 1988), or under elevated

hydrostatic pressures (Bothun et al. 2004) has been shown to increase ethanol/acetate ratios, and decrease H_2 yields. Unfortunately, these studies did not report CO_2. Furthermore, while many of these studies (Lamed and Zeikus 1980a; Lamed et al. 1988; Bothun et al. 2004) were performed on *C. thermocellum* strains (YS, LQRI, JW20, or AS-39; a mutant of strain ATCC 27405), formate production has only been reported in strain ATCC 27405 (Islam et al. 2006; Sparling et al. 2006; Rydzak et al. 2009). In agreement with these studies, we found that ethanol/acetate ratios increased from 1.3 to 1.7 when the 100% N_2 headspace was replaced with 100% H_2. Furthermore, in the presence of 100% H_2 we observed that formate production increased and CO_2 production decreased, suggesting that carbon and electron flux was shifted from PFOR to PFL. Although H_2 production could not be measured accurately, our calculations show a 34% decrease in H_2 production further supporting a decrease in flux through PFOR, resulting in lower concentrations of reduced ferredoxin for H_2 production.

Given that the H_2-mediated shift from H_2 + acetate to ethanol production was accompanied by a fourfold decrease in NADH-H_2ase activity and an increase in aldehyde dehydrogenase and ADH activity in *T. tengcongensis* (Soboh et al. 2004), we set out to determine if changes in enzyme expression are responsible for the product shifts observed. Although PFOR activities did not change significantly in the absence or presence of H_2, PFL activity increased 42% in the presence of H_2, complementing our end-product data which suggests that carbon and electron flux is directed away from PFOR towards PFL. However, none of the H_2ase or ADH activities detected fluctuated

in the presence of H_2 with the exception of MV-H_2ase activity, which seemed to increase in the presence of H_2. Unfortunately, standard deviations were high between biological samples and thus the statistical relevance of this change cannot be ensured.

The increase in ethanol yields in the presence of H_2 may be attributed to increases in NADH/NAD^+ ratios. Tritium addition to the gas phase has demonstrated that the H_2ases in *C. thermocellum* YS can catalyze the uptake of H_2 (Lamed et al. 1988) via NAD(P)H-dependent H_2ases and/or NAD(P)H:ferredoxin oxidoreductase, while pulsing of steady-state cultures of *T. ethanolicus* 39E (previously known as *C. thermohydrosulfuricum*) with H_2 has been shown to increase NADH/NAD^+ ratios . Pei *et al*. have recently isolated a redox sensing protein in *T. ethanolicus* which senses NADH, leading to conformational changes in the protein alleviating repression of ADH transcription (Pei et al. 2011). Differential transcription of ADHs with different kinetic properties has also been reported (Pei et al. 2010). We thus suggest that exogenous H_2 may lead to generation of NAD(P)H, which is in turn is reoxidized via ADHs resulting in higher ethanol yields. While changes in ADH activities in response to H_2 were insignificant, kinetic characterization and determination of expression profiles of the various ADHs encoded by *C. thermocellum*, in conjunction with measurements of NAD(P)H/$NAD(P)^+$ pools, will provide a more comprehensive response mechanism to H_2.

3.5.3 Response to exogenous acids and ethanol

The addition of sodium acetate had negligible effects on CO_2 and formate production, marginally decreased H_2 production, and increased ethanol production by 28%. This was in agreement with He *et al*. (He et al. 2009) who report increased ethanol production by *T. ethanolicus* and *C. thermocellum* ATCC 27405 in response to exogenous acetate (Table 3.4). End-product profiles suggest that exogenous acetate does not affect carbon and electron flux at the *Pyruvate/Lactate/Acetyl-CoA* branchpoint, but does redirect acetyl-CoA towards ethanol, during which NADH is reoxidized, leaving fewer reducing equivalents for H_2 production. Despite the changes in product yields, no changes in enzyme activity were observed in response to exogenous sodium acetate additions, indicating that carbon and electron fluxes are not necessarily regulated by the amount of enzyme expressed. This is not surprising given that carbon fluxes in *C. cellulolyticum* have been shown to be, in most cases, much lower than the available enzyme activity in the cell (Desvaux et al. 2001c). Interestingly, addition of sodium formate had no significant effect on final end-product ratios despite the observed 57% decrease in PFOR activity and the 42% increase in PFL activity. This further supports that available enzyme activities are higher than the carbon and electron flux flowing through them, and thus changes in expression of these enzymes may not necessarily impose changes in final end-product yields under the growth conditions tested.

Addition of exogenous ethanol had little impact on CO_2 and formate production despite the 2-fold increase in PFL activity. Exogenous ethanol, however, did redirect

acetyl-CoA at the *Acetyl-CoA/Ethanol/Acetate* branchpoint towards acetate. Furthermore, addition of ethanol increased H_2 production by 54% which was accompanied by a 2-fold increased in Fd-H_2ase activity. The increase in H_2 and acetate in the presence of ethanol mirrors end-product shifts of *T. ethanolicus* when grown on pyruvate with increasing concentrations of ethanol (Lovitt et al. 1988). As with H_2, pulsing of steady-state cultures of *T. ethanolicus* with ethanol increased NADH/NAD^+ ratios (Lovitt et al. 1988) thus enhancing the thermodynamic feasibility of H_2 production from NADH. The increase in acetate/ethanol ratios and increase in H_2 yields in the presence of exogenous ethanol (Table 3.4) were analogous to the effects observed previously of decreasing dissolved H_2 partial pressures via stirring (Lamed et al. 1988), agitation (Freier et al. 1988), or by co-culturing with methanogens (Weimer and Zeikus 1977; Ben-Bassat et al. 1981).

The presence of exogenous sodium lactate resulted in a significant increase in acetate, H_2, CO_2, and ethanol production while the production of formate decreased by 77%. This was not observed with *T. ethanolicus* (He et al. 2009) possibly due to differences in lactate production between *T. ethanolicus* and *C. thermocellum* ATCC 27405 under the conditions tested. Given that only trace amounts of lactate are produced under carbon-limited conditions by *C. thermocellum*, the change in final end-product concentrations could not be explained by inhibition of lactate production, but rather by lactate consumption. In fact our calculations indicate that 1 mM lactate was consumed. Despite the strong regulatory effect observed on formate production in response to added

sodium lactate, no significant changes in the activity of PFL or other enzymes involved in pyruvate catabolism or end-product formation were observed.

Here we demonstrated that manipulation of fermentation parameters (e.g. addition of exogenous end-products) can redirect carbon and electron flow resulting in changes in final product yields. While changes in end-product concentrations in response to exogenous end-products are dependent on strain and growth conditions (Table 3.4), the trends are seemingly the same. Although changes in enzyme activities in response to the presence of exogenous end-product additions were observed, the lack of correlation of these changes in enzyme activities with changes in final end-product yields suggests that the changes in final end-product yields may be governed by thermodynamic considerations that redirect carbon and electron flow rather than levels of enzyme expressed under the conditions tested. While the economic feasibility of industrial fermentations relies preferably on the use of raw lignocellulosic materials, we used cellobiose as our model carbon substrate, and thus bioenergetic parameters, flux distributions, and resulting end-product ratios may not represent cellulose fermentations (Strobel 1995; Strobel et al. 1995; Zhang and Lynd 2005a), since cellotriose and cellotetraose rather than cellobiose form a significant proportion of the soluble sugars taken up during cellulose fermentation. Nevertheless, these results have practical implications in commercial anaerobic bacterial fermentations as build-up of a desired end-product, for example ethanol or H_2, may negatively impact on its further production. Modification of fermenter design permitting selective end-product recycling or removal

may be used to manipulate bacterial metabolism in order to selectively improve desired product yields.

3.6 Authors' contributions

Thomas Rydzak performed growth experiments, end-product analysis, enzyme activity analysis, and authored the manuscript. Richard Sparling, Nazim Cicek, and David B. Levin conceived of the study, participated in its design, and helped draft the manuscript.

3.7 Acknowledgements

This work was supported by funds provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944-04) and the BIOCAP Canada Foundation.

Chapter 4: Role of Transcription and Enzyme Activities in Redistribution of Carbon and Electron Flux in Response to N_2 and H_2 Sparging of Open-Batch Cultures of *Clostridium thermocellum* ATCC 27405[12]

4.1 Abstract

Growth, end-product synthesis, enzyme activities, and transcription of select genes associated with the "malate shunt", pyruvate catabolism, H_2 synthesis, and ethanol production were studied in the cellulolytic anaerobe, *Clostridium thermocellum* ATCC 27405, during open-batch fermentation of cellobiose to determine the effect of elevated N_2 and H_2 gas sparging on metabolism using a 14 liter (L) fermenter with a 7 L working volume. The metabolic shift from acetate, H_2, and CO_2 to ethanol and formate in response to high H_2 versus high N_2 sparging (20 mL s^{-1}) was accompanied by (i) a 2-fold increase in NADH-dependent alcohol dehydrogenase (ADH) activity, (ii) a 10-fold increase in *adhE* transcription, and (iii) a 3-fold decrease *adhZ* transcription. A similar, but less pronounced, metabolic shift was also observed when the rate of N_2 sparging was

---

[12] Contributing authors: Thomas Rydzak[†], Carlo Carere†, Nazim Cicek, David B. Levin and Richard Sparling. *Applied microbiology and biotechnology* (Submitted). ([†]Equal contributors)

decreased from 20 mL s^{-1} to 2 mL s^{-1}, during which (i) NADH-dependent ADH and pyruvate:ferredoxin oxidoreductase (PFOR) activities increased by ~1.5-fold, (ii) *adhY* transcription increased 6-fold, and (iii) transcription of selected *pfor* genes increased 2-fold. Here we demonstrate that transcription of genes involved in ethanol metabolism is tightly regulated in response to gas sparging. We discuss the potential impacts of dissolved H_2 on electron carrier (NADH, NADPH, ferredoxin) oxidation, and how these electron carriers can redirect carbon and electron flux and regulate *adhE* transcription.

4.2 Introduction

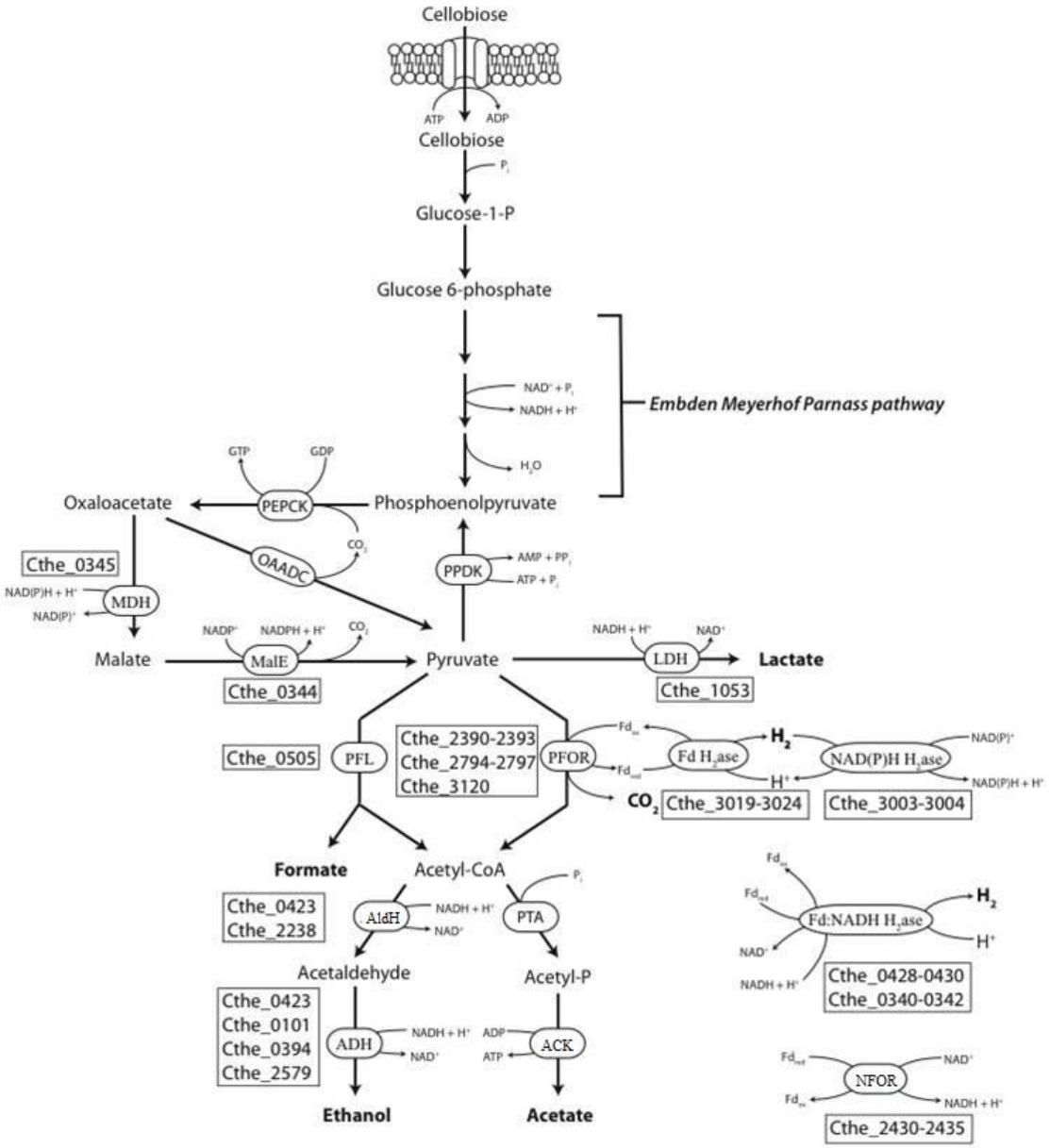
Successful development of lignocellulosic biofuels using (hemi)cellulolytic bacteria such as *Clostridium thermocellum,* depends heavily on a detailed understanding of fermentative metabolism. As a consequence of highly branched catabolic pathways (Rydzak et al. 2009; Rydzak et al. 2011) (Fig. 4.1), yields of fuel chemicals fall far below theoretically maximal values (2 mol-ethanol mol-glucose^{-1} or 4 mol-H_2 mol-glucose^{-1}). Growth conditions can significantly influence end-product yields through regulation of gene and gene product expression (Soboh et al. 2004; Stevenson and Weimer 2005), and modulation of metabolic flux and intracellular metabolite levels (Strobel 1995; Guedon et al. 1999a). Variations in substrate availability/dilution rate (Strobel 1995; Guedon et al. 1999b; Guedon et al. 2000b; Desvaux et al. 2001c; Desvaux et al. 2001b; Desvaux and Petitdemange 2001; Stevenson and Weimer 2005; Islam et al. 2006), substrate composition (Ben-Bassat et al. 1981; Girbal and Soucaille 1994; Vasconcelos et al. 1994; Strobel 1995; Strobel et al. 1995; Zhang and Lynd 2005b; Levin et al. 2006), media

composition (Guedon et al. 1999b), pH (Desvaux et al. 2001c), gas partial pressure (Lamed et al. 1988; Bothun et al. 2004; Soboh et al. 2004; Nguyen et al. 2010), growth phase (Willquist et al. 2010), osmolarity (Ljunggren et al. 2011), and end-product accumulation (Lamed and Zeikus 1980a; Ben-Bassat et al. 1981; Freier et al. 1988; Lamed et al. 1988; Rydzak et al. 2011) are all known to dramatically influence end-product yields.

Figure 4.1: Genetically deduced metabolic pathway of pyruvate catabolism and end-product synthesis in *C. thermocellum* ATCC 27405. Enzyme designations and corresponding locus tags for genes analysed by RT-qPCR analysis are provided. PEPCK, phosphoenolpyruvate carboxykinase; OAADC, oxaloacetate decarboxylase; PPDK, phosphoenolpyruvate dikinase; MDH, malate dehydrogenase; MalE, malic enzyme; LDH, lactate dehydrogenase; PFL, pyruvate:formate lyase; PFOR, pyruvate:ferredoxin oxidoreductase; AldH, acetaldehyde dehydrogenase; ADH, alcohol dehydrogenase; Fd H_2ase, ferredoxin-dependent hydrogenase; NAD(P)H H_2ase, NAD(P)H-dependent hydrogenase; Fd:NADH H_2ase, "*ferredoxin and NADH-dependent bifurcating hydrogenase*[13]"; NFOR, NADH:ferredoxin oxidoreductase; PTA, phosphotransacetylase; ACK, acetate kinase.

---

[13] The trimeric *'bifurcating H_2ase'* was first introduced in section 1.5.5.2 of the literature review. It was first purified in *Thermotoga maritima* in 2009 by Schut and Adams (2009) and was shown to simultaneously oxidize reduced Fd and NADH to produce H_2. During initial characterization of *C. thermocellum* (Chapters 2 and 3), the presence of this hydrogenase was not considered. However, it may have major implications in electron balance and reduced end-product synthesis (specifically H_2 and ethanol; alluded to in section 1.7.7, and more thoroughly discussed in chapters 5, 6, and 7). Subsequent sections discuss the thermodynamic implications of their presence (section 5.4.2.4), and the impact they may have on reduced end-product yields due to the co-utilization of both NADH and reduced Fd (subsequent chapters, emphasized in chapter 7).


Cellobiose
ATP
ADP
Cellobiose
Glucose-1-P
Glucose 6-phosphate
Embden Meyerhof Parnass pathway
GTP
GDP
PEPCK
Oxaloacetate
Phosphoenolpyruvate
OAADC
PPDK
Cthe_0345
MDH
Malate
MalE
Cthe_0344
Pyruvate
LDH
Lactate
Cthe_1053
Cthe_0505
PFL
Cthe_2390-2393
Cthe_2794-2797
Cthe_3120
PFOR
Fd H2ase
NAD(P)H H2ase
Cthe_3019-3024
Cthe_3003-3004
Formate
Acetyl-CoA
Cthe_0423
Cthe_2238
AldH
PTA
Acetaldehyde
Acetyl-P
Cthe_0423
Cthe_0101
Cthe_0394
Cthe_2579
ADH
ACK
Ethanol
Acetate
Fd:NADH H2ase
Cthe_0428-0430
Cthe_0340-0342
NFOR
Cthe_2430-2435

Elevated dissolved H_2 concentrations (H_{2aq}) may inhibit H_2 evolving hydrogenases (H_2ases) within fermentative microorganisms resulting in increased ratios of reduced electron carriers relative to their oxidized counterparts (nicotinamide adenine dinucleotide, NADH; ferredoxin, Fd) and metabolic shifts towards other reduced end-products such as lactate and ethanol (Hallenbeck 2002). Gas stripping with N_2 has been shown to increase H_2 production in *Caldicellulosiruptor saccharolyticus* cultures (Ljunggren et al. 2011) and sparging with H_2 has been shown to increase ethanol yields in cultures of both *Clostridium thermocellum* (Li et al. 2012a) and *Caldanaerobacter subterraneus* subsp. *tengcongensis* (Soboh et al. 2004). While Li *et al*. examined the effect of N_2 and H_2 sparging on the production of ethanol and volatile fatty acids in *C. thermocellum*, the authors did not quantify formate production or look at changes in the expression of genes and/or proteins associated with pyruvate catabolism and product formation (Li et al. 2012a). Therefore, in this study we sought to determine if changes in end-product distribution in response to high N_2, low N_2, and high H_2 sparging rates are reflected by changes in enzyme activities and gene transcription.

4.3 Materials and methods

4.3.1 Microorganism and media

Clostridium thermocellum strain ATCC 27405 was used for all growth experiments. Culture maintenance, media composition and preparation, and growth

experiments in closed-batch cultures were performed as previously described[14] by Rydzak *et al*. (Rydzak et al. 2009); however cells were grown and maintained on 2 g L^{-1} cellobiose (versus 1.1 g L^{-1} cellobiose), which still reflected carbon limited conditions.

4.3.2 Experimental design

Batch culture growth was carried out using a 14 L New Brunswick Scientific Bioflow 110 fermenter with a working volume of 7.0 L. Reactor temperature was maintained at 60 °C and pH was monitored in real time. Fermentation experiments were conducted under low gas sparging rates with N_2 (2 mL s^{-1}) and under conditions of high gas sparging rates (20 mL s^{-1}) of either N_2 or H_2. Fermentation broth was continuously mixed (250 rpm) by an axial flow impellor under direction of the fermenter control module to maintain a homogenous suspension. Gas effluent passed through a condenser and condensate was collected and analyzed for end-products. Samples for biomass and end-product analysis (1.5 ml) were taken throughout growth, whereas samples for determination of gene expression (50 ml) and enzyme activities (2 L) were harvested anaerobically during exponential phase (OD ~ 0.4).

4.3.3 Cell growth and end-product analysis

Cell growth measurements, sample processing, pH measurement, product gas, protein, sugar, and end-product analyses, and biomass and gas calculations were

---

[14] Refer to section 2.3.1 for details

performed as previously described[15] (Rydzak et al. 2009). Data represent the average of three technical replicates for each run. A minimum of 2 biological runs were performed under each condition. Final H_2 and CO_2 concentrations were calculated using biomass and fermentation products assuming a carbon recovery of ~ 90 % and a C1/C2 ratio and O/R index of 0.9-1.1.

4.3.4 Preparation of cell extracts and enzyme assays

Two liters of mid-exponential (OD_{600} ~ 0.4) *C. thermocellum* cells were anaerobically transferred from the fermenter to an anaerobic 4 liter reagent bottle (Bellco Glass Inc., Vineland, NJ) via a rubber tube using a peristaltic pump. Prior to cell harvesting, the reagent bottle was sealed with a butyl rubber stopper equipped with a glass tube attached to a clamped rubber tube, and was sparged with N_2 for 20 minutes. During cell transfer, the rubber tube was connected to the fermenter efflux tube, de-clamped, and cells were transferred into the reagent bottle via a peristaltic pump. All subsequent cell manipulations were performed under anaerobic conditions in a Thermo Forma Anaerobic Chamber (Model 1025, Thermo Fisher Scientific, Waltham, MA, USA) with an atmosphere of 5 % hydrogen, 95 % nitrogen. Cell extracts were prepared as previously described[16] by Rydzak *et al* (Rydzak et al. 2009). Cell extracts re-suspended and reactivated in reactivation buffer were used for pyruvate:formate lyase assays and all

[15] Refer to section 2.3.2 for details

[16] Refer to section 2.3.3 for details

hydrogenase assays as described[17] by Rydzak *et al*. (Rydzak et al. 2011). All other activities were measured using extracts of cells re-suspended in cell suspension buffer.

All assay conditions and specific activity calculations were performed as previously described[18] by Rydzak *et al*. (Rydzak et al. 2009) with the exception of malate dehydrogenase (MDH) and malic enzyme (MalE), which were not previously determined. MDH was measured in the direction of malate formation (NAD(P)H oxidation). The standard assay mixture contained 0.1 M Tris-Cl (pH 6.5), 10 mM DTT, 2 mM oxaloacetate, 0.3 mM NAD(P)H (Sridhar et al. 2000). Malic enzyme was measured in the direction of pyruvate formation. The standard assay mixture contained 0.1 M Tris-Cl (pH 6.5), 10 mM DTT, 2 mM malate, 0.5 mM $NAD(P)^+$, 5 mM $MnCl_2$, 20 mM NH_4Cl (Lamed and Zeikus 1981). Each assay was done in triplicate on two different cell extracts.

4.3.5 RNA extraction and reverse transcription quantitative PCR (RT-qPCR) analysis

Exponential phase cells (OD~0.4) were collected from 50 mL of sample by centrifugation (4500 x g, 20 min) at 4 °C. Cells were re-suspended in RNA*later* solution (Ambion, Carlsbad CA) as per the manufacturers recommended protocol for RNA extraction and RT-qPCR analysis. Total RNA was isolated using the Invitrogen TRIzol

[17] Refer to section 3.3.3 for details

[18] Refer to section 2.3.4 for details

Reagent kit (Invitrogen, Carlsbad CA) according to the manufacturer's guidelines. Cells were pelleted by centrifugation (4500 x g, 12 minutes, 4 °C) prior to RNA extraction. RNA pellets were dissolved in DEPC-treated RNase free H_2O containing RNase inhibitor (Invitrogen) and dithiothreitol (DTT) at concentrations of 0.5 U μL^{-1} and 1 mM, respectively. RNA samples were treated with DNase I (0.1 U μL^{-1}) for 15 min at 20 °C in reactions containing 2 mM $MgCl_2$, 20 mM Tris-Cl (pH 8.4) and 50 mM KCl, prior to cDNA synthesis. Reactions were halted by the addition of ethylenediaminetetraacetic acid (EDTA) to a concentration of 2.5 mM and incubation at 65 °C for 10 min.

cDNA synthesis was accomplished with SuperScript II Reverse Transcriptase (Invitrogen), following the manufacture's recommended protocol with random hexamer primers. Each reaction was performed under the following conditions: between 1-5 μg of total RNA, 2.5 μM random hexamer primers, 0.5 mM dNTPs, 5 mM DTT, 1.25 mM $MgCl_2$, 0.5x RT buffer, 2.5 U μL SuperScript II Reverse Transcriptase and RNase H. An Eppendorf Mastercycler EP thermocycler was used for cDNA synthesis. Samples containing RNA, dNTPs and random hexamer primers were incubated at 65 °C for 5 min prior to the addition of RT buffer, and DTT. Contents were next incubated for 2 min at 42 °C before the addition of SuperScript II reverse transcriptase. Following incubation at 42 °C for 50 min, reactions were terminated by heating samples at 70 °C for 15 min. Finally, RNA complementary to cDNA was removed by addition of RNase H and incubation at 37°C for 20 min.

All primers used in RT-qPCR reactions were designed using Primer3 software (http://primer3.sourceforge.net/) (Rozen and Skaletsky 2000) with amplicon sizes averaging 200 bp and T_m values ranging between 58 and 60 °C. Primer nucleotide sequences are outlined in Table 4.1. All amplicons were cloned into plasmid pGEM-t (Promega, Madison, WI) following the manufacturer's recommended protocol and insert identities were confirmed subsequently by DNA sequencing. Cloned vector DNA was extracted from *E. coli* DH5α cells using a QIAprep Spin miniprep kit (QIAGEN, Valencia, CA) and used to construct calibration curves for RT-qPCR analysis.

Table 4.1: Primers used for RT-qPCR analysis

Gene Locus	Annotation	Accession no.	Forward Primer (5'→ 3')	Reverse Primer (5'→3')
Cthe__R0041	16s rDNA	CP000568	TGACGGGCGGTGTGTACAAGG	GGTGGGGACGACGTCAAATCA
Cthe__0344	*malE*	YP_001036775	CGGTGAGTGGAAGGGTAAAA	CGGTGAGTGGAAGGGTAAAA
Cthe__0345	*mdh*	YP_001036776	GCCGGAGCCAACAGAAAACCT	CGTCAACGCCCAATTTTTCGC
Cthe__1053	*ldh*	YP_001037478	CAAAGACTGTGCCGGATCCGA	GGCTGTCCAAAACCGTTCC
Cthe__2796	*pfor*	YP_001039188	GGGAAATGAAGCAGTGGCGGA	AGCCCCTGGGATGAAGTTGCC
Cthe__2392	*pfor*	YP_001038787	ACCGATGATGCCGATGTTGCC	AACAAAGGTCCTCCGGCTGCA
Cthe__3120	*pfor*	YP_001039508	GCAGGGGCATTGACGACCACT	TAATGGCCGAAAGATGCGAA
Cthe__0505	*pfl*	YP_001036935	CCGAAGCTTATGGCCACAGTG	GCAATCAGCCTGTCAACGCCA
Cthe__2238	*aldH*	YP_001038633	GACATGCATAGCTCCGGATT	ACAAACTTTTGCCCCCTCTT
Cthe__0423	*adhE*	YP_001036854	TTTCTTTGAGGCGGCAATGGC	CCGCAATAACCCCCAAAGGCT
Cthe__2579	*adhZ*	YP_001038972	TGCTGGTTACGGGAAAGAAC	CTCCTCCGATTCCGATGATA
Cthe__0394	*adhY*	YP_001036825	GACATTGAGGCACGGTCAAA	CTCCTGCGTTGGAATTGGTAA
Cthe__0101	*adh*	YP_001036535	TGTGCAGTTTGCCGTAAGAG	GCATTTTGACGCCATTTCTT
Cthe__0430	[FeFe] *H_2ase*	YP_001036861	AGGAGCCGGTGTGATATTTG	CGCTGACCACTGCAACTTTA
Cthe__3003	[FeFe] *H_2ase*	YP_001039392	CCGGAATACTGTTTGGTGCT	TCTTAAGGCCGCTGACAACT
Cthe__0342	[FeFe] *H_2ase*	YP_001036773	AGGAGCCGTCTGAACGACTA	TGTGTCAGCATGTGCAAGAA
Cthe__3020	[NiFe] *H_2ase*	YP_001039409	GCCAGTGCATATCCCTCATT	GGTTGGAGTACGAACCCTGA

Cthe__2430	*nfor*	YP_001038825	GTATTGCCAAAGGAGCAAGC	TAAGCACAGCTGCCACATTC

Ten (10) ng of cDNA (~2 μL) was used as a template for subsequent RT-qPCR experiments. Reactions were run on a Stratagene Mx3005p thermocycler using 12 μL iTaq SYBR Green Supermix with Rox (2X reaction buffer, 0.4mM dNTPs, 50 U mL^{-1} iTaq polymerase, 6 mM $MgCl_2$, SYBR Green I dye, 1 μM ROX; BioRad, Mississauga, ON), 2.5 μL each of forward and reverse primer (final concentration 1 mM), and ddH_2O to a total volume of 25 μL. All RT-qPCR reactions were run in 8-strip tubes with flat caps (Axygen, Union city, CA). RT-qPCR reactions consisted of an initial 2 min denaturation step at 95 °C followed by 40 cycles of: 95 °C for 15 s, 55 °C for 45 s and 72 °C for 45 s. After cDNA amplification a thermal disassociation curve was generated to ensure amplification specificity. Calibration curves for all assayed genes displayed a linear dynamic range between 10^2-10^8 copies $reaction^{-1}$. Average determined r^2 values of 16s rRNA calibration curves were 0.98 (SD +/- 0.01) with an average PCR efficiency of 98 % (SD +/- 0.04). Assayed genes of interest (GOI) displayed an average r^2-value of 0.99 (SD +/- 0.01) for calibration curves with a calculated PCR efficiency of 1.03 (SD +/- 0.06). The average of intra-assay variation (repeatability) was calculated to be 1.52 % (SD +/- 2.01 %, n=513). "No-reverse transcription" controls, containing 10 ng RNA, for each assayed gene were performed to validate extracted RNA as DNA free.

The genes examined by RT-qPCR in this study were divided into four classes. The first class included genes involved in the '*malate shunt*' (*mdh, malE*). The second class included genes directly involved in pyruvate catabolism, including *ldh*, *pfl*, and three *pfor* genes. For *pfor* analysis, primers were designed to amplify regions of the α–

subunit regardless of whether they were identified as multi-subunit PFORs (Cthe_2392, Cthe_2796) or a single subunit PFOR (Cthe_3120). The third class included genes potentially involved in ethanol metabolism, including *aldH* (Cthe_2238), three type IV *adh* genes (Cthe_0101, Cthe_0394, Cthe_2579), and a bifunctional acetaldehyde/alcohol dehydrogenase (*adhE*; Cthe_0423). The designations *adhY* and *adhZ*, put forth by Stevenson *et al*., 2005, correspond to Cthe_0394 and Cthe_2579, respectively, and are maintained in this study to facilitate comparison (Stevenson and Weimer 2005). The fourth class of genes analyzed are directly or indirectly involved in H_2 metabolism and include two putative bifurcating [FeFe] H_2ases (Cthe_0342, Cthe_0430), a putative NAD(P)H-dependent [FeFe] H_2ases (Cthe_3003), a [NiFe] H_2ase (Cthe_3020), and a NADH:ferredoxin oxidoreductase (*nfor*; Cthe_2430). H_2ase genes listed above correspond to H_2ase catalytic (large) subunits, whereas the *nfor* gene listed corresponds to subunit C of NFOR. 16S rRNA expression was used as a reference gene for all analysis. Relative expression of genes analyzed in this study was calculated using the Pfaffl method (Pfaffl 2001) with the 20 mL s^{-1} N_2 gas sparging condition used as the calibrator. All data analysis was performed using the MxPro QPCR software package with the statistical significance of expression ratios determined using the BootstRatio web application (Cleries et al. 2012).

4.4 Results

4.4.1 Cell growth, pH and cellobiose consumption

Growth of *C. thermocellum* ATCC 27405 was monitored in open-batch cultures containing 2 g L^{-1} (5.9 mM) cellobiose in 1191 media continuously sparged with high N_2 (20 mL s^{-1}), low N_2 (2 mL s^{-1}), or H_2 (20 mL s^{-1}; Figure 4.2). Cell densities measured at 600 nm correlated well with biomass production as measured by protein content. Final cell biomass (Table 4.2) concentrations reached 3.6, 3.3, and 2.5 mM for high N_2, low N_2, high H_2 sparged cultures, respectively. In all studies, cellobiose was completely utilized upon entry in stationary phase indicating that growth of these cultures was carbon limited. *C. thermocellum* exponential phase generation times (2.3, 2.2, and 2.7 h for cultures sparged at high N_2, low N_2, and high H_2, respectively) revealed only minor differences in growth rates. However, cellobiose consumption and end-product synthesis rates were much slower during high gas sparging, regardless of gas used (Figure 4.2). Final culture pH was measured to be 6.8, 6.6, and 6.9 in stationary phase cultures sparged with high N_2, low N_2, and high N_2, respectively.

Table 4.2: Effect of gas sparging on final end-product concentrations of *C. thermocellum* grown on 2 g L^{-1} cellobiose. Standard deviations of three replicates are indicated in brackets. Values indicated in bold represent calculated values assuming a carbon recovery ~ 90 % and C1/C2 ratio and an O/R index of 0.9-1.1.

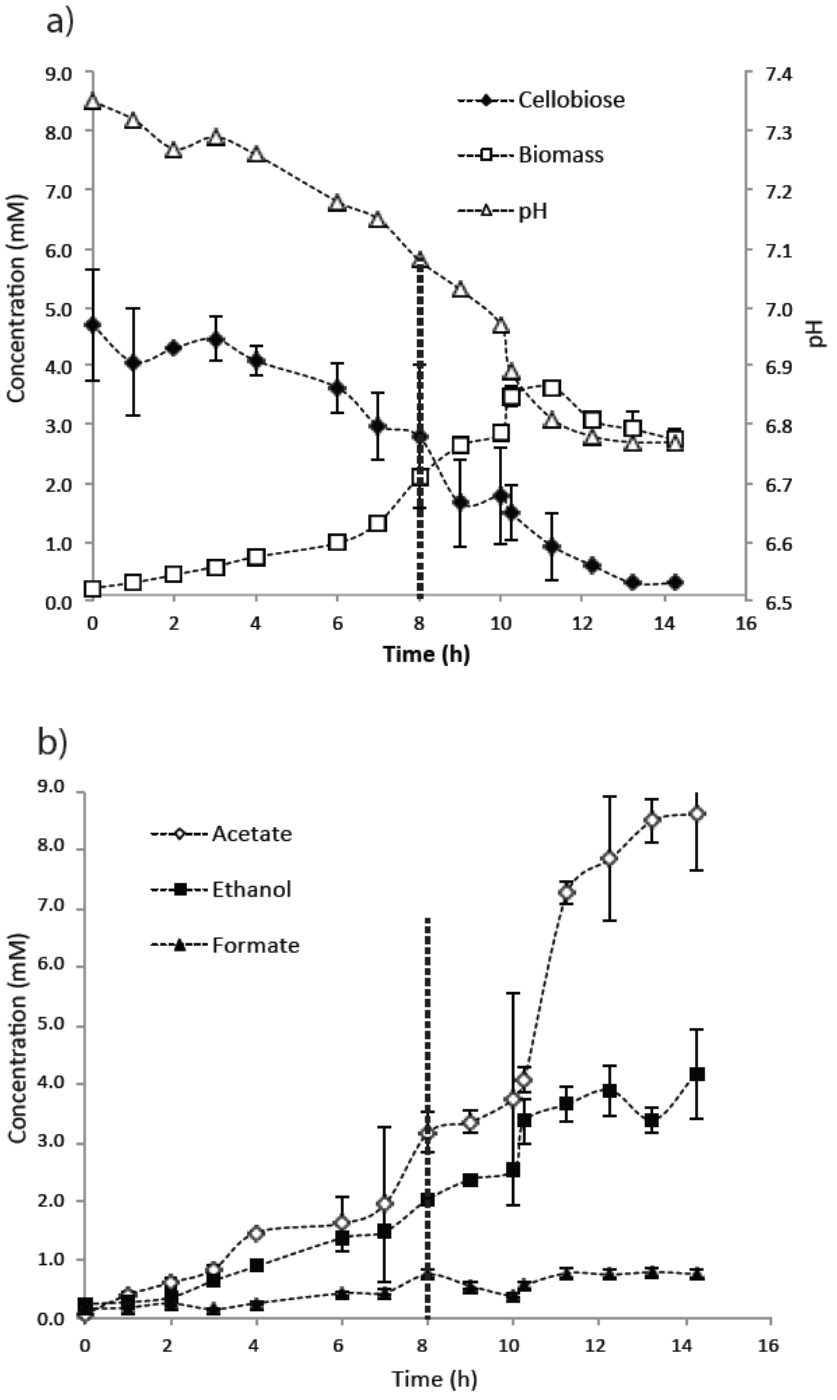
Product (mM)	Reactor Condition		
	20 mL s^{-1} N_2	2 mL s^{-1} N_2	20 mL s^{-1} H_2
Biomass[a]	3.6 (0.2)	3.3 (0.3)	2.5 (0.1)
Acetate	8.6 (0.9)	7.4 (0.2)	4.9 (0.1)
Formate	0.8 (0.1)	1.0 (0.1)	1.5 (0.1)
Lactate	< 0.2[c]	< 0.2[c]	< 0.2[c]
Ethanol[b]	4.3 (0.8)	4.8 (0.2)	7.0 (0.2)
Condensate (% of total produced)	3.46 (0.01)	0.00 (0.00)	9.52 (0.07)
H_2	**16.4**	**13.9**	**10.4**
CO_2	**12.1**	**11.4**	**9.8**
Generation time (h)	2.3 (0.1)	2.2 (0.4)	2.7 (0.4)
Final pH	6.8 (0.1)	6.6 (0.2)	6.9 (0.1)
E:A[b]	0.5	0.6	1.4

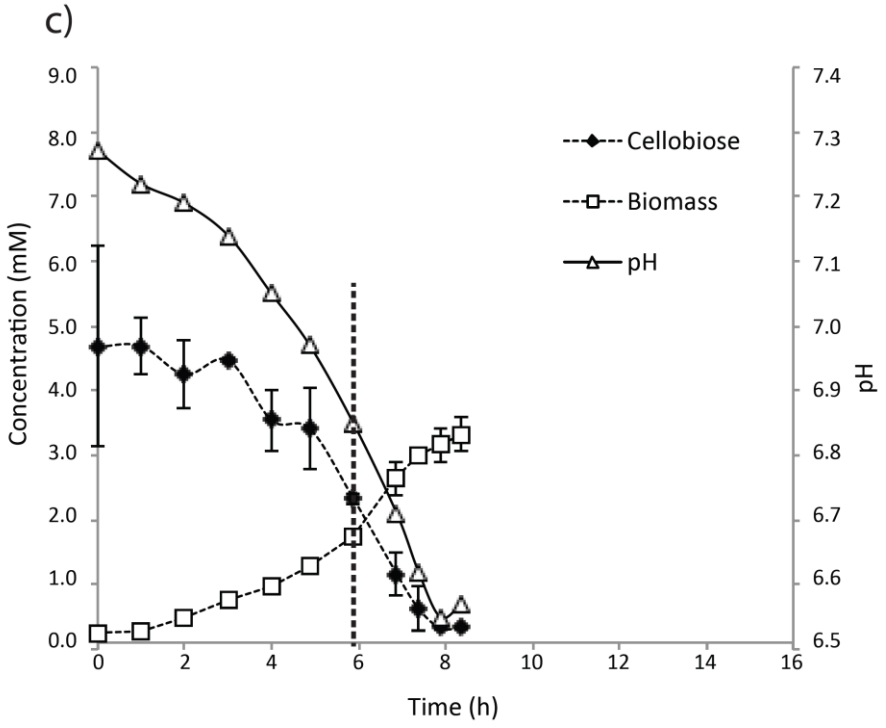
[a]Represents maximal biomass detected prior to cell lysis

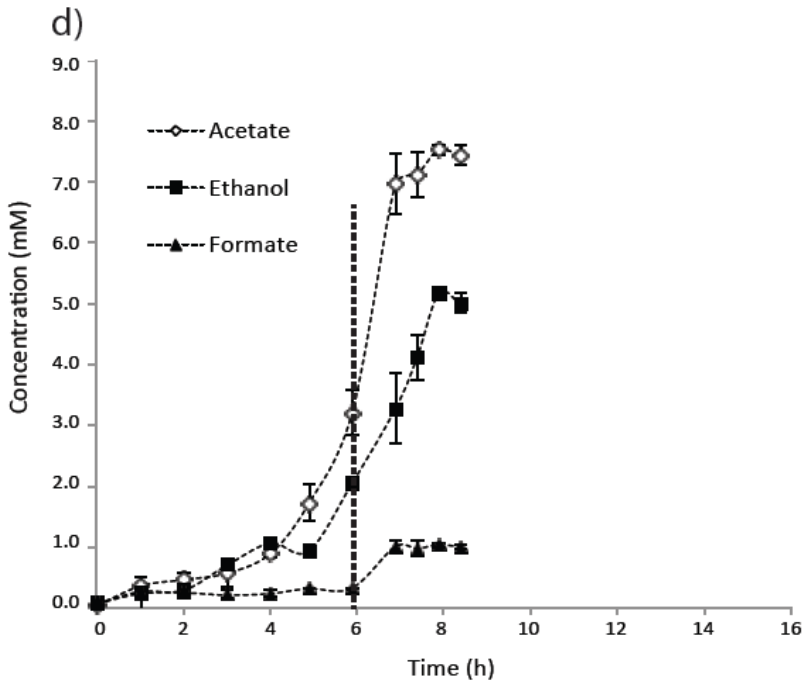
[b]Reflects total amount of ethanol produced (in fermenter and condensate)

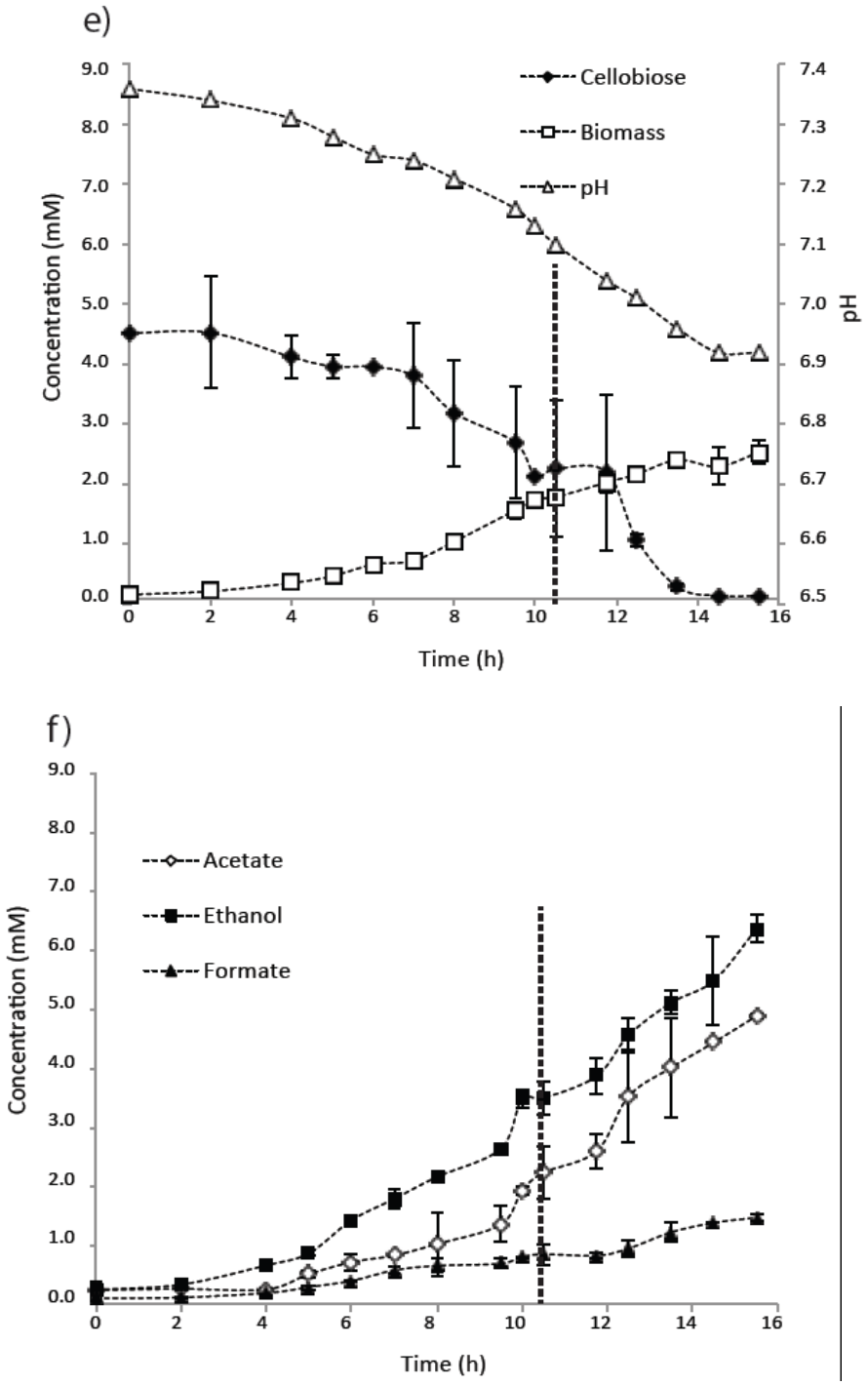
[c] Below detectable limits

Figure 4.2: Growth and fermentation kinetics of *C. thermocellum* in response to gas sparging. Cultures sparged with: 20 mL s^{-1} N_2 (a and b)' 2 mL s^{-1} N_2 (c and d); and 20 mL s^{-1} H_2 (e and f). Cells harvested for enzyme assay and RT-qPCR determination is indicated by a vertical dotted line.


a)
Cellobiose
Biomass
pH
Concentration (mM)
pH
Time (h)
b)
Acetate
Ethanol
Formate
Concentration (mM)
Time (h)


c)
Concentration (mM)
pH
Time (h)
Cellobiose
Biomass
pH
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
7.4
7.3
7.2
7.1
7.0
6.9
6.8
6.7
6.6
6.5
0
2
4
6
8
10
12
14
16


d)
Concentration (mM)
Time (h)
Acetate
Ethanol
Formate
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0
2
4
6
8
10
12
14
16


e)
Cellobiose
Biomass
pH
Concentration (mM)
pH
Time (h)
f)
Acetate
Ethanol
Formate
Concentration (mM)
Time (h)

4.4.2 End-product synthesis

Acetate, ethanol, formate, CO_2 and H_2 were the major end-products of cellobiose fermentation under all sparging conditions (Figure 4.2, Table 4.2). Although both lactate dehydrogenase (LDH) enzyme activity and *ldh* transcription were detected at significant levels, negligible lactate was produced (< 0.2 mM) during fermentation (discussed below).

When batch cultures of *C. thermocellum* were sparged with N_2 at 2 mL s^{-1}, 7.4 (± 0.2) mM acetate, 4.8 (± 0.2) mM ethanol, and 1.00 (± 0.03) mM formate were produced (Table 4.2). Final acetate concentrations increased by 16 % when gas sparging with N_2 was increased to 20 mL s^{-1}. Additionally, final ethanol and formate concentrations decreased by 14 % and 31 %, respectively, when compared to cultures sparged at 2 mL s^{-1}. Sparging cultures with H_2 at 20 mL s^{-1} had the most dramatic effect on end-product distribution profiles. Compared to cultures of *C. thermocellum* sparged at 20 mL s^{-1} N_2, acetate production was reduced by 52 %, while ethanol and formate production increased by 34 % and 47 % respectively (Table 4.2).

Ethanol:acetate (E:A) ratios are illustrative of how carbon and electron flux was affected by gas sparging with either N_2 or H_2. When sparged with high N_2 and low N_2, final molar ethanol-to-acetate ratios were 0.5, and 0.6. In contrast, when *C. thermocellum* was sparged with H_2 at 20 mL s^{-1}, an E:A ratio of 1.4 was observed. Given that the bioreactor was subjected to continuous gas sparging and was openly vented, we collected condensates for end-product determinations. Ethanol collected in the

condensate accounted for 3.5% and 9.5% of the total ethanol produced during high N_2 and high H_2 sparging, respectively. Ethanol in condensate acquired during low N_2 sparging accounted for less than 0.001% of total ethanol produced. Acetate, lactate and formate concentrations in collected condensates were below detectable limits in each of the samples.

4.4.3 Enzyme activities

Enzyme activities involved in the malate shunt, pyruvate catabolism, H_2 metabolism, and ethanol metabolism were assessed on mid-exponential phase cell extracts grown under low N_2, high N_2, and high H_2 gas sparging (Table 4.3). MDH and MalE enzyme activities did not change in response to sparging conditions. However, MDH activity was demonstrated to be both NADH and NADPH-dependent. MalE activity was strictly NADPH-dependent. Although LDH activity was detected at relatively uniform levels under all conditions, lactate production was not detected. PFOR activity increased only marginally (1.4-fold) under low sparging conditions versus high sparging conditions (regardless of gas used), and PFL activity did not change in response to sparging conditions. No significant changes in methyl viologen-, NADH-, NADPH-, or ferredoxin-dependent H_2ase activities were observed. Bifurcating H_2ase activities were not assessed in crude cell extracts given that this type of H_2ase was not discoverd until after enzyme activity experiments were performed. While NADPH-dependent ADH activity did not change in response to gas sparging, NADH-dependent ADH activity increased 2-fold when cultures were sparged with high H_2 versus high N_2.

Table 4.3: Effect of gas sparging on exponential phase enzymes activities involved in the malate shunt, pyruvate catabolism, and H_2 and ethanol fermentation. Changes greater that 2-fold are indicated in bold. Standard deviations are provided in parentheses

Enzyme	Substrate	Specific Activity (nmol min^{-1} mg-$protein^{-1}$)		
		20 mL s^{-1} N_2	2 mL s^{-1} N_2	20 mL s^{-1} H_2
Malic enzyme[a]	Malate: NADP+	78 (2)	86 (2)	81 (2)
Malate dehydrogenase	Oxaloacetate: NADH	46 (2)	ND[1]	46 (6)
	Oxaloacetate: NADPH	20 (2)	27 (2)	21 (2)
Lactate dehydrogenase	Pyruvate: NADH: FDP	11 (3)	13 (2)	10 (4)
Pyruvate:formate lyase	Pyruvate: CoA	11 (3)	11 (1)	13 (2)
Pyruvate:ferredoxin oxidoreductase	Pyruvate: CoA: MV	218 (8)	327 (21)	256 (30)
Hydrogenase	H_2: MV	445 (82)	547 (90)	506 (56)
	H_2: Fd: metronidazole	1 (<1)	1 (< 1)	1 (< 1)
	H_2: NAD^+	4 (1)	6 (2)	3 (1)
	H_2: $NADP^+$	10 (3)	8 (1)	7 (1)
Alcohol dehydrogenase	Acetaldehyde: NADH	**40 (14)**	61 (5)	**84 (29)**
	Acetaldehyde: NADPH	15 (1)	15 (1)	18 (3)

[1]ND: Not determined

4.4.4 Gene expression

The transcription of 17 protein-encoding genes involved in the malate shunt, pyruvate catabolism, and H_2 and ethanol metabolism were assessed under different gas sparging conditions using RT-qPCR. Relative expression values were normalized to the transcription of reference gene, 16S rRNA, which typically was 2 to 4 orders of magnitude greater than the expression of the assayed genes. Gene expression ratios relative to cultures sparged with 20 mL s^{-1} N_2 gas are provided in Table 4.4. Most genes analyzed displayed less than a 2-fold change in expression in response to gas sparging conditions. However, transcription of *adhE* increased 10-fold in response to high H_2 sparging, when compared to high N_2 sparging, while transcription of *adhZ* decreased 2.7-fold. Transcription of *adhY* increased by more than 5-fold in cells grown under low N_2 sparging when compared to high N_2 sparging. A moderate (~2-fold) increase in two *pfor* genes (Cthe__2392 and Cthe__3120) analyzed was observed under low N_2 sparging when compared to high N_2 or H_2 gas sparging.

Table 4.4: Relative expression of genes involved in pyruvate catabolism, ethanol and hydrogen production in *C. thermocellum* ATCC 27405 during mid-exponential phase. Significant changes greater than 2-fold are indicated in bold.

Gene	Locus	Relative Expression Ratios Normalized to 20 mL s^{-1} N_2					
		2 mL s^{-1} N_2			20 mL s^{-1} H_2		
		RE[a]	SE[b]	P[c]	RE[a]	SE[b]	P[c]
Malate shunt							
malE	Cthe_0344	0.846	(0.187)	NS	0.645	(0.142)	NS
mdh	Cthe_0345	0.544	(0.018)	<0.01	0.685	(0.040)	<0.01
Pyruvate catabolism							
ldh	Cthe_1053	1.317	(0.131)	NS	1.456	(0.176)	<0.05
pfl	Cthe_0505	0.774	(0.352)	NS	0.746	(0.206)	NS
pfor	Cthe_2392	**2.114**	(0.059)	<0.01	1.747	(0.204)	<0.01
pfor	Cthe_2796	1.754	(0.100)	<0.01	1.499	(0.019)	<0.01
pfor	Cthe_3120	**2.298**	(0.236)	<0.01	1.899	(0.186)	<0.01
Ethanol Synthesis							
aldH	Cthe_2238	1.233	(0.810)	<0.05	3.076	(0.719)	NS
Fe-adh	Cthe_0101	0.555	(0.013)	<0.01	0.610	(0.102)	<0.01
adhY	Cthe_0394	**5.758**	(0.343)	<0.01	1.714	(0.233)	<0.05
adhE	Cthe_0423	4.245	(2.144)	NS	**10.074**	(2.753)	<0.01

adhZ	Cthe_2579	0.522	(0.154)	NS	**0.376**	(0.134)	<0.05
Hydrogen metabolism							
[FeFe] H_2ase	Cthe_0342	0.758	(0.086)	NS	0.778	(0.124)	NS
[FeFe] H_2ase	Cthe_0430	1.313	(0.307)	NS	1.681	(0.204)	<0.01
[FeFe] H_2ase	Cthe_3003	1.866	(0.219)	<0.01	1.817	(0.332)	<0.05
[NiFe] H_2ase	Cthe_3020	1.442	(0.227)	NS	1.078	(0.144)	NS
Electron transfer							
nfor	Cthe_2430	0.749	(0.108)	<0.05	0.606	(0.108)	<0.01

[a] RE: median observed relative expression ratio normalized to 16s rRNA expression and calibrated against 20 mL s^{-1} N_2 gas sparging cultures;

[b] SE: standard error of the observed gene expression ratio;

[c] P: Approximation to p-value for bootstrap sampling of mean ratio. If the median boot ratio is > 1 then P is the relative frequency of median ratio < 1. If median boot ratio is < 1 then P is the relative frequency of median boot ratio > 1; NS, not significant (P > 0.05)

4.5 Discussion

Growth of *C. thermocellum* requires both ATP, which is generated via substrate level phosphorylation (during glycolysis and acetate production), and the maintenance of cellular redox. NADH reduced during glycolysis may be oxidized directly through either lactate, ethanol, or H_2 production via LDH, AldH/ADH, and NAD(P)H-dependent hydrogenases, respectively. Alternatively, ferredoxin (Fd) reduced via PFOR can be oxidized via Fd-dependent hydrogenase, or via bifurcating H_2ases, which simultaneously oxidize reduced Fd and NADH to produce H_2 (Schut and Adams 2009; Rydzak et al. 2012). H_2ase mediated oxidation of reducing equivalents, however, becomes thermodynamically unfavourable when H_2 is permitted to accumulate, which may lead to changes in electron (and carbon) flux distribution (Lamed et al. 1988).

Increases in formate and ethanol production and a decrease in acetate production was observed in response to sparging cultures of *C. thermocellum* with H_2 compared to N_2 (20 mL s^{-1}). Ethanol-to-acetate ratios for N_2 sparged cultures (0.5-0.6) were consistent with reports of stirred *C. thermocellum* cultures grown on cellobiose under 2.5 atm N_2 (Lamed et al. 1988). Additionally, the shift towards ethanol production that was observed when cultures were sparged with H_2 was similar to the findings of Lamed *et al.* (Lamed et al. 1988) when *C. thermocellum* strains YS, AS-39, and LQRI were grown under 2.5 atm H_2 with stirring. Li *et al* (Li et al. 2012a) reported a 35 % decrease in acetate production and a 350 % increase in ethanol production, corresponding to a shift in E:A ratio of 0.12 to 0.72, when the sparging gas of steady-state cultures of *C. thermocellum* cultures was switched from N_2 to H_2.

This shift in metabolic flux is likely a consequence of thermodynamic barriers to the oxidation of NAD(P)H and/or reduced Fd to H_2 (Thauer et al. 1977; Angenent et al. 2004). Accumulation of reduced Fd may result in a bottleneck at PFOR, redirecting flux through PFL, as supported by decreased CO_2 and increased formate production in response to H_2 sparging. Additionally, increased NAD(P)H pools must be oxidized via ethanol formation, rather than H_2 production, which in turn leads to decreases in acetate production from acetyl-CoA. The observed decrease in maximal biomass production correlated with decreased acetate production, and may be a consequence of reduced ATP production per mole cellulose fermented.

The critical parameter associated with H_2ase inhibition during fermentation has, traditionally, been partial H_2 pressure (P_{H2}) (Lamed et al. 1988). More recent studies, however, are demonstrative of liquid-to-gas mass-transfer limitations that may cause H_2 super-saturation of the fermentation medium and, consequently, H_{2aq} is a more appropriate predictor of hydrogenase inhibition (Pauss et al. 1990; Kraemer and Bagley 2006). Indeed, replacement of a N_2 headspace with H_2 in closed-batch cultures of *C. thermocellum* (no stirring) grown on cellobiose was shown to increase ethanol production by only 10 % (Rydzak et al. 2011). Angenenet *et al*. (Angenent et al. 2004) have suggested that, in addition to the 2 mol H_2 that may be evolved from the ferredoxin reduced during the PFOR-mediated oxidative decarboxylation of pyruvate, H_2 may also be produced from NADH (either directly using NADH-dependent H_2ases or indirectly using NADH:ferredoxin oxidoreductase/Fd-H_2ase) at $H_{2aq} < 0.5\mu M$ (<60 Pa). However, in a report examining super-saturation of dissolved gasses during fermentative growth,

the common assumption, that sparging with N_2 may increase H_2 yields due to lower dissolved H_2, was challenged by assertions that H_{2aq} may not be decreased sufficiently to permit reduction of ferredoxin by NADH:Fd oxidoreductase (Kraemer and Bagley 2006). This explanation, however, neither considers the redox implication resultant from a recently described (Schut and Adams 2009) and highly prevalent (Carere et al. 2012) bifurcating type hydrogenase, nor the potential for metabolic regulation. Ethanol production was virtually eliminated in continuous open-batch cultures of *Caldanaerobacter subterraneus* subsp. *tengcongensis* sparged with N_2, suggesting that 4 moles of H_2 were produced per mol glucose (Soboh et al. 2004). Although the maximum H_2 yields reported here, were below 2 mol per mol glucose, a high rate of N_2 sparging shifted metabolism from the production of ethanol and formate towards acetate, H_2, and CO_2; likely as a consequence of decreased H_{2aq}.

While the majority of enzyme activities did not change significantly in response to sparging conditions, consistent with negligible changes in transcription levels, NADH-dependent ADH activity increased 2.1-fold in response to high H_2 sparging compared high N_2 sparging, whereas PFOR activity was 1.3-1.5-fold higher during low gas sparging compared to high gas sparging (Table 4.3). Given that *C. thermocellum* encodes multiple PFORs and ADHs, RT-qPCR analysis was performed to resolve which genes are regulated and responsible for observed changes in enzyme activities in response to gas sparging.

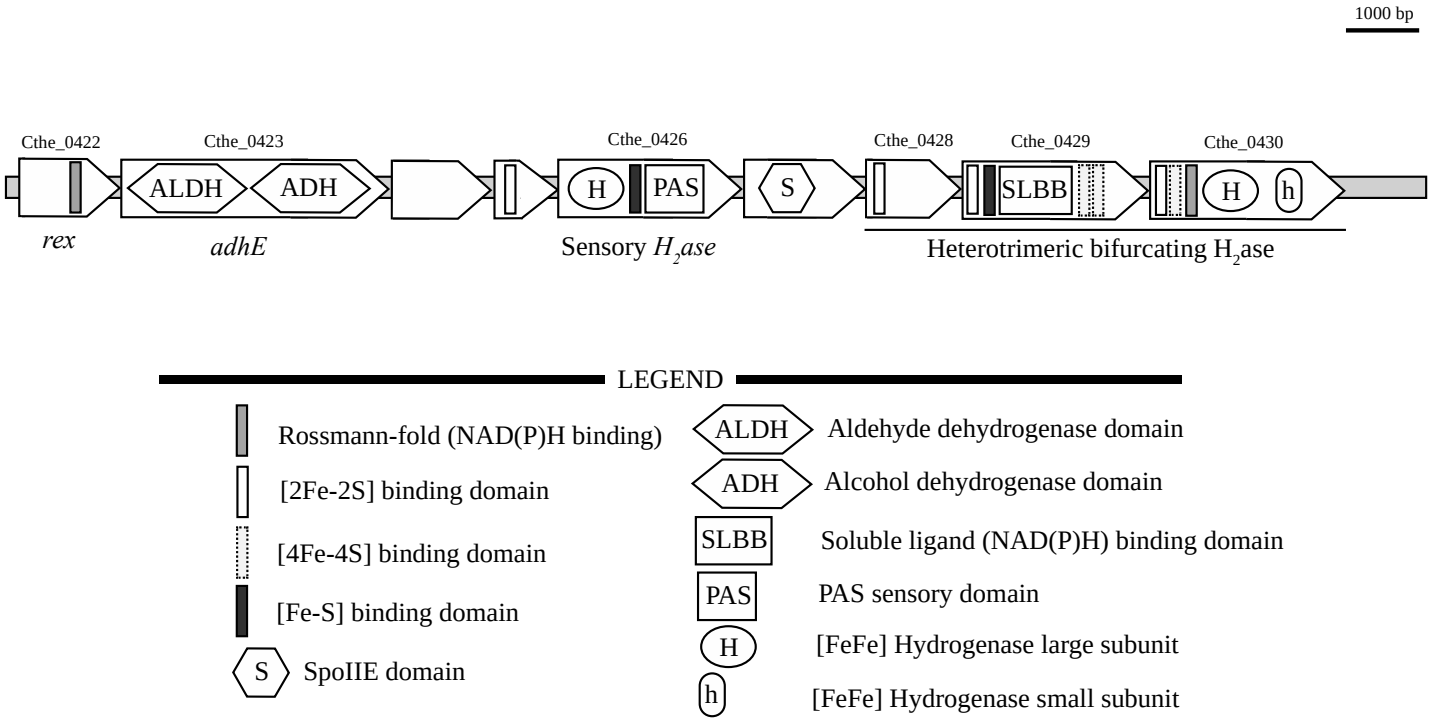
While *C. thermocellum* encodes multiple PFORs (Cthe__2390-2393, Cthe__3120, Cthe__0866, Cthe__0614, Cthe__2796-2797), 2D-HPLC-MS/MS and microarray

expression profiles show that Cthe__2390-2393 and Cthe__3120 are the principal PFORs responsible for pyruvate catabolism (Raman et al. 2011; Rydzak et al. 2012). Transcription of both Cthe__2396 and Cthe__3120, was 2-fold greater in cultures sparged with low N_2 verses high N_2, consistent with a 1.5-fold increased in PFOR activity. While it is unclear as to why expression of these genes/proteins is elevated during low sparging conditions, it is tempting to suggest that higher dissolved CO_2 and/or bicarbonate levels may induce their expression.

The genome of *C. thermocellum* encodes multiple iron-containing alcohol dehydrogenases including Cthe__0394 (*adhY*), Cthe__2579 (*adhZ*), Cthe__0101 (*adh*) and a bifunctional acetaldehyde/alcohol dehydrogenase (Cthe__0423; *adhE*). In addition to *adhE*, the conversion of acetyl-CoA to acetaldehyde may be catalyzed by an NADH-dependent aldehyde dehydrogenase (Cthe__2238; *aldH*). While crude cell free extract displays both NADH- and NADPH-dependent alcohol dehydrogenase activities, identifying cofactor specificity on the basis of sequence analysis is problematic. Proteomic analysis of the core catabolic enzymes of *C. thermocellum*, however, has shown AdhE to be the most abundant ADH expressed. The alcohol dehydrogenases encoded by Cthe__0394, Cthe__0101, and Cthe__2579 were reported at 78 %, 24 % and 9 % abundance relative to the level of AdhE expression, respectively (Rydzak et al. 2012). The relative abundance of AdhE and the low level of reported AldH expression suggest that AdhE is the primary enzyme involved in the reduction of acetyl-CoA to ethanol.

Transcription of *adhE*, which increased 10-fold during high H_2 sparging when compared to high N_2 sparging, may be regulated via a '*Rex*'-type redox-sensing transcriptional repressor (Cthe__0422, Figure 4.3). These redox-sensing proteins have been implicated in regulating the expression of genes important for fermentative growth in many Gram-positive bacteria (Brekasis and Paget 2003; Gyan et al. 2006). In *Bacillus subtilis* Rex senses NADH/NAD^+ and represses gene expression by binding to 20-bp AT-rich semi-palindromic operator region downstream of the *yjlC-ndh* promoter. Repression is greatest under low NADH/NAD^+ ratios with de-repression occurring as NADH/NAD^+ ratios increase (Gyan et al. 2006). The 11.5 kbp region of *C. thermocellum's* genome that encompasses Cthe__0422 to Cthe__0431 includes genes that encode the Rex-type redox sensing protein (Cthe__0422), AdhE (Cthe__0423), and both a PAS-fold containing sensory H_2ase and a heterotrimeric bifurcating H_2ase (Cthe__0426, Cthe__0428-0430). This gene cassette appears poised to regulate ethanol and H_2 metabolism in response to cellular redox with de-repression of *adhE* transcription likely occurring with increasing H_{2aq}. Poor homology of '*Rex*'-type operator sequences, however, makes it difficult to predict to what extent ADH expression may be influenced.

Figure 4.3: Illustration of 11.5 kbp genomic region of *Clostridium thermocellum* ATCC 27405 that includes ORFs Cthe_0422-Cthe_0430. Included in this gene cassette is a redox sensing transcription regulator (*rex*, Cthe_0422), a bifunctional acetaldehyde/alcohol dehydrogenase (*adhE*, Cthe_0423), and both a PAS-fold containing sensory H_2ase and a heterotrimeric bifurcating H_2ase (Cthe_0426 and Cthe_0428-0430, respectively).



In *C. thermocellum*, there is some precedent suggesting regulation of ADHs. Transcriptomic and RT-qPCR analysis, have reported increased expression of *aldh* and reduced expression of Cthe__0101 in stationary phase cultures (Raman et al. 2011) and differential expression of *adhY* in response to growth rate (Stevenson and Weimer 2005). The analysis reported here revealed *adhY* (Cthe__0394) transcription increased ~ 5-fold under low sparging conditions when compared to high sparging and a 2.7-fold decrease in *adhZ* (Cthe__2579) transcription in response to high H_2 sparging, an enzyme of low abundance in previous microarray and proteomic studies (Raman et al. 2011; Rydzak et al. 2012). It is possible that ethanol metabolism in *C. thermocellum* is regulated, as in *Thermoanaerobacter ethanolicus* JW200, with multiple ADHs differentially expressed in response to increased ethanol concentrations and an accumulation of oxidized reducing equivalents (Pei et al. 2010). In *T. ethanolicus*, a redox sensing protein (RSP) regulates the operons for ethanol fermentation in response to redox signals (Pei et al. 2010). RSP, which shares 41 % amino acid identity with Rex, also binds to palindromic operators sequences of various ADHs, repressing their transcription. Further research must be performed to elucidate the kinetic properties of various ADHs encoded by *C. thermocellum* and the factors that affect their gene expression.

4.6 Conclusions

C. thermocellum possesses numerous mechanisms to maintain cellular redox that ultimately results in lactate, ethanol, or H_2 production. These mechanisms employ various enzymes that are, in many cases, redundant. Manipulation of gas sparging rates, and thus H_{2aq}, had a significant impact on fermentation product yields. While the

majority of enzyme activities and transcript levels did not change in response to gas sparging, suggesting that thermodynamics play a significant role in manipulation of carbon and electron fluxes, changes in ADH transcript levels suggest that ethanol production is in fact regulated by changes in gene/enzyme expression. The observed changes in ADH expression patterns are likely regulated by redox signals that modulate transcriptional regulators, and merit further investigation. The shifts in enzyme use observed in this study, while subtle, appear have an important affect on metabolic fluxes and could impact flux models under different growth conditions.

4.7 Authors` contributions

Thomas Rydzak and Carlo R. Carere drafted the manuscript (equal contributions). TR and CRC performed the fermentations and end-product analysis. Thomas Rydzak performed enzyme activity analysis with help from Carlo R, Carere. Carlo R. Carere performed RT-qPCR analysis with help from Thomas Rydzak. Richard Sparling, Nazim Cicek, and David B. Levin conceived of the study, participated in its design, and helped draft the manuscript.

4.8 Acknowledgments

This work was supported by funds provided by Genome Canada, through the Applied Genomics Research in Bioproducts or Crops (ABC) program for the grant titled, “Microbial Genomics for Biofuels and CoProducts from Biorefining Processes”, and by the Province of Manitoba, Agricultural and Rural Development Initiative (ARDI), grant 09-986.

Chapter 5: Linking genome content to biofuel production yields: A meta-analysis of major catabolic pathways among select H_2 and ethanol-producing bacteria[19]

5.1 Abstract

Fermentative bacteria offer the potential to convert lignocellulosic waste-streams into biofuels such as hydrogen (H_2) and ethanol. Current fermentative H_2 and ethanol yields, however, are below theoretical maxima, vary greatly among organisms, and depend on the extent of metabolic pathways utilized. For fermentative H_2 and/or ethanol production to become practical, biofuel yields must be increased. We performed a comparative meta-analysis of (i) reported end-product yields, and (ii) genes encoding pyruvate metabolism and end-product synthesis pathways to identify suitable biomarkers for screening a microorganism's potential of H_2 and/or ethanol production, and to identify targets for metabolic engineering to improve biofuel yields. Our interest in H_2 and/or ethanol optimization restricted our meta-analysis to organisms with sequenced genomes and limited branched end-product pathways. These included members of the Firmicutes, Euryarchaeota, and Thermotogae. Bioinformatic analysis revealed that the absence of genes encoding acetaldehyde dehydrogenase and bifunctional

---

[19] Thomas Rydzak[†], Carlo Carere[†], Tobin J. Verbeke, Nazim Cicek, David B. Levin and Richard Sparling. *BMC Microbiology* 12, no. 1 (2012): 295. ([†]Equal contributors)

acetaldehyde/alcohol dehydrogenase (AdhE) in *Caldicellulosiruptor*, *Thermococcus*, *Pyrococcus*, and *Thermotoga* species coincide with high H_2 yields and low ethanol production. Organisms containing genes (or activities) for both ethanol and H_2 synthesis pathways (i.e. *Caldanaerobacter subterraneus* subsp. *tengcongensis*, *Ethanoligenens harbinense*, and *Clostridium* species) had relatively uniform mixed product patterns. The absence of hydrogenases in *Geobacillus* and *Bacillus* species did not confer high ethanol production, but rather high lactate production. Only *Thermoanaerobacter pseudethanolicus* produced relatively high ethanol and low H_2 yields. This may be attributed to presence of genes encoding proteins that promote NADH production. Lactate dehydrogenase and pyruvate:formate lyase are not conducive for ethanol and/or H_2 production. While the type(s) of encoded hydrogenases appear to have little impact on H_2 production in organisms that do not encode ethanol producing pathways, they do influence reduced end-product yields in those that do. Here we show that composition of genes encoding pathways involved in pyruvate catabolism and end-product synthesis pathways can be used to approximate potential end-product distribution patterns. We have identified a number of genetic biomarkers for streamlining ethanol and H_2 producing capabilities. By linking genome content, reaction thermodynamics, and end-product yields, we offer potential targets for optimization of either ethanol or H_2 yields through metabolic engineering.

5.2 Introduction

Fuel derived from waste-stream lignocellulosic biomass via consolidated bioprocessing is a renewable and carbon-neutral alternative to current petroleum-based

fuels (Levin et al. 2004; Lynd et al. 2005; Carere et al. 2008b). Consequently, considerable effort is being made to characterize species capable of efficiently converting lignocellulosic substrates into biofuels. An ideal biofuel producing microorganism should posses several key features, including: (i) high yields of the desired product, (ii) simultaneous utilization of sugars (cellulose, hemicellulose, pectin), and (iii) growth at elevated temperatures, and (iv) low product inhibition. Recent studies have focused on the characterization of numerous cellulose and hemicellulose degrading species of bacteria (Desvaux 2005; Islam et al. 2009; Yang et al. 2009). To fully exploit the biofuel producing potential of these organisms, several genomes have been sequenced and are now available for analysis (http://genome.jgi-psf.org/). While some hemicellulolytic or cellulolytic microorganisms are capable of hydrogen (H_2) or ethanol production via fermentation, end-product yields typically are far lower than their maximum theoretical values (4 mol H_2 or 2 mol ethanol per mol glucose) when cells are grown in pure culture. This is due to the presence of branched catabolic pathways that divert carbon and/or electrons away from a particular desired end-product (Hallenbeck 2002). Strategies that optimize yields for a single biofuel (H_2 or ethanol) can only be developed through a detailed knowledge of the relationships between genome content, gene and gene product expression, pathway utilization, and end-product synthesis patterns.

Given that our primary focus is to optimize H_2 and/or ethanol yields, we restricted our meta-analysis to sequenced organisms with limited branched end-product pathways (i.e. organisms that do not produce butyrate, butanol, propionate, propanol, and acetoin) for which end-product data was available. These included members of the Firmicutes

(*Clostridium*, *Caldicellulosiruptor*, *Thermoanaerobacter*, *Caldanaerobacter*, *Ethanoligenens*, *Geobacillus*, and *Bacillus* species), Euryarchaeota (*Thermococcus* and *Pyrococcus* species), and Thermotogae (*Thermotoga* species). A list of species analyzed and corresponding GenBank accession numbers are summarized in Table 5.1. With the exception of *Caldanaerobacter subterraneus* subsp. *tengcongensis*, *Thermoanaerobacter pseudethanolicus*, *Pyrococcus furiosus*, *Geobacillus thermoglucosidasius*, and *Bacillus cereus*, all organisms were capable of cellulose and/or xylan saccharification.

Table 5.1: H_2 and ethanol producing organisms included in meta-analysis of end-product yields and genome content. National Center for Biotechnology Information taxon IDs, GenBank accession numbers, corresponding sequencing centers responsible for the generation of the genome sequences data analyzed in this study are provided. Phyla (F; Firmicutes: E; Euryarchaeota: T; Thermotogae), and polymeric carbon sources degraded (S; starch: C; cellulose: X; xylose) by each organism are indicated.

Organism	Synonyms	Taxon ID	GenBank #	Sequencing Center	Phyla	C sources
Caldicellulosiruptor saccharolyticus DSM 8903		351627	NC_009437	DOE Joint Genome Institute	F	S,C,X
Caldicellulosiruptor besci DSM 6725	*Anaerocellum thermophilum*; Z-1320	521460	NC_012036	DOE Joint Genome Institute	F	S,C,X
Pyrococcus furiosus DSM 3638		186497	AE009950	Univ of Maryland, Univ of Utah	E	S,C,X
Thermococcus kodakaraensis KOD1		69014	NC_006624	Kwansei Gakuin Univ, Kyoto University	E	S
Thermotoga neapolitana DSM 4359	ATCC 49049; JCM 10099; NS-E	309803	NC_011978	Genotech corp.	T	S,C
Thermotoga petrophila RKU-1		390874	NC_009486	DOE Joint Genome Institute	T	S,C,X
Thermotoga maritima MSB8	DSM 3109	243274	NC_000853	J. Craig Venter Institute	T	S,C,X

Caldanaerobacter subterraneus subsp. *tengcongensis* MB4	*Thermoanaerobacter tengcongensis*	273068	NC_003869	Beijing Genomics Institute, The Institute of Microbiology, China	F	S
Ethanoligenens harbinense YUAN-3T	DSM 18485	663278	NC_014828	DOE Joint Genome Institute	F	S,C
Clostridium cellulolyticum H10		394503	NC_011898	DOE Joint Genome Institute	F	S,C,X
Clostridium phytofermentans ISDg	ATCC 700394	357809	NC_010001	DOE Joint Genome Institute	F	S,C,X
Clostridium thermocellum ATCC 27405	DSM 1237	203119	NC_009012	DOE Joint Genome Institute, University of Rochester	F	S,C,X
Clostridium thermocellum DSM 4150	JW20	492476	ABVG00000000	DOE Joint Genome Institute	F	S,C,X
Thermoanaerobacter pseudethanolicus 39E	ATCC 33223	340099	NC_010321	DOE Joint Genome Institute	F	S,X
Geobacillus thermoglucosidasius C56-YS93		634956	NC_015660	DOE Joint Genome Institute	F	S
Bacillus cereus ATCC 14579	DSM 31	226900	NC_004721	Integrated Genomics Inc.	F	S

We focused on the various metabolic branches involved in pyruvate formation from phosphoenolpyruvate (PEP) and subsequent catabolism of pyruvate into end-products. Although studies comparing the H_2 and ethanol-producing potential of several cellulose degrading bacteria have been previously published (Desvaux 2006; Bruggemann and Gottschalk 2008; Rydzak et al. 2009), a comprehensive comparison of the major biofuel producing pathways at the genome level has not yet been reported. Here we present a comparison of the genes encoding proteins involved in (i) pyruvate metabolism, (ii) ethanol synthesis, and (iii) H_2 metabolism, in order to rationalize reported end-product yields. Results indicate that the presence or absence of specific genes dictating carbon and electron flow towards end-products may be used to infer end-product synthesis patterns and help develop informed metabolic engineering strategies for optimization of H_2 and ethanol yields. Furthermore, certain genes may be suitable biomarkers for screening novel microorganisms' capability of producing optimal H_2 or ethanol yields, and may be suitable targets for metabolic engineering strategies for optimization of either ethanol or H_2 yields

5.3 Materials and methods

5.3.1 Comparative analysis of genome annotations

All sequence data and gene annotations were accessed using the Joint Genome Institute's Integrated Microbial Genomes (IMG) database (Markowitz et al. 2006). Gene annotations presented in this paper reflect the numbering of the final assembly or most recent drafts available (July, 2012). Comparative analyses were performed using the

IMG database. In brief, analyses of all genomes (Table 5.1) were conducted using three annotation databases independently: i) Clusters of Orthologs Groups (COGs)(Tatusov et al. 2003), ii) KEGG Orthology assignments (KO) (Kanehisa et al. 2008), and (iii) TIGRFAMs (Haft et al. 2001). Genes identified using a single database were cross-referenced against the others to identify genes of interest. Functional annotations of the identified genes were evaluated on a case-by-case basis and decisions regarding the annotation accuracy were made using a combination of manual analysis of genomic context, literature searches, and functional prediction through RPS-BLAST using the Conserved Domain Database website (Altschul et al. 1990).

Hydrogenases were classified based on phylogenetic relationships of hydrogenase large subunits according to Calusinska *et al*. (Calusinska et al. 2010). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 50 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 863 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). Thermodynamic calculations were performed using values provided by Thauer *et al.* (Thauer et al. 1977) and the CRC Handbook of Chemistry and Physics (Chemical Rubber Company. 1977; Thauer et al. 1977). BioEdit v.7.0.9.0 (Hall 1999) was used to perform sequence alignments.

5.4 Results and discussion

5.4.1 Survey of end-product yields

A literature survey of end-product yields (normalized to mol end-product per mol hexose equivalent) of the species surveyed in this study is summarized in Table 5.2. While it is difficult to perform a direct comparison of end-product yields from available literature due to different growth conditions employed (ex. growth substrate, carbon loading, reactor conditions, etc.), and further difficult to validate these data due to incomplete end-product quantifications and lack of corresponding carbon balances and oxidation/reduction (O/R) ratios, it still provides a good approximation of molar end-product yields based on substrate utilization. Calculated end-product yields reveal that the *Caldicellulosiruptor*, *Pyrococcus*, *Thermococcus*, and *Thermotoga* species surveyed, produced, in most cases, near-maximal H_2 yields with concomitant CO_2 and acetate production, and little or no ethanol, formate, and lactate (Svetlichnyi et al. 1990; Schicho et al. 1993; Kengen et al. 1994; Schröder et al. 1994; Bredholt et al. 1999; Nguyen et al. 2001; Takahata et al. 2001; Kadar et al. 2004; Kanai et al. 2005; Chou et al. 2007; de Vrije et al. 2007; Eriksen et al. 2008; Nguyen et al. 2008; Kataeva et al. 2009; Munro et al. 2009; Lakhal et al. 2010; Nguyen et al. 2010). It is important to note that while some studies (Schicho et al. 1993; Kengen et al. 1994; Chou et al. 2007; Eriksen et al. 2008; Lakhal et al. 2010; Nguyen et al. 2010) report lower overall end-product yields, likely due to a large amount of carbon flux being directed towards biomass production under a given growth condition, H_2:ethanol ratios remain high. *Cal. subterraneus* subsp. *tengcongensis*, *E. harbinense*, and *Clostridium* species displayed mixed end-product

fermentation patterns, with comparatively lower H_2, CO_2, and acetate yields, higher ethanol yields, and generally low formate and lactate yields (Freier et al. 1988; Xue et al. 2001; Warnick et al. 2002; Soboh et al. 2004; Islam et al. 2006; Xing et al. 2006; Ren et al. 2007; Rydzak et al. 2009). *Ta. pseudethanolicus* produced the highest ethanol yields of the organisms surveyed with little concomitant H_2, acetate, and lactate production, and no formate synthesis (Wiegel and Ljungdahl 1981; Lacis and Lawford 1988; Lacis and Lawford 1991). *G. thermoglucosidasius* and *B. cereus* produced the highest lactate and formate yields, moderate ethanol and acetate yields, and low H_2 and CO_2 yields (Ouhib-Jacobs et al. 2009; Tang et al. 2009) .

Table 5.2: Summary of end-product yields, optimal growth temperatures, total molar reduction values of H_2 + ethanol (RV_{EP}), and growth conditions employed.

Organism	Growth temp (°C)	End Products (mol/mol hexose equivalent)[A]							Growth condition	Ref
		H_2	CO_2	Acetate	Ethanol	Formate	Lactate	RV_{EP}		
Ca. saccharolyticus DSM 8903	70	4.0	1.8	NR	ND	ND	ND	4.0	Cont., 1.1 g L^{-1} glucose (D=0.09 h^{-1})	(de Vrije et al. 2007)
		3.6	1.5	1.6	ND	ND	ND	3.6	Cont., 4.1 g L^{-1} glucose (D=0.1 h^{-1})	(de Vrije et al. 2007)
		3.5	NR	2.1	NR	NR	NR	3.5	Batch, 10 g L^{-1} sucrose	(Bredholt et al. 1999)
		2.5	1.4	1.4	ND	ND	0.1	2.5	Batch, 10 g L^{-1} glucose	(Kadar et al. 2004)
Ca. bescii DSM 6725	75	✓	✓	✓	NR	NR	✓	NA		(Svetlichnyi et al. 1990; Kataeva et al. 2009)
P. furiosus DSM 3638	90	3.8	1.9	1.5	0.1	NR	NR	4.0	Cont, cellobiose (D=0.45 h^{-1})	(Chou et al. 2007)[B]
		3.5	1.0	1.4	ND	NR	ND	3.5	Batch, 1.9 g L^{-1}, maltose	(Kengen et al. 1994)[B]
		2.9	1.9	0.8	0.1	NR	ND	3.1	Batch, 2 g L^{-1} maltose	(Schicho et al. 1993)[C]
		2.8	0.9	1.2	ND	NR	ND	2.8	Batch, 3.5 g L^{-1}, cellobiose	(Kengen et al. 1994)[B]
		2.6	1.4	1.0	ND	NR	NR	2.6	Cont, maltose (D=0.45 h^{-1})	(Chou et al. 2007)[B]
Th. kodakaraensis KOD1	85	3.3	1.8	1.1	NR	NR	NR	3.3	Cont, starch (D=0.2 h^{-1})	(Kanai et al. 2005)[D]

T. neapolitana DSM 4359	80-85	3.8	2.0	1.8	ND	NR	0.1	3.8	Batch, 2.5 g L^{-1} glucose	(Munro et al. 2009)
		3.2	NR	1.9	NR	NR	NR	3.2	Batch (N_2 sparged), 7.0 g L^{-1} glucose	(Nguyen et al. 2010)
		2.4	NR	1.1	NR	NR	0.7	2.4	Batch, 1.1 g L^{-1} glucose	(Eriksen et al. 2008)
		1.8	NR	1.0	NR	NR	NR	1.8	Batch, 7.5 g L^{-1} glucose	(Nguyen et al. 2008)
		1.8	NR	1.5	NR	NR	NR	1.8	Batch, 7.0 g L^{-1} glucose	(Nguyen et al. 2010)
T. petrophila RKU-1	80	3.7	0.4	1.8	NR	NR	0.3	3.7	Batch, 1 g L^{-1} glucose	(Takahata et al. 2001)
T. maritima MSB8	80	4.0	2.0	2.0	NR	ND	NR	4.0	Batch, 2 g L^{-1} glucose	(Schröder et al. 1994)
		2.2	1.1	1.0	ND	NR	0.3	2.2	Batch, 3 g L^{-1} glucose	(Lakhal et al. 2010)
		1.7	NR	1.0	NR	NR	NR	1.7	Batch, 7.5 g L^{-1} glucose	(Nguyen et al. 2008)
Cal. subterraneus subsp. *tengcongensis* MB4	75	2.8	NR	1.4	0.6	NR	ND	4.0	Cont, starch (D=0.27 h^{-1})	(Soboh et al. 2004)
		NR	NR	2.0	ND	NR	ND	NA	Cont (N_2 sparged), glucose (D=0.24 h^{-1})	(Soboh et al. 2004)
		0.3	1.5	1.0	0.7	NR	ND	1.7	Batch, 4 g L^{-1} glucose	(Xue et al. 2001)
E. harbinense YUAN-3T	35	2.8	✓	0.7	1.1	ND	ND	5.0	Batch, 20 g L^{-1} glucose	(Xing et al. 2006)
C. cellulolyticum H10	37	1.6	1.0	0.8	0.3	ND	NR	2.2	Batch, 5 g L^{-1} cellulose	(Ren et al. 2007)
		1.8	1.1	0.8	0.4	ND	NR	2,6	Batch, 5 g L^{-1} cellobiose	(Ren et al. 2007)

C. phytofermentans ISDg	35-37	Major	Major	0.6	1.4	0.1	0.3	NA	Batch, 34 g L^{-1} cellobiose	(Warnick et al. 2002)
		1.0	0.9	0.6	0.5	0.1	NR	2.0	Batch, 5 g L^{-1} cellulose	(Ren et al. 2007)
		1.6	1.2	0.6	0.6	ND	NR	2.8	Batch, 5 g L^{-1} cellobiose	(Ren et al. 2007)
C. thermocellum ATCC 27405	60	0.8	1.1	0.7	0.8	0.3	ND	2.4	Batch, 1.1 g L^{-1} cellobiose	(Rydzak et al. 2009)
		1.0	0.8	0.8	0.6	0.4	0.4	2.2	Batch, 4.5 g L^{-1} cellobiose	(Islam et al. 2006)
C. thermocellum DSM 4150	60	1.8	1.7	0.9	0.8	ND	0.1	3.4	Batch, 2 g L^{-1} glucose	(Freier et al. 1988)
		0.6	1.8	0.3	1.4	ND	0.2	3.4	Batch, 27 g L^{-1} cellobiose	(Freier et al. 1988)
Ta. Pseudethanolicus 39E	65	0.1	2.0	0.1	1.8	NR	0.1	3.7	Batch, 8 g L^{-1} glucose	(Wiegel and Ljungdahl 1981)
		NR	NR	NR	1.6	NR	<0.1	3.2	1 g L^{-1} xylose	(Lacis and Lawford 1988)
		NR	NR	0.4	1.0	NR	<0.1	2.0	Batch, 20 g L^{-1} xylose	(Lacis and Lawford 1991)
		NR	NR	0.2	0.4	NR	1.1	0.8	Batch, 20 g L^{-1} glucose	(Lacis and Lawford 1991)
G. thermoglucosidasius M10EXG[E]	60	NR	NR	0.6	0.4	1.0	0.9	0.8	Batch, 10 g L^{-1} glucose	(Tang et al. 2009)
B cereus ATCC 14579	35	NR	0.1	0.2	0.2	0.3	1.1	0.4	Batch, 3.6 g L^{-1} glucose	(Ouhib-Jacobs et al. 2009)

[A]Abbreviations: NR, not reported; ND, not detected; NA, not applicable; Major, reported as major product without absolute values; ✓, reported as present with no values indicated; Cont, continuous culture; D, dilution rate.

[B]~0.5 mol alanine per mol-hexose produced on cellobiose and maltose

[C]Produces H_2, CO_2, volatile fatty acids, and NH_3 on peptides in the absence of carbon source

[D]~0.5 mol alanine per mol-hexose produced on starch

[E]Only *G. thermoglucosidasuis* strain C56-TS93 has been sequenced but no end-product data is available. Strain M10EXG was used for end-product yield comparisons instead.

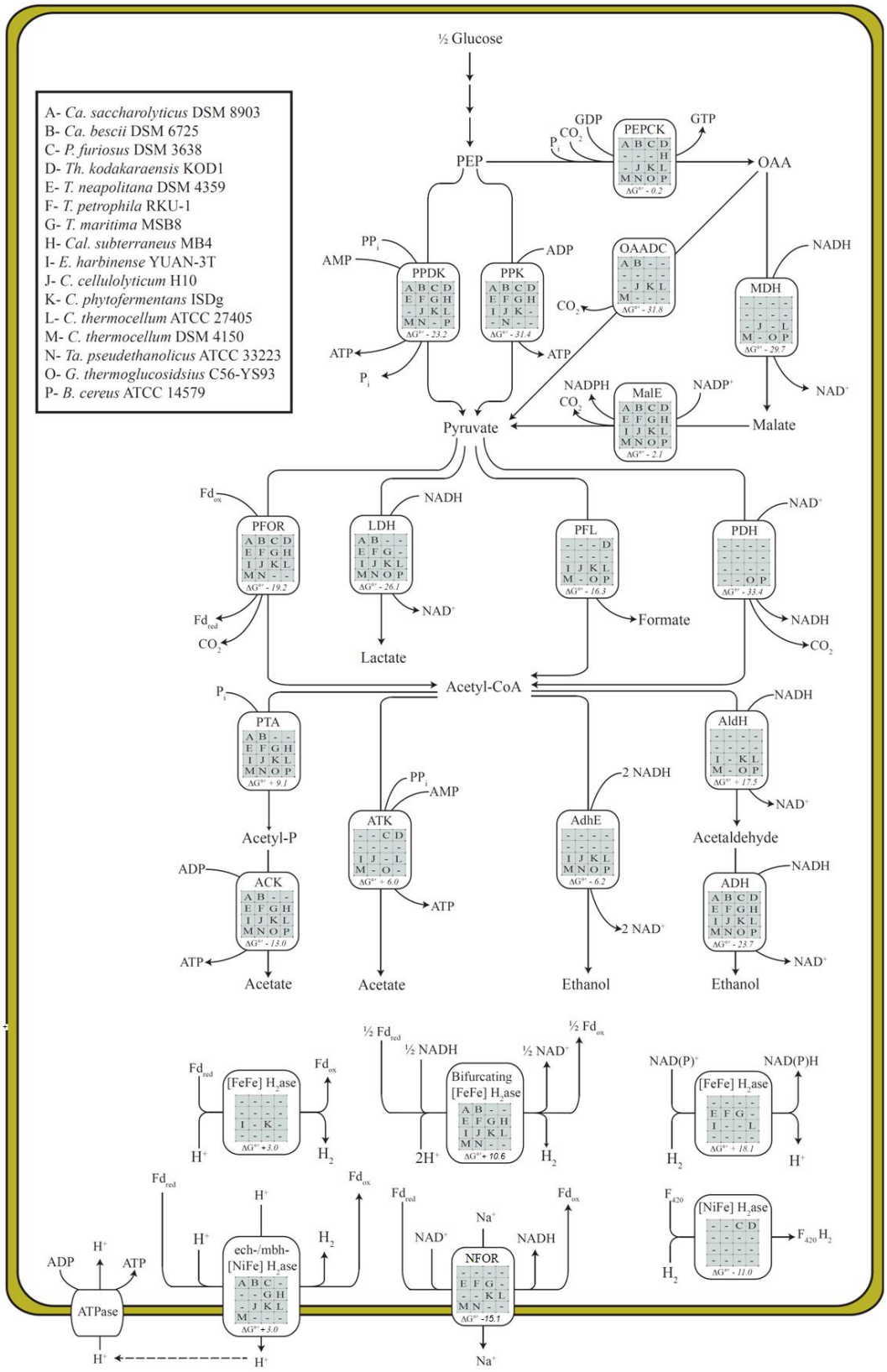
While reported yields vary considerably for each organism, it is important to note that different growth conditions may influence end-product yields through regulation of gene and gene product expression (Soboh et al. 2004; Stevenson and Weimer 2005), and modulation of metabolic flux and intracellular metabolite levels (Strobel 1995; Guedon et al. 1999a) that may act as allosteric regulators (Özkan 2004; Willquist and van Niel 2010). Variations in fermentation conditions including substrate availability/dilution rates (Strobel 1995; Guedon et al. 1999a; Guedon et al. 2000b; Desvaux et al. 2001b; Desvaux et al. 2001a; Desvaux and Petitdemange 2001; Stevenson and Weimer 2005; Islam et al. 2006), substrate composition (Ben-Bassat et al. 1981; Girbal and Soucaille 1994; Vasconcelos et al. 1994; Strobel 1995; Zhang and Lynd 2005a; Levin et al. 2006), media composition (Guedon et al. 1999b), pH (Desvaux et al. 2001c), gas partial pressures (Lamed et al. 1988; Bothun et al. 2004; Soboh et al. 2004; Nguyen et al. 2010), growth phase (Willquist et al. 2010), and accumulation of end-products (Lamed and Zeikus 1980a; Ben-Bassat et al. 1981; Freier et al. 1988; Lamed et al. 1988; Rydzak et al. 2011) have been shown to influence end-product yields. Hence, while genome content alone cannot be used to predict end-product yields with accuracy, it can reflect end-product distribution profiles.

5.4.2 Genome comparison of pyruvate metabolism and end-product synthesis pathways

The assemblage of genes encoding proteins involved in pyruvate metabolism and end-product synthesis dictate, in part, how carbon and electron flux is distributed between the catabolic, anabolic, and energy producing pathways of the cell. The flow of carbon and

electrons from PEP towards end-products may be separated into branch-points or nodes which include (i) the PEP/oxaloacetate/pyruvate node, (ii) the pyruvate/lactate/acetyl-CoA node, (iii) the acetyl-CoA/acetate/ethanol node, and the (iv) ferredoxin/NAD(P)H/H_2 node (Sauer and Eikmanns 2005). Several different enzymes may be involved in the conversion of intermediate metabolites within these nodes. These enzymes, and the presence of corresponding genes encoding these proteins in each of the organisms surveyed, are summarized in Figure 5.1. The oxidation of electron carriers (NADH and/or reduced ferredoxin) is required for maintaining glycolytic flux and leads to the ultimate production of reduced products (ethanol, lactate, and H_2). Thus, distribution of carbon and electron flux among different pathways can influence levels of reduced electron carrier pools, which in turn can dictate end-product distribution patterns. Genome content can be used to resolve the relationship between carbon and electron flux and end-product distribution.

Figure 5.1: Comparison of putative gene products involved in pyruvate metabolism and end-product synthesis among select hydrogen and ethanol-producing species. Presence of putative gene products are indicated in matrix with respective letters corresponding to selected organism (see legend). Numbers indicate standard free energies of reaction (△G°') corresponding to a particular enzyme. Abbreviations: PEPCK, phosphoenolpyruvate carboxykinase; OAADC, oxaloacetate decarboxylase; MDH, malate dehydrogenase; MalE, malic enzyme; PPK, pyruvate kinase; PPDK, pyruvate phosphate dikinase; LDH, lactate dehydrogenase; PFL, pyruvate formate lyase; PFOR, pyruvate:ferredoxin oxidoreductase; PDH, pyruvate dehydrogenase; ADH, alcohol dehydrogenase; AldH, acetaldehyde dehydrogenase; AdhE, bifunctional acetaldehyde/alcohol dehydrogenase; ACK, acetate kinase; PTA, phosphotransacetylase; NFOR, NADH:Fd oxidoreductase.


A- Ca. saccharolyticus DSM 8903
B- Ca. bescii DSM 6725
C- P. furiosus DSM 3638
D- Th. kodakaraensis KOD1
E- T. neapolitana DSM 4359
F- T. petrophila RKU-1
G- T. maritima MSB8
H- Cal. subterraneus MB4
I- E. harbinense YUAN-3T
J- C. cellulolyticum H10
K- C. phytofermentans ISDg
L- C. thermocellum ATCC 27405
M- C. thermocellum DSM 4150
N- Ta. pseudethanolicus ATCC 33223
O- G. thermoglucosidsius C56-YS93
P- B. cereus ATCC 14579
½ Glucose
PEP
PEPCK
GDP
GTP
OAA
PPDK
PPK
OAADC
MDH
NADH
NAD+
MalE
NADPH
NADP+
Malate
Pyruvate
PFOR
LDH
PFL
PDH
Formate
Lactate
Acetyl-CoA
PTA
Acetyl-P
ACK
Acetate
ATK
AldH
Acetaldehyde
AdhE
ADH
Ethanol
[FeFe] H2ase
Bifurcating [FeFe] H2ase
[NiFe] H2ase
ech-/mbh- [NiFe] H2ase
NFOR
ATPase

5.4.2.1 ***Genes involved in pyruvate synthesis***

All organisms considered in this study utilize the Embden-Meyerhof-Parnas pathway for conversion of glucose to PEP with the following notable variations. Alignments key residues of phosphofructokinase (PFK) according to Bapteste *et al.* (Bapteste et al. 2003; Bielen et al. 2010), suggest that *P. furiosus*, *Th. kodakaraensis*, *Cal. subterraneus* subsp. *tengcongensis*, *E. harbinense*, *G. thermoglucosidasius*, and *B. cereus* encode an ATP-dependent PFK, while *Thermotoga*, *Caldicellulosiruptor*, *Clostridium*, and *Thermoanaerobacter* species encode both an ATP-dependent PFK, as well as a pyrophosphate (PP_i)-dependent PFK (Bapteste et al. 2003; Bielen et al. 2010) (Supplementary Table 1). Furthermore, while bacteria catalyze the oxidation of glyceraldehyde-3-P to 3-phosphoglycerate (yielding NADH and ATP) with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK), archaea (*P. furiosus* and *Th. kodakaraensis*) preferentially catalyze the same reaction via glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPFOR). This enzyme reduces ferredoxin (Fd) rather than NAD^+ and does not produce ATP (Mukund and Adams 1995).

In contrast to the generally conserved gene content required for the production of PEP, a number of enzymes may catalyze the conversion of PEP to pyruvate (Sauer and Eikmanns 2005) (Figure 5.1; Table 5.3). PEP can be directly converted into pyruvate via an ATP-dependent pyruvate kinase (PPK), or via an AMP-dependent pyruvate phosphate dikinase (PPDK). All strains considered in this review encode both *ppk* and *ppdk*, with the exception of *C. thermocellum* strains, which do not encode a *ppk*, and *E. harbinense*,

G. thermoglucosidasius, and *B. cereus*, which do not encode *ppdk*. Given that the formation of ATP from ADP and P_i is more thermodynamically favorable than from AMP and PP_i ($\triangle G°' = 31.8$ vs. 41.7 kJ mol^{-1}), production of pyruvate via PPK is more favorable than via PPDK (Thauer et al. 1977).

Table 5.3: Genes encoding proteins involved in interconversion of phosphenolpyruvate and pyruvate.

Organism	Gene[A]						
	eno	***ppk***	***ppdk***	***pepck***	***oaadc***	***mdh***	***malE***
Standard free energy (△G°')	ND	-31.4	-23.2	-0.2	-31.8	-29.7	-2.1
Ca. saccharolyticus DSM 8903	Athe_1403	Athe_1266	Athe_1409	Athe_0393	Athe_1316-1319		Athe_1062
Ca. bescii DSM 6725	Csac_1950	Csac_1831	Csac_1955	Csac_0274	Csac_2482-2485		Csac_2059
P. furiosus DSM 3638	PF0215 PF1641	PF1188	PF0043	PF0289			PF1026
Th. kodakaraensis KOD1	TK1497 TK2106	TK0511	TK0200 TK1292	TK1405			TK1963
T. neapolitana DSM 4359	CTN_1698	CTN_0477	CTN_0413				CTN_0126
T. petrophila RKU-1	Tpet_0050	Tpet_0716	Tpet_0652				Tpet_0379
T. maritima MSB8	TM0877	TM0208	TM0272				TM0542
Cal. subterraneus subsp. *tengcongensis* MB4[A]	TTE1759	TTE1815	TTE0164 TTE0981	TTE1783			TTE2332
E. harbinense YUAN-3T	Ethha_2662	Ethha_0305					Ethha_0739

C. cellulolyticum H10	Ccel_2254	Ccel_2569	Ccel_2388	Ccel_0212	Ccel_1736-1738	Ccel_0137	Ccel_0138
C. phytofermentans ISDg	Cphy_3001	Cphy_0741 Cphy_2900	Cphy_0651	Cphy_3853	Cphy_2433-2434		Cphy_0409
C. thermocellum ATCC 27405	Cthe__0143		Cthe__1253 Cthe__1308	Cthe__2874	Cthe__0699-0701	Cthe__0345	Cthe__0344
C. thermocellum DSM 4150			Cthe_rDRAFT_1661 Cthe_rDRAFT_1896	Cthe_rDRAFT_1742	Cthe_rDRAFT_0819-0822	Yes[B]	Yes[B]
Ta. pseudethanolicus 39E	Teth39_0735	Teth39_0684	Teth39_1358 Teth39_2098	Teth39_0711			Teth39_0337
G. thermoglucosidasius C56-YS93	Geoth_0446	Geoth_0898		Geoth_0811		Geoth_0904 Geoth_3508	Geoth_1713 Geoth_2444
B.cereus ATCC 14579	BC5135	BC3323 BC4599	BC3087	BC4762		BC4592 BC2959	BC0580 NAD) BC1741 (NAD) BC4604 (NADP)

[A]Abbreviations: ND, not determined; *eno*, enolase; *ppk*, pyruvate kinase; *ppdk*, pyruvate phosphate dikinase; *pepck*, phosphoenolpyruvate carboxykinase; *oaadc*, oxaloacetate decarboxylase; *mdh*, malate dehydrogenase; *malE*, malic enzyme

[B]Genes have been verified by PCR amplification (*unpublished*)

Flux balance analysis integrated with RNAseq data suggests higher carbon and electron flow in *C. thermocellum* ATCC 27405 is directed through enzymes capable of direct, rather than indirect, conversion of PEP to pyruvate (Gowen and Fong 2010). However, *C. cellulolyticum* mutation studies suggests that a portion of PEP can also be converted to pyruvate via the "malate shunt" (Li et al. 2012b). This PPK/PPDK bypass system utilizes either (i) phosphoenolpyruvate carboxykinase (PEPCK), malate dehydrogenase (MDH), and malic enzyme (MalE), or (ii) PEPCK and oxaloacetate decarboxylase (OAADC), for the interconversion of PEP and pyruvate (Figure 5.1). While PEPCK provides a pathway for energy conservation via ATP (or GTP) production, MDH and MalE permit transhydrogenation from NADH to $NADP^+$ (Lamed and Zeikus 1980a), generating additional reducing equivalents required for biosynthesis. *G. thermoglucosidasius*, *B. cereus*, *C. thermocellum* (ATCC 27405), and *C. cellulolyticum* contain *pepck*, *mdh* and *malE* suggesting that they are capable of transhydrogenation using these proteins. Although the draft genome of *C. thermocellum* DSM 4150 does not include genes encoding MDH and MalE, we have verified their presence via PCR amplification (unpublished results). Deletion of *mdh* in *C. cellulolyticum* resulted in significant increases in lactate, and to a lesser extent ethanol yields, and reduced acetate production when grown on cellulose demonstrating carbon and electron flux through MDH in wild type strains (Li et al. 2012b). It seems evident that in the absence of MDH, transhydrogenation was reduced, and thus the resulting increase in NADH:NADPH ratios promote lactate and ethanol production, while decreasing NADPH levels for biosynthesis.

A number of organisms analyzed encode *pepck* and *oaadc* (*Ca. bescii*, *Ca. saccharolyticus*, *C. cellulolyticum*, *C. phytofermentans*, and *C. thermocellum*), also allowing for indirect conversion of PEP to pyruvate via an oxaloacetate intermediate. While the redirection of carbon and electron flux through this pathway likely has little effect on product yields, synthesis of GTP, versus ATP, may promote transcription and protein synthesis. Finally, *Cal. subterraneus*, *E. harbinense*, *P. furiosus*, *Th. kodakaraensis*, *Ta. pseudethanolicus*, and *Thermotoga* species do not encode all of the proteins required for a "malate shunt" and consequentially the catalysis of PEP to pyruvate must be achieved via PPK and/or PPDK.

5.4.2.2 ***Genes involved in pyruvate catabolism***

The pyruvate/lactate/acetyl-CoA node plays an important role in regulating carbon flux and electron distribution and dramatically affects end-product distribution. The NADH-dependent reduction of pyruvate to lactate via fructose-1,6-bisphosphate activated lactate dehydrogenase (LDH) (Özkan 2004) diverts reducing equivalents away from biofuels such as H_2 and ethanol. Alternatively, the oxidative decarboxylation of pyruvate to acetyl-CoA via pyruvate dehydrogenase (*pdh*) or pyruvate:ferreodoxin oxidoreductase (*pfor*) generate NADH and reduced Fd, respectively. These reducing equivalents may then be oxidized during the production of H_2 or ethanol (Figure 5.1). Pyruvate may also be catabolised to acetyl-CoA via pyruvate:formate lyase (*pfl*) yielding formate in the process. In some enterobacteria, formate is further oxidized to CO_2, releasing H_2, through the action of a multisubunit formate:hydrogen lyase (FHL) complex (Axley 1990). However, *pfl* was not encoded in any of the organisms analysed.

With the exception of *Cal. subterraneus* subsp. *tengcongensis*, *P. furiosus*, and *Th. kodakaraensis*, *ldh* genes were identified in all organisms studied (Table 5.4). Surprisingly, while the production of lactate from pyruvate is highly favorable thermodynamically ($\triangle G^{\circ\prime} = -26.1$ kJ mol^{-1}), only *B. cereus*, *G. thermoglucosidasius*, and, under some conditions, *Ta. pseudethanolicus* and *T. neapolitana* produce high yields of lactate (> 0.5 mol mol-glucose^{-1}). In all other organisms surveyed lactate production was either a minor end-product, not detected, or not reported under the reported growth conditions (Table 5.2). This suggests that the presence of *ldh* cannot be used to predict lactate production.

Table 5.4: Genes encoding proteins directly involved in pyruvate catabolism.

Organism	Gene[A]			
	ldh	***pdh***	***pfor***	***pfl***
Standard free energy (△G°')	-26.1	-33.4	-19.2	-16.3
Ca. saccharolyticus DSM 8903	Csac_1027		Csac_1458-1461 Csac_2248-2249	
Ca. bescii DSM 6725	Athe_1918		Athe_0874-0877 Athe_1708-1709	
P. furiosus DSM 3638			PF0965-PF0967, PF0971	
Th. kodakaraensis KOD1			TK1978, TK1982-1984	TK0289
T. neapolitana DSM 4359	CTN_0802		CTN_0680-CTN_0683	
T. petrophila RKU-1	Tpet_0930		Tpet_0905-Tpet_0908	
T. maritima MSB8	TM1867		TM0015-TM0018	
Cal. subterraneus subsp. *tengcongensis* MB4			TTE0445 TTE0960	
E. harbinense YUAN-3T	Ethha_1350 Ethha_2705		Ethha_0231-0234	Ethha_1657

C. cellulolyticum H10	Ccel_2485		Ccel_0016 Ccel_1164	Ccel_2224 Ccel_2582
C. phytofermentans ISDg	Cphy_1117 Cphy_1232		Cphy_0603 Cphy_3558	Cphy_1174 Cphy_1417 Cphy_2823
C. thermocellum ATCC 27405	Cthe__1053		Cthe__2390-2393 Cthe__2794-2797 Cthe__3120	Cthe__0505
C. thermocellum DSM 4150	Cthe_rDRAFT_2943		Cthe_rDRAFT_0414-0417 Cthe_rDRAFT_1182-1185 Cthe_rDRAFT_1311	Cthe_rDRAFT_2234
Ta. pseudethanolicus 39E	Teth39_1997		Teth39_0289 Teth39_1842	
G. thermoglucosidasius C56-YS93	Geoth_3351	Geoth_0237-0239 Geoth_1595-1597 Geoth_2366-2368 Geoth_2479-2480 Geoth_2860-2863		Geoth_3895

	BC1924		
B.cereus ATCC 14579	BC4870	BC3970-3973	BC0491
	BC4996		

[A]Abbreviations: *ldh*, lactate dehydrogenase; *pdh*, pyruvate dehydrogenase; *pfor*, pyruvate:ferredoxin oxidoreductase; *pfl*, pyruvate formate lyase

LDH is, in fact, allosterically activated by fructose-1,6-bisphosphate in *C. thermocellum* ATCC 27405, *Ca. saccharolyticus*, and *Thermoanaerobacter brockii* (Garvie 1980; Ben-Bassat et al. 1981; Özkan 2004; Willquist and van Niel 2010). While enzyme assays reveal high LDH activity in *C. thermocellum* (Rydzak et al. 2009; Rydzak et al. 2011), most studies report only trace amounts of lactate. Islam *et al*. (Islam et al. 2006), however, demonstrated that lactate production was triggered in stationary-phase batch cultures only under excess cellobiose conditions. In *Thermoanaerobacter brockii,* Ben-Bassat *et al*. reported elevated lactate production as a consequence of accumulated intracellular fructose-1,6-bisphosphate (FDP) when cultures were grown on glucose compared to starch (Ben-Bassat et al. 1981). Finally, Willquist and van Niel (Willquist and van Niel 2010) reported that LDH in *Ca. saccharolyticus* was activated by FDP and ATP, and inhibited by NAD^+ and PP_i. An increase in fructose-1,6-bisphosphate, NADH:NAD^+ ratios, and ATP:PP_i ratios was observed during the transition from exponential to stationary phase in *Ca. saccharolyticus* cultures, and was accordingly accompanied by lactate production (Willquist and van Niel 2010).

All organisms analyzed encode either *pdh* or *pfor*, but not both (Table 5.4). While *G. thermoglucosidasius* and *B. cereus* encode *pdh*, all other organisms analyzed encode *pfor*. Although *Caldicellulosiruptor*, *Clostridia*, and *Thermoanaerobacter* species studied appear to encode a putative *pdh*, there has been no enzymatic evidence to support the presence of PDH in these species. Thus far, only PFOR activity has been verified in *C. cellulolyticum* (Desvaux et al. 2001c; Desvaux et al. 2001b) and *C. thermocellum* (Rydzak et al. 2009; Rydzak et al. 2011). The putative E1, E2, and E3

subunits of the *pdh* complex (Csac_0874-0872) in *Ca. saccharolyticus* were designated simply as a keto-acid dehydrogenase by van de Werken *et al*. (van de Werken et al. 2008). Similarly, while genes encoding a putative *pdh* (Teth_0790-0793) are present in *Ta. pseudethanolicus*, genomic context strongly supports that this putative *pdh* is part of an acetoin dehydrogenase complex, despite the absence of reported acetoin production. In *Clostridia* species, putative *pdh'*s (Cthe__3449-3450, Cthe__1543) may actually encode 2-oxo acid dehydrogenase complexes, which share a common structure and homology to pyruvate dehydrogenase. These include 2-oxoglutarate dehydrogenase, branched-chain alpha-keto acid dehydrogenase, acetoin dehydrogenase complex, and the glycine cleavage complex. All organisms that encode a *pfor* also encode a Fd-dependent hydrogenase (H_2ase), bifurcating H_2ase, and/or a NADH:Fd oxidoreductase (NFOR), and are thus capable of reoxidizing reduced Fd produced by PFOR. Conversely, *G. thermoglucosidasius* and *B. cereus*, which encode *pdh* but not *pfor*, do not encode enzymes capable of reoxidizing reduced Fd, and thus do not produce H_2. While the presence of PDH allows for additional NADH production that could be used for ethanol production, *G. thermoglucosidasius* and *B. cereus* end-product profiles suggest that this NADH is preferentially rexodized through lactate production rather than ethanol production. Pyruvate decarboxylase, a homotetrameric enzyme that catalyzes the decarboxylation of pyruvate to acetaldehyde was not encoded by any of the species considered in this study.

Given the requirement of reduced electron carriers for the production of ethanol/H_2, the oxidative decarboxylation of pyruvate via PDH/PFOR is favorable over

PFL for the production of these biofuels. Genome analyses revealed that a number of organisms, including *P. furiosus*, *Ta. pseudethanolicus*, *Cal. subterraneus* subsp. *tengcongensis*, and all *Caldicellulosiruptor* and *Thermotoga* species considered, did not encode PFL. In each of these species, the production of formate has neither been detected nor reported. Unfortunately, many studies do not report formate production, despite the presence of PFL. This may be a consequence of the quantification methods used for volatile fatty acid detection. When formate is not produced, the total oxidation value of 2 CO_2 per mole glucose (+4), must be balanced with the production of H_2 and/or ethanol. Thus, the "total molar **r**eduction **v**alues of reduced end-products (H_2 + ethanol)", termed RV_{EP}, should be -4, providing that all carbon and electron flux is directed towards end-product formation and not biosynthesis. Indeed, RV_{EP}s were usually greater than 3.5 in organisms that do not encode *pfl* (*T. maritima*, *Ca. saccharolyticus*), and below 3.5 in those that do encode *pfl* (*C. phytofermentans*, *C. thermocellum*, *G. thermoglucosidasius*, and *B. cereus*; Table 5.2). In some studies, RV_{EP}s were low due to a large amount of carbon and electron flux directed towards biosynthesis. In *G. thermoglucosidasius* and *B. cereus* RV_{EP}'s of H_2 plus ethanol ranged from 0.4 to 0.8 due to higher reported formate yields. The large differences in formate yields between organisms that encode *pfl* may be due to regulation of *pfl*. In *Escherichia coli* (Membrillo-Hernandez et al. 2000; Zhu and Shimizu 2005) and *Streptococcus bovis* (Asanuma and Hino 2000; Asanuma et al. 2004), *pfl* expression has been shown to be negatively regulated by AdhE. Thus presence of *pfl* alone is not a good indicator of formate yields.

5.4.2.3 ***Genes involved in acetyl-CoA catabolism, acetate production, and ethanol production***

The acetyl-CoA/acetate/ethanol node represents the third major branch-point that dictates how carbon and electrons flow towards end-products (Figure 5.1). Acetyl-CoA may be converted to acetate, with the concomitant production of ATP, either indirectly through an acetyl phosphate intermediate using phosphotransacetylase (*pta*) and acetate kinase (*ack*), or directly via acetate thiokinase *(atk*). Although both reactions produce ATP, the former uses ADP and P_i whereas the latter uses AMP and inorganic PP_i as substrates for ATP synthesis. As a result, acetate production via *pta* and *ack* is more thermodynamically favorable than via *atk* ($\triangle G°'$ = -3.9 vs. +6.0 kJ/mol, respectively) which is typically used for acetate assimilation. Of the organisms surveyed, *E. harbinense*, *G. thermodenitrificans*, *C. cellulolyticum*, both *C. thermocellum* strains, and *G. thermoglucosidasius* contain all three genes capable of converting pyruvate to acetate (Table 5.5). Conversely, *Cal. subterraneus* subsp. *tengcongensis*, *Thermotoga* and *Caldicellulosiruptor* species, *C. phytofermentans*, *Ta. pseudethanolicus*, and *B. cereus* encode only *pta* and *ack*, whereas *P. furiosus* and *Th. kodakaraensis* encode only *atk*.

Table 5.5: Genes encoding proteins involved in end-product synthesis from acetyl-CoA

Organism	Gene[A]					
	pta	***ack***	***atk***	***aldH***	***adh***	***adhE***
Standard free energy (△G°')	9.1	-13.0	6.0	17.5	-23.7	-6.2
Ca. saccharolyticus DSM 8903	Csac_2041	Csac_2040			Csac_0407 Csac_0554 Csac_0622 Csac_0711 Csac_1500	
Ca. bescii DSM 6725	Athe_1494	Athe_1493			Athe_0928 Athe_0224	
P. furiosus DSM 3638			PF1540 PF1787		PF0075 PF0608	
Th. kodakaraensis KOD1			TK0465 TK0665		TK1008 TK1569	

T. neapolitana DSM 4359	CTN_0945	CTN_0411	CTN_0257
			CTN_0369
			CTN_0385
			CTN_0580
	CTN_1440		CTN_1655
			CTN_1756
T. petrophila RKU-1	Tpet_1042	Tpet_0650	Tpet_0007
			Tpet_0107
			Tpet_0484
			Tpet_0508
			Tpet_0563
	Tpet_1615		Tpet_0614
			Tpet_0813

T. maritima MSB8	TM1130 TM1755	TM0274			TM0111 TM0298 TM0412 TM0436 TM0820 TM0920	
Cal. subterraneus subsp. *tengcongensis* MB4	TTE1482	TTE1481			TTE0313 TTE0695 TTE0696 TTE1591	
E. harbinense YUAN-3T	Ethha_2711	Ethha_2004	Ethha_1333	Ethha_0578 Ethha_0625	Ethha_0051 Ethha_0580 Ethha_1164 Ethha_2217 Ethha_2239	Ethha_1385

C. cellulolyticum H10	Ccel_2137	Ccel_2136	Ccel_0494 Ccel_1469		Ccel_0894 Ccel_1083 Ccel_3337	Ccel_3198
C. phytofermentans ISDg	Cphy_1326	Cphy_1327		Cphy_0958 Cphy_1178 Cphy_1416 Cphy_1428 Cphy_2418 Cphy_2642 Cphy_3041	Cphy_1029 Cphy_1421 Cphy_2463 Cphy_2463	Cphy_3925
C. thermocellum ATCC 27405	Cthe_1029	Cthe_1028	Cthe_0551	Cthe_2238	Cthe_0101 Cthe_0394 Cthe_2579	Cthe_0423
C. thermocellum DSM 4150[B]	Cthe_rD_2741	Cthe_rD_2742	Cthe_rD_2349	Cthe_rD_1042	Cthe_rD_0189 Cthe_rD_0616 Cthe_rD_2833	Cthe_rD_1096
Ta. pseudethanolicus 39E	Teth39_1296	Teth39_1295			Teth39_0220 Teth39_1597	Teth39_0206

					Teth39_1979	
G. thermoglucosidasius C56-YS93	Cthe__3862	Geoth_0875	Geoth_0855 Geoth_0879 Geoth_2349	Geoth_0268 Geoth_0652 Geoth_3494	Geoth_1572 Geoth_1941 Geoth_0631	Geoth_3879
B. cereus ATCC 14579	BC5387	BC4637		BC2832 BC3555 BC1285	BC0802 BC2529 BC2220	BC4365

[A]Abbreviations: *pta*, phosphotransacetylase; *ack*, acetate kinase; *atk*, acetate thiokinase; *aldH*, acetaldehyde dehydrogenase; *adh*, alcohol dehydrogenase; *adhE*; bifunctional acetylaldehyde/alcohol dehydrogenase

[B]Note that Cthe_rDRAFT gene annotation was changed to Cthe_rD in this table due to space limitations.

Alternatively, acetyl-CoA may be converted into ethanol, during which 2 NADH (or NADPH) are oxidized, either directly via a fused acetaldehyde/alcohol dehydrogenase encoded by *adhE*, which has been proposed to be the key enzyme responsible for ethanol production (Brown et al. 2011; Trinh et al. 2011), or indirectly through an acetaldehyde intermediate via acetaldehyde dehydrogenase (*aldH*) and alcohol dehydrogenase (*adh*). While all organisms surveyed encoded multiple class IV Fe-containing ADHs (Table 5.5), the functions of these ADHs may vary with respect to substrate specificity (aldehyde length and substitution), coenzyme specificity (NADH vs. NADPH), and the catalytic directionality favored (ethanol formation vs. consumption) (Lovitt et al. 1988; Burdette and Zeikus 1994; Desvaux et al. 2001c; Desvaux and Petitdemange 2001; Liu et al. 2009; Rydzak et al. 2009; Pei et al. 2010; Willquist et al. 2010; Rydzak et al. 2011). Although there are reports of *in silico* determinations of substrate and cofactor specificity amongst ADHs, such resolutions are problematic with divergent species (Nair et al. 1994; Bernard et al. 1995). Often times, the gene neighborhoods of identified ADHs were suggestive that the physiological role of many enzymes was not ethanol production. This is evident in *Ca. saccharolyticus*, which does not produce ethanol despite reported NADPH-dependent ADH activity (Willquist et al. 2010).

P. furiosus, *Th. kodakaraensis*, and all *Thermotoga* and *Caldicellulosiruptor* species do not encode *adhE* or *aldH*, and therefore produce negligible or no ethanol. Given the absence of ethanol producing pathways in these species, reducing equivalents are disposed of through H_2 production via H_2ases and/or lactate production via LDH. Surprisingly, while *Cal. subterraneus* subsp. *tengcongensis* also does not appear to

encode *aldH* or *adhE*, NADPH-dependent AldH and both NADH and NADPH-dependent ADH activities, as well as ethanol production, have been reported by Soboh *et al*. (Soboh et al. 2004). Similarly, *Caldicellulosiruptor obsidiansis*, which does not encode *aldH* or *adhE*, does produce trace levels of ethanol, suggesting that the various encoded ADHs may have broad substrate specificities (Hamilton-Brehm et al. 2010). Although *C. cellulolyticum* and *Ta. pseudethanolicus* do not encode *aldH*, they do encode *adhE*, and thus are capable of ethanol production. Of the organisms surveyed, only *G. thermoglucosidasius* and *C. cellulolyticum* encoded *aldH* and *adh* but no *adhE*, and produced moderate amounts of ethanol (~0.4 mol per mol hexose). Conversely, a number of organisms (*E. harbinense*, *C. phytofermentans*, both *C. thermocellum* strains, *G. thermoglucosidasius*, and *B. cereus*) encoded *aldH*, *adh*, and *adhE*, all of which produce varying ethanol yields.

5.4.2.4 ***Hydrogenases***

In addition to disposal of reducing equivalents via alcohol and organic acid production, electrons generated during conversion of glucose to acetyl-CoA can be used to produce molecular H_2 via a suite of [FeFe] and/or [NiFe] H_2ases. The incredible diversity of H_2ases has been extensively reviewed by Vignais *et al*. and Calusinska *et al*. (Vignais 2001; Vignais 2008; Calusinska et al. 2010). H_2ases may be (i) monomeric or multimeric, (ii) can catalyze the reversible production of H_2 using various electron donors, including reduced Fd and NAD(P)H, or (iii) can act as sensory H_2ases capable of regulating gene expression (Buhrke et al. 2004). While most H_2ases can reversibly shuttle electrons between electron carriers and H_2, they are typically committed to either

H_2-uptake or evolution, depending on reaction thermodynamics and the requirements of the cell *in vivo* (Vignais 2001). While Fd-dependent H_2 production remains thermodynamically favorable at physiological concentrations ($\triangle G°'$ ~ -3.0 kJ mol^{-1}), potential production of H_2 from NAD(P)H ($\triangle G°'$ = +18.1 kJ mol^{-1}) becomes increasingly unfavorable with increasing H_2 partial pressure (Angenent et al. 2004). Hence, Fd-dependent H_2ases are associated with H_2 evolution, whereas NAD(P)H-dependent H_2ases are more likely to catalyze H_2 uptake. Recent characterization of a heterotrimeric "bifurcating" H_2ase from *Thermotoga maritma* demonstrated that it can simultaneously oxidize reduced Fd and NADH to H_2 ($\triangle G°'$ ~ +10.6 kJ mol-H_2^{-1}), which drives the endergonic production of H_2 from NADH ($\triangle G°'$ +18.1 kJ mol^{-1}) by coupling it to the more favorable oxidation of reduced Fd ($\triangle G°'$ ~ -3.0 kJ mol^{-1}) (Schut and Adams 2009).

With the exception of *G. thermoglucosidasius* and *B. cereus*, which did not contain putative H_2ase genes, the genomes of all of the organisms surveyed encode multiple H_2ases. These H_2ases were classified based on i) the phylogenetic relationship of H_2ase large subunits (Supplementary Figures 1 and 2), according to Calusinska *et al*. (Calusinska et al. 2010), ii) H_2ase modular structure, and iii) subunit composition, based on gene neighbourhoods. Encoded [NiFe] H_2ases fell into 3 major subgroups including: (i) Fd-dependent, H_2-evolving, membrane-bound H_2ases (Mbh) and/or energy conserving [NiFe] H_2ases (Ech) capable of generating sodium/proton motive force (Group 4) (Soboh et al. 2004), (ii) Soluble cofactor-dependent (F_{420} or NAD(P)H), bidirectional, cytoplasmic, heteromultimeric H_2ases (Group 3), and (iii) H_2-uptake, membrane bound

H_2ases (Group 1) (Vignais 2008) (Supplementary Figure 1). Similarly, encoded [FeFe] H_2ases fell into 5 major subgroups including: (i) heterotrimeric bifurcating H_2ases, (ii) dimeric, NAD(P)H-dependent uptake H_2ases, (iii) monomeric, putatively Fd-dependent H_2ases, (iv) dimeric sensory H_2ases containing PAS/PAC sensory domains which may be involved in redox sensing, and (v) monomeric sensory H_2ases (Supplementary Figure 2). These sensory H_2ases are usually encoded upstream of trimeric bifurcating H_2ases (Table 5.6) and are often separated by a histidine/serine kinase suggesting a regulatory relationship between these two enzymes (Calusinska et al. 2010).

Table 5.6: Genes encoding putative hydrogenases, sensory hydrogenases, and NADH:Fd oxidoreductases using ferredoxin, coenzyme F420, and NAD(P)H as electron carriers.

Organism	Hydrogenase and NADH:Fd Oxidoreductase Classification and Corresponding Genes[A]						
	[NiFe] H_2ase		[FeFe] H_2ase				NFOR
	Fd-dependent *ech* and *mbh* [G4]	F_{420}-dependent [G3] and other [G1]	Bifurcating	Sensory[B]	NAD(P)H-dependent	Fd-dependent	*rnf*-type
Standard free energy (△G°')*	-3.0	11	10.6**	NA	18.1	-3.0	-15.1***
Ca. bescii DSM 6725	Athe_1082-Athe_1087		Athe_1297-Athe_1299 [A1 TR(M3)]	Athe_1292 [D M2e]			
Ca. saccharolyticus DSM 8903	Csac_1534-Csac_1539		Csac_1862-Csac_1864 [A1 TR(M3)]	Csac_1857 [D M2e]			
P. furiosus DSM 3638	PF1423-PF1436	PF0891-PF0894 [G3] PF1329-PF1332 [G3]					

Th. kodakaraensis KOD1	TK2080-TK2093	TK2069-TK2072 [G3]					
T. neapolitana DSM 4359			CTN_1067-CTN1069 [TTH]	CTN_1071-CTN_1072 [CD(M2f)]	CTN_0485 [TTH]		CTN_0437-CTN_0442
T. petrophila RKU-1			Tpet_1367-Tpet_1369 [TTH]	Tpet_1371-Tpet_1372 [CD(M2f)]	Tpet_0723 [TTH]		Tpet_0675-Tpet_0680
T. maritima MSB8			TM1424-TM1426 [TTH]	TM1420-TM1422 [CD(M2f)]	TM0201 [TTH]		TM0244-TM0249
Cal.subterraneus subsp. *tengcongensis* MB4	TTE0123-TTE0134		TTE0892-TTE0894 [A1 TR(M3)]	TTE0887 [D M2e] TTE0697 [CD(M2f)]			
E. harbinense YUAN-3T			Ethha_2614-Ethha_2616 [A8 TR(M3)]	Ethha_0052 [CD(M2f)]	Ethha_2293 [A7 D(M3)]	Ethha_0031 [B2 M2a]	
C. cellulolyticum H10	Ccel_1686-Ccel_1691 Ccel_3363-Ccel_3371	Ccel_1070-Ccel_1071 [G1]	Ccel_2303-Ccel_2305 [A8 TR(M3)] Ccel_2232-Ccel_2234 [A1 TR(M3)]	Ccel_2300-Ccel_2301 [CD(M2f)]		Ethha_2695 [B3 M3a]	

		Ccel_2467- Ccel_2468 A1 TR(M3)				
C. phytofermentans ISDg	Cphy_1730- Cphy_1735	Cphy_0087- Cphy_0089 A8 TR(M3) Cphy_3803- Cphy_3805 A1 TR(M3)	Cphy_0092- Cphy_0093 **CD(M2f)** Cphy_3798 **D M2e**		Cphy_2056 A5 M2c Cphy_0090 B1 M3a	Cphy_0211- Cphy_0216
C. thermocellum ATCC 27405	Cthe_3013- Cthe_3024	Cthe_0428- Cthe_0430 A8 TR(M3) Cthe_0340- Cthe_0342 A1 TR(M3)	Cthe_0425- Cthe_0426 **CD(M2f)** Cthe_0335 **D M2e**	Cthe_3003- Cthe_3004		Cthe_2430- Cthe_2435
C. thermocellum DSM 4150[D]	Cthe_rD_2162- Cthe_rD_2173	Cthe_rD_1101- Cthe_rD_1103 A8 TR(M3) Cthe_rD_2978 A1 TR(M3)	Cthe_rD_1098- Cthe_rD_1099 **CD(M2f)**	Yes[C]		Cthe_rD_0369- Cthe_rD_0375
Ta. pseudethanolicus 39E		Teth39_1456- Teth39_1458 A1 TR(M3)	Teth39_0221 **CD(M2f)** Teth39_1463 **D M2e**			Teth39_2119- Teth39_2124

G. thermoglucosidasius C56-YS93	
B. cereus ATCC 14579	

[A]Characterization of hydrogenase specificity was based metallocenter composition ([NiFe] or [FeFe]), modular structure, subunit composition, and large (catalytic) subunit phylogeny according to Vignais *et al*. and Calusinska *et al*. (Vignais 2001; Vignais 2008; Calusinska et al. 2010). Phylogenetic cluster groupings are indicated in superscript, and corresponding phylogenetic trees are provided in Supplementary Figures 1 and 2. Abbreviations: H_2ase, hydrogenase; NFOR, NADH:ferredoxin oxidoreductase; *ech*, energy conserving hydrogenase; *mbh*, membrane bound hydrogenase; *rnf*, *Rhodobacter* nitrogen fixation.

[B]Group D M2e hydrogenases are poorly characterized and do not contain a PAS/PAC-sensory domain. However, given their proximity to protein kinases and bifurcating hydrogenases, and their phylogenetic proximity to group C D(M2f) sensory hydrogenases (Additional file 3) we have classified them as sensory hydrogenases.

[C]Verified by microarray and proteomic analysis (*unpublished*).

[D]Note that Cthe_rDRAFT gene annotation was changed to Cthe_rD in this table due to space limitations.

With the exception of *P. furiosus* and *Th. kodakaranesis*, which encode only Fd-dependent and putative F_{420}-dependent [NiFe] H_2ases, all other H_2ase encoding organisms surveyed are capable of H_2ase-mediated oxidation/reduction of both Fd and NAD(P)H. This seems fitting given that *P. furiosus* and *Th. kodakaraensis* preferentially catalyze the oxidation of glyceraldedhyde-3-P via GAPFOR rather than GAPDH and PGK, and thus must reoxidize reduced Fd, rather than NADH, during fermentative product synthesis. All other H_2ase encoding organisms produce NADH during glycolysis and reduced Fd via PFOR. In these organisms, the oxidation of these electron carriers may be carried out using various different types of H_2ases. All of these species encoded at least a single putative bifurcating H_2ase (Table 5.6). The majority of these bifurcating H_2ases were found downstream dimeric or monomeric sensory [FeFe] H_2ases that may be involved in their regulation (Table 5.6). Soboh *et al.* have demonstrated that NADH-dependent H_2ase activities in *Cal. subterraneus* subsp. *tengcongensis* are affected by H_2 partial pressures (Soboh et al. 2004) suggesting possible regulation of these H_2ases via a two-component signal transduction mechanism in response changes in redox levels (Buhrke et al. 2004; Calusinska et al. 2010). It is important to note that these NADH-dependent H_2ase activities may reflect bifurcating H_2ase activities given that *Cal. subterraneus* subsp. *tengcongensis* encodes only a Fd-dependent and a putative bifurcating H_2ase, and no NAD(P)H-dependent H_2ases.

While *Ta. pseudethanolicus* only encodes a bifurcating H_2ase, all other organisms that encode a bifurcating H_2ase also encode Fd-dependent H_2ases. Putative Fd-dependent, [NiFe] Ech/Mbh-type H_2ases were identified in the genomes of *Cal.*

subterraneus subsp. *tengcongensis*, *P. furiosus*, *Th. kodakaraensis*, and all *Caldicellulosiruptor* and *Clostridium* species (Table 5.6). A pair of putative Fd-dependent [FeFe] H_2ases were identified in both *E. harbinense* and *C. phytofermentans*. With the exception *Ta. pseudethanolicus*, *Cal. subterraneus* subsp. *tengcongensis*, and *Caldicellulosiruptor* species, all organisms surveyed containing a bifurcating H_2ase also appear to be capable of NADH and/or NADPH oxidation using NADH/NADPH-dependent H_2ases. As with ADHs, however, we could not determine H_2ase cofactor specificity exclusively using *in silico* sequence analysis, stressing the importance of activity characterization of enzyme substrate specificity. While *C. cellulolyticum* achieves NAD(P)H oxidation using a putative H_2-uptake [NiFe] H_2ases, *E. harbinense*, *Thermotoga* species, and *C. thermocellum* ATCC 27405 achieve this using [FeFe] H_2ases. Although the draft genome of *C. thermocellum* DSM 4150 does not encode an NAD(P)H-dependent H_2ase, our proteomic and microarray data reveal the presence of Cthe__3003/Cthe__3004 homologues (Rydzak, *unpublished results*).

In addition to H_2ases-mediated electron transfer between Fd and/or NADH and H_2, electrons may be transferred directly between Fd and NAD(P)H via an Rnf-like (Rhodobacter nitrogen fixation) NADH:ferredoxin oxidoreductase (NFOR), a membrane-bound enzyme complex capable of generating a sodium motive force derived from the energy difference between reduced Fd and NADH. Only *Thermotoga* species, *C. phytofermentans*, *C. thermocellum*, and *Ta. pseudethanolicus* encode putatively identified NFOR. Proteomic analysis of *C. thermocellum*, however, revealed low, or no, expression

of NFOR subunits, suggesting it does not play a major factor in electron exchange between Fd and NADH (Rydzak et al. 2012).

While the presence/absence of genes encoding pathways that lead to reduced fermentation products (i.e. formate, lactate, and particularly ethanol) is a major determinant of H_2 yields, we can make some inferences with respect to H_2 yields based on the types of H_2ases encoded. Given the thermodynamic efficiencies of H_2 production using different cofactors, we can say that Fd-dependent H_2ases are conducive for H_2 production while NAD(P)H-dependent H_2ases are not. However, organisms that do not encode ethanol-producing pathways (i.e. *Caldicellulosiruptor* and *Thermotoga* species) may generate high intracellular NADH:NAD^+ ratios, making NADH-dependent H_2 production thermodynamically feasible under physiological conditions. Conversely, in organisms capable of producing both H_2 and ethanol (*Ethanoligenens*, *Clostridium*, and *Thermoanaerobacter* species), the presence of Fd-dependent H_2ases appears to be beneficial for H_2 production. For example, *E. harbinense* and *Clostridium* species, which encode Fd-dependent, as well as bifurcating and NAD(P)H-dependent H_2ases, produce much higher H_2 yields when compared to those of *Ta. pseudethanolicus*, which encodes only one bifurcating H_2ase and no Fd or NAD(P)H-dependent H_2ases. Interestingly, organisms that do not encode H_2ases (*G. thermoglucosidasius* and *B. cereus*) produce low ethanol and high lactate (and/or formate yields), suggesting that H_2 production can help lower NADH:NAD^+ ratios, and thus reduce flux through LDH.

5.4.3 Influence of overall genome content on end-product profiles

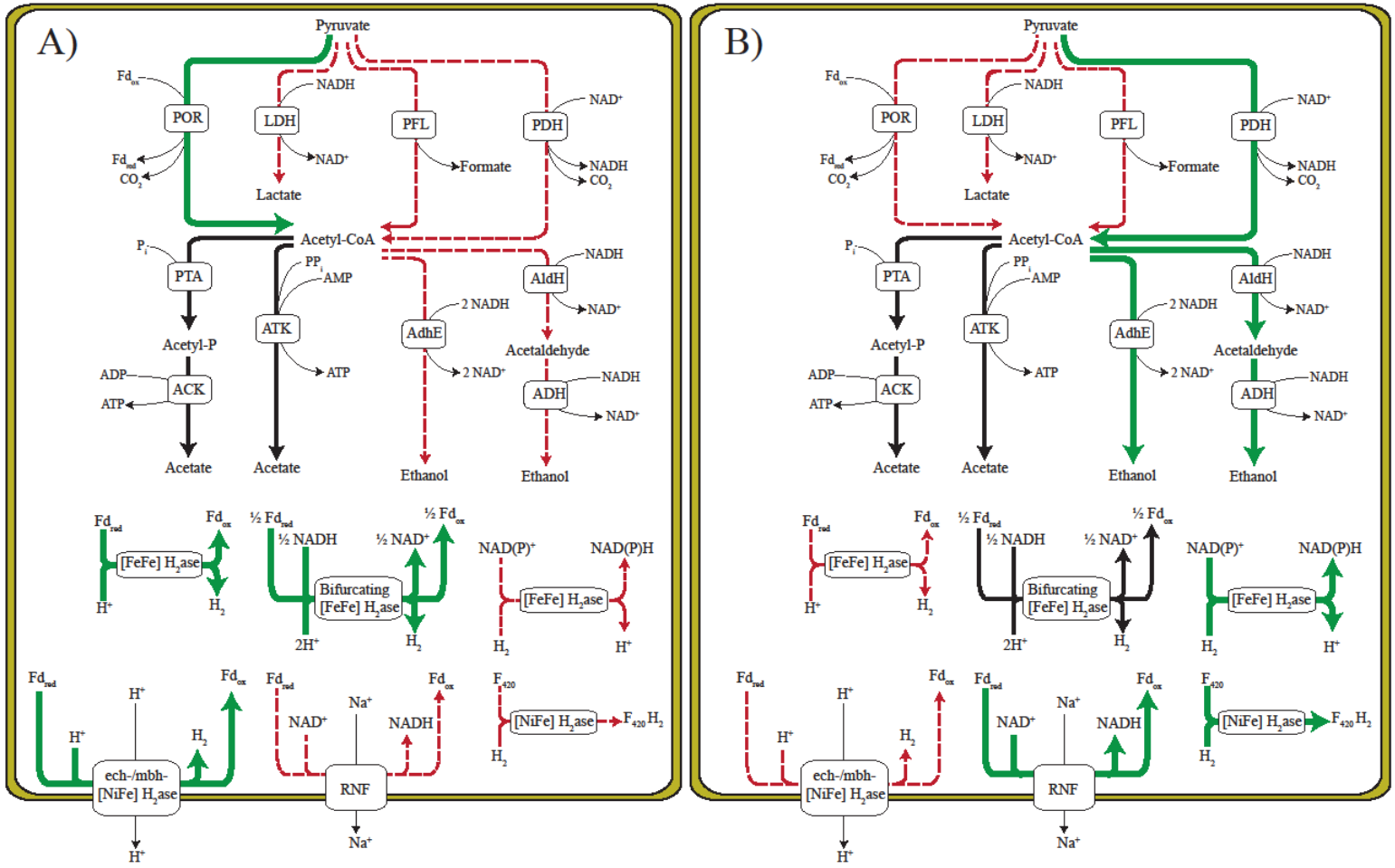
The presence and absence of genes encoding proteins involved in pyruvate metabolism and end-product synthesis may be used as an indicator of end-product distribution. By comparing genome content to end-product yields, we identified key markers that influence ethanol and H_2 yields. These include (i) MDH (ii) LDH, (iii) PFL vs. PFOR and/or PDH (iv) Aldh and AdhE, and (v) bifurcating, Fd-dependent, and NAD(P)H dependent H_2ase.

While it is difficult to elucidate how differences in "malate shunt" genes affect end-product synthesis patterns by comparing reported yields, eliminating MDH has been shown to increase lactate and ethanol production, and decrease acetate production in *C. cellulolyticum* (Li et al. 2012b). The elimination of this transhydrogenation pathway may increase NADH:NAD^+ ratios for reduced end-product synthesis and reduce NADPH:$NADP^+$ ratios for biosynthesis. While presence of LDH is not a good predictor of lactate yields, LDH, when activated, diverts reducing equivalents away from H_2 and ethanol. In contrast to PFL, PFOR and PDH produce additional reducing equivalents (reduced Fd and NADH, respectively), and thus promote reduced end-product synthesis. Organisms that do not encode *pfl* generally produce more ethanol and H_2 (based on sum redox value) compared to those that do encode *pfl*. Of the organisms surveyed, those that did not encode (or express) both *adhE* and *aldH* produced near-maximal H_2 yields and little to no ethanol. While the type(s) of encoded H_2ases appear to have little impact in organisms that do not encode ethanol producing pathways, they do seem to influence reduced end-product yields in those that do. For example, *Ta. pseudethanolicus*, which

encodes an *adhE*, NFOR, and a single bifurcating H_2ase, but no discernible Fd or NAD(P)H-dependent H_2ases, generates low H_2 and near-optimal ethanol yields. The inability to oxidize reduced Fd via Fd-dependent H_2ases may elevate reduced Fd levels, which in turn can be used by NFOR to produce additional NADH for ethanol synthesis. Interestingly, in the absence of H_2ases, lactate production was favoured over ethanol production, suggesting that H_2 production can help lower NADH:NAD^+ ratios, and thus reduce flux through LDH.

Given the impact that MDH, PFL, Aldh, AdhE, and the different H_2ases have on end-product yields, screening for these biomarkers can streamline ethanol and H_2 producing potential of sequenced and novel organisms through *in silico* gene mining and the use of universal primers, respectively. Furthermore, understanding how end-product yields are affected by (i) the framework of genes encoding pathways catalyzing pyruvate into end-products, and (ii) thermodynamic efficiencies of these reactions, we can begin to develop informed metabolic engineering strategies for optimization of either ethanol or H_2 (Figure 5.2). For example, in order to optimize either ethanol or H_2, we would recommend elimination of *ldh* and *pfl* in order to allow accumulation of additional reducing equivalents. Given that ethanol and H_2 compete for reducing equivalents, elimination of one product should direct carbon/and or electron flux towards the other.

Figure 5.2: Differentiation between fermentation pathways that favor (a) hydrogen and (b) ethanol production based on comparative genomics and end-product profiles. Pathways that favor (green lines), disfavor (broken red lines), and appear to have little impact (black lines) on production of H_2 or ethanol are indicated. Correlation of reaction thermodynamics and genome content with reported end-product yields suggest that reduction, and subsequent reoxidation, of ferredoxin via PFOR and Fd-dependent (and/or bifurcating) H_2ases, respectively, support H_2 production. Alternatively, reduction, of NAD^+ via PDH (and/or NADH generating uptake H_2ases) generate NADH conducive for ethanol production. Abbreviations (see figure 5.1 legend).


A)
B)
Pyruvate
POR
LDH
PFL
PDH
Lactate
Formate
Acetyl-CoA
PTA
ATK
AdhE
AldH
Acetyl-P
Acetaldehyde
ACK
ADH
Acetate
Ethanol
[FeFe] H2ase
Bifurcating [FeFe] H2ase
[NiFe] H2ase
ech-/mbh- [NiFe] H2ase
RNF

For optimization of H_2 yields (Figure 5.2a), deletion of *aldH* and *adhE* is likely most effective. Although conversion of pyruvate to acetyl-CoA is more thermodynamically favorable using PDH versus PFOR ($\triangle G°'$ = -33.4 vs. -19.2 kJ mol^{-1}), production of H_2 from NADH is highly unfavorable compared to the use of reduced Fd ($\triangle G°'$ = +18.1 vs. -3.0 kJ mol^{-1}). This in turn demonstrates that reduction of Fd via PFOR and subsequent H_2 production via a Fd-dependent H_2ase ($\triangle G°'$ = -21.2 kJ mol^{-1}) is more favorable than NADH production via PDH and subsequent H_2 production via NAD(P)H-dependent H_2ases ($\triangle G°'$ = -15.3 kJ mol^{-1}). Therefore, we propose that conversion of pyruvate to acetyl-CoA via PFOR is favorable for H_2 production, and *pdh* (and *pfl*) should be deleted. Given that 2 NADH (per glucose) are produced during glycolysis in most anaerobic microorganisms, the presence of a bifurcating H_2ase, which would simultaneously oxidize the 2 NADH generated during and 2 reduced Fd produced by PFOR, would be required to achieve theoretically maximal H_2 yields of 4 mol per mol glucose. A Fd-dependent H_2ase would also be conducive for H_2 production during times when reducing equivalents generated during glycolysis are redirected towards biosynthetic pathways, resulting in a disproportionate ratio of reduced ferredoxin to NAD(P)H. Alternatively, in organisms such as *P. furiosus* and *Th. kodakaraensis*, which generate high levels of reduced Fd and low levels of NADH, the presence of Fd-dependent H_2ases, rather than bifurcating H_2ases, would be more conducive for H_2 production. In all cases, NFOR and NAD(P)H-dependent H_2ases should be deleted to prevent oxidation of reduced Fd and uptake of H_2, respectively, which would generate NAD(P)H.

The metabolic engineering strategies employed for optimization of ethanol (Figure 5.2b) are much different than those used for the production of H_2. First, *adhE* and/or *aldH* and *adh* genes that encode enzymes with high catalytic efficiencies in the direction of ethanol formation should be heterologously expressed. Given that ethanol production is NAD(P)H dependent, increasing NADH production should be optimized, while Fd reduction should be eliminated. Through deletion of *pfl* and *pfor*, and expression of *pdh*, up to 4 NADH can be generated per glucose, allowing for the theoretical maximum of 2 mol ethanol per mol glucose to be produced. To prevent NADH reoxidation, lactate and H_2 production should be eliminated by deleting *ldh* and NAD(P)H-dependent H_2ases. While this strategy is theoretically sound, low AldH/ADH catalytic efficiencies may cause NADH/NAD^+ ratios to rise so high that they may impede glycolysis. In these situations, the presence of a NFOR or NAD(P)H-dependent H_2ase may intermittently alleviate these high NADH/NAD^+ ratios through generation of reduced Fd pools or H_2 production, respectively, albeit it would decrease reducing equivalents for ethanol production.

While some attempts to increase H_2 and/or ethanol yields through genetic engineering have been successful in a number of lignocellulolytic organisms (reviewed elsewhere; (Chang and Yao 2011)) engineering of strains discussed here has only been marginally successful. Heterologous expression of *Zymomonas mobilis* pyruvate decarboxylase and ADH in *C. cellulolyticum* increased cellulose consumption and biomass production, and decreased lactate production and pyruvate overflow due to a more efficient regulation of carbon and electron flow at the pyruvate branchpoint

(Guedon et al. 2002). However, despite higher levels of total ethanol produced, ethanol yields (per mol hexose consumed) actually decreased when compared to the wild-type strain. Similarly, deletion of PTA in *C. thermocellum* drastically reduced acetate production, but had minimal impact on lactate or ethanol production (Tripathi et al. 2010). This suggests that genome content alone cannot exclusively dictate the extent of end-product yields observed in literature, and thus growth conditions must be optimized in order to moderate regulatory mechanisms that direct carbon and electron flux. This could only be attained through a thorough understanding of regulatory mechanisms that mediate gene and gene-product expression and activity levels under various growth conditions through a combination of genomics, transcriptomics, proteomics, metabolomics, and enzyme characterization.

5.5 Conclusions

Fermentative bacteria offer the potential to convert biomass into renewable biofuels such as H_2 and ethanol through consolidated bioprocessing. However, these bacteria display highly variable, branched catabolic pathways that divert carbon and electrons towards unwanted end products (i.e. lactate, formate). In order to make fermentative H_2 and/or ethanol production more economically feasible, biofuel production yields must be increased in lignocellulolytic bacteria capable of consolidated bioprocessing. While the cellulolytic and, to a lesser extent, H_2 and ethanol producing capabilities of cellulolytic bacteria have been reviewed (Desvaux 2006; Ren et al. 2007; Bruggemann and Gottschalk 2008), a comprehensive comparison between genome content and corresponding end-product distribution patterns has not been reported. While reported

end-product yields among vary considerably in response to growth conditions, which may influence gene and gene product expression and metabolic flux, we demonstrate that composition of genes encoding pyruvate catabolism and end-product synthesis pathways alone can be used to approximate potential end-product distribution patterns. We have identified a number of genetic biomarkers, including (i) MDH (ii) LDH, (iii) PFL vs. PFOR and/or PDH (iv) Aldh and AdhE, and (V) bifurcating, Fd-dependent, and NAD(P)H dependent H_2ases, that can be used for streamlining H_2 and/or ethanol producing capabilities in sequenced and novel isolates. By linking genome content, reaction thermodynamics, and end-product yields, we offer potential targets for optimization of either ethanol or H_2 yields via metabolic engineering. Deletion of LDH and PFL could potentially increase both H_2 and ethanol yields. While deletion of ethanol producing pathways (*aldH*, *adh*, *adhE*), increasing flux through PFOR, overexpression of Fd-dependent H_2ases, and elimination of potential H_2-uptake (NAD(P)H-dependent) H_2ases could lead to increased H_2 production, eliminating H_2 production and redirecting flux through PDH would be beneficial for ethanol production. Although gene and gene-product expression, functional characterization, and metabolomic flux analysis remains critical in determining pathway utilization, insights regarding how genome content affects end-product yields can be used to direct metabolic engineering strategies and streamline the characterization of novel species with potential industrial applications.

5.6 Authors` contributions

Thomas Rydzak and Carlo R. Carere co-authored the manuscript. Thomas Rydzak and Carlo R. Carere performed genomic meta-analysis with help from Tobin Verbeke.

Thomas Rydzak performed end-product comparisons and thermodynamic calculations. Carlo R. Carere performed phylogenetic analysis. Richard Sparling, Nazim Cicek, and David B. Levin conceived of the study, participated in its design, and helped draft the manuscript.

5.7 Acknowledgements

This work was supported by funds provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944-04), by Genome Canada, through the Applied Genomics Research in Bioproducts or Crops (ABC) program for the grant titled, "Microbial Genomics for Biofuels and CoProducts from Biorefining Processes", and by the Province of Manitoba, Agricultural and Rural Development Initiative (ARDI), grant 09-986.

Chapter 6: Relative Protein Expression Profiles and Growth Phase-Dependent Changes in Core Metabolic Proteins in *Clostridium thermocellum* DSM 1237 using 2D-HPLC-MS/MS[20]

6.1 Abstract

Clostridium thermocellum produces H_2, ethanol, and concomitant volatile fatty acids directly from cellulosic biomass, and is therefore an attractive model for biofuel production via consolidated bioprocessing. Shotgun 2D-HPLC-MS/MS was performed on closed-batch cellobiose-grown exponential phase *C. thermocellum* cell lysates, to determine relative protein expression profiles of core metabolic proteins involved in (i) carbohydrate hydrolysis, (ii) cellodextrin transport, (iii) glycolysis, (iv) energy storage, (v) pentose phosphate pathway, (vi) pyruvate catabolism, (vii) end-product synthesis, and (viii) energy generation and pyrophosphate metabolism. Exponential and stationary phase closed-batch cellobiose-grown cell lysates labelled with isobaric tags for relative and absolute quantification were compared using 4-plex 2D-HPLC-MS/MS to determine

[20] Contributing authors: Thomas Rydzak, Peter D. McQueen, Oleg V. Krokhin, Vic Spicer, Peyman Ezzati, Ravi C. Dwivedi, Dmitry Shamshurin, David B. Levin, John A. Wilkins, Richard Sparling. *BMC microbiology* 12, no 1 (2012):214.

changes in core metabolic proteins in response to growth phase. Relative abundance profiles revealed differences in levels of putative enzymes capable of catalyzing parallel pathways. The majority of proteins suspected to be involved in pyruvate catabolism and end-product synthesis were detected with high abundance, with the exception of aldehyde dehydrogenase, ferredoxin-dependent Ech-type NiFe-hydrogenase, and RNF-type NADH:ferredoxin oxidoreductase. Using 4-plex 2D-HPLC-MS/MS, 24% of the 144 core metabolism proteins detected demonstrated differential expression during transition from exponential to stationary phase. Notably, proteins involved in pyruvate synthesis decreased in stationary phase, whereas proteins involved in pyruvate catabolism and end-product synthesis increased in stationary phase. Several proteins directly involved in pyruvate catabolism and end-product synthesis, including pyruvate:ferredoxin oxidoreductases, alcohol dehydrogenases, and a putative bifurcating hydrogenase, demonstrated differential expression during transition from exponential to stationary phase. We identified proteins and metabolic networks predominantly utilized during conversion of cellobiose to end-products based on relative abundance profiles. Low detection levels of ferredoxin-dependent NiFe-hydrogenase, NADH:ferredoxin oxidreductatase, and aldehyde dehydrogenase suggest that H_2 synthesis occurs via bifurcating hydrogenase and ethanol synthesis is predominantly catalyzed by a bifunctional aldehyde/alcohol dehydrogenase. Differential expression in pyruvate:ferredoxin oxidoreductase, bifurcating hydrogenase, and alcohol dehydrogenases in response to growth phase may dictate carbon and electron flux towards end-products. Combined knowledge of relative protein expression levels and

their changes in response to physiological conditions may aid in targeted metabolic engineering strategies and optimization of fermentation condition for improvement of biofuels production.

6.2 Introduction

Clostridium thermocellum ATCC 27405, an anaerobic, Gram-positive thermophilic bacterium, is capable of cellulosome-mediated breakdown of (hemi)cellulose (Freier et al. 1988; Bayer et al. 2004) and simultaneous fermentation of resulting cello-oligosaccharides into hydrogen (H_2) and ethanol (Islam et al. 2006; Rydzak et al. 2009). This reduces the need for separate cellulase production, cellulose hydrolysis, and fermentation, which could improve economic viability of industrial cellulosic biofuel production (Thauer et al. 1977; Lynd et al. 2005; Rydzak et al. 2009). Among cellulolytic microorganisms, *C. thermocellum* exhibits one of the highest growth rates on cellulose (Lynd and Grethlein 1987; Lynd et al. 1989; Lynd et al. 2002). Its high temperature growth optimum aids in H_2 recovery (Angenent et al. 2004), and the availability of an annotated genome sequence (GenBank accession number ZP_00312459.1) allows an initial deduction of possible gene products involved in metabolic pathways *in silico*, expression studies by microarray and proteomic analysis, and genetic engineering (Tyurin et al. 2004; Tyurin et al. 2005; Tripathi et al. 2010) proving it to be an attractive model for biofuel production via consolidated bioprocesing.

Despite these appealing characteristics, *C. thermocellum* normally produces ethanol and H_2 yields (~0.6 and 1.3 mol per mol hexose, respectively) well below the 'Thauer limit' of either 2 moles of ethanol or 4 moles of H_2 per mol hexose, respectively

(Thauer et al. 1977; Rydzak et al. 2009). This is due to branched fermentative pathways that lead to the production of both ethanol and H_2 (with concomitant production of CO_2 and acetate), as well as branches leading to formic acid and lactic acid that compete for carbon and/or electrons required for the production of either ethanol or H_2 (Thauer et al. 1977; Lynd et al. 2005; Rydzak et al. 2009). Metabolic engineering strategies to improve product yields in *C. thermocellum* (Lynd and Cruz 2010) and related species (Guedon et al. 2002) have been only moderately successful and at times resulted in unpredicted changes in product yields (Tripathi et al. 2010). This may be due to the complexity of metabolic networks in which multiple gene products may catalyze parallel reactions (Rydzak et al. 2009), the presence of response regulators that modulate gene and gene-product expression (Buhrke et al. 2004; Soboh et al. 2004; Calusinska et al. 2010), and modulation of enzyme activity via intracellular metabolite levels (Özkan 2004; Willquist and van Niel 2010). While many of the genes and proteins involved in pyruvate catabolism and product formation have been verified via RT-PCR (Carere et al. 2008a), enzyme activity assays (Rydzak et al. 2009), and purification (Lin et al. 1998a; Özkan 2004), a more thorough understanding of metabolic and regulatory networks must be attained.

A number of studies have demonstrated the ability of *C. thermocellum* to control scaffoldin and cellulase mRNA (Mishra et al. 1991; Dror et al. 2003a; Dror et al. 2003b; Dror et al. 2005) and protein (Zhang and Lynd 2005b; Gold and Martin 2007; Zverlov and Schwarz 2008; Raman et al. 2009) levels in response to substrate type and growth rate, whereby cellulosome gene expression is positively regulated through binding of

cellulose and xylan to anti-σ factors, preventing their binding to alternative σ factors required for cellulosome expression (Nataf et al. 2009; Kahel-Raifer et al. 2010), and negatively regulated by cellobiose via a carbon catabolite repression mechanism (Mishra et al. 1991; Zhang and Lynd 2005b). A few studies have looked at expression levels of genes encoding proteins involved in central metabolism and end-product formation. Stevenson and Weimer have looked at expression levels of 17 genes involved in cellulose degradation, intracellular phosphorylation, catabolite repression, and fermentation end-product formation in response to substrate and growth rate (Stevenson and Weimer 2005). More recently, microarray studies have looked at overall gene expression levels and global changes in mRNA levels in response to substrate and dilution rate (Riederer et al. 2011) and growth phase in cellulose-grown batch cultures (Raman et al. 2011). To date, there have been no reports of global protein expression levels of *C. thermocellum*.

We have now completed the first proteomic study of cellobiose-grown batch culture *C. thermocellum* lysates to determine relative abundances of metabolic proteins and responses in their expression levels during different growth phases. Shotgun two-dimensional high performance liquid chromatography-tandem mass spectrometry (2D-HPLC-MS/MS) was used to determine protein relative abundance indexes (RAI) to deduce key metabolic pathways and corresponding enzymes utilized in the conversion of cellobiose to end-products.

6.3 Materials and methods

6.3.1 Organism, media, and growth

Clostridium thermocellum strain DSM 1237 was used for all growth experiments. Culture maintenance, media composition and preparation, and growth experiments in closed-batch cultures were performed as previously described[21] by Rydzak *et al*. (Rydzak et al. 2009); however cells were grown and maintained on 2 g L^{-1} cellobiose (versus 1.1 g L^{-1} cellobiose), which still reflected carbon limited conditions. Samples for end-product, cell biomass, and pH measurements were taken throughout growth, while samples for proteomic analysis were taken in exponential and stationary phase (OD_{600} ~ 0.37 and ~0.80, respectively).

[21] Refer to section 2.3.1 for details

6.3.2 Cell growth, pH, and end-product analysis

Cell growth measurements, sample processing, pH measurement, product gas, protein, sugar, and end-product analyses, and biomass and gas calculations were performed as previously described[22] (Rydzak et al. 2009). Data represent the average of a minimum of three biological replicates.

6.3.3 Preparation of cell extracts for proteomic analysis

Exponential and stationary phase cell cultures (10.5 mL) were centrifuged (10000 x g, 5 minutes, 4˚C). Cells pellets were washed 3 times in 500 µL 1x PBS buffer and then frozen at -80˚C. Cell pellets were re-suspended in 540 µL lysis buffer (Tris-HCl, 10 mM, pH 7.4; $CaCl_2$, 3 mM; 2 mM $MgCl_2$, 2 mM; bacterial protease inhibitor, 1.0 %; Tergitol NP-40, 0.1%) and sonicated 5 rounds for 15 seconds each round with cooling on ice in between rounds. Unlysed cells were removed by centrifugation (14000 x g, 10 minutes) and protein concentration of supernatant was determined Bicinchononic Acid (BCA) Protein Assay Kit (Pierce Biotechnology, Rockford, IL) as outlined by the manufacturer. The supernatant was stored at -80˚C. An aliquot corresponding to 200 µg of protein was mixed with 100 mM ammonium bicarbonate, reduced with dithiothreitol (10 mM), and incubated for 30 minutes at 57˚C. Proteins were then alkylated with iodoacetamide (50 mM) for 30 minutes at room temperature in the dark. Excess iodoacetamide was quenched with dithiothreitol (16 mM). Peptides were digested in a 1:50 trypsin/protein

[22] Refer to section 2.3.2 for details

ratio (Promega, Madison, WI) for 10 hours at 37˚C. Samples were then acidified with an equal volume of 3% trifluoroacetic acid (TFA), lyophilized, and re-suspended in 270 µL of 0.1% TFA. Samples were loaded on a C18 X-Terra column (1 × 100 mm, 5 µm, 100 Å; Waters Corporation, Milford, MA, USA), desalted using 0.1% TFA, and peptides were eluted with 50% acetonitrile. Desalted samples were stored at -80˚C for 2D-HPLC-MS/MS analysis. For comparative proteomic analysis of exponential and stationary phase cells, each trypsinized protein sample (100 µg) was labelled with isobaric Tags for Relative and Absolute Quantification (iTRAQ) reagent (Applied Biosystems, Foster City, CA, USA) as outlined by the manufacturer. Samples differentially labelled with isobaric tags of different masses (exponential phase replicate A [iTRAQ tag 114], exponential phase replicate B [iTRAQ tag 115], stationary phase replicate A [iTRAQ tag 116], stationary phase replicate B [iTRAQ tag 117]) were mixed in equal proportions and subjected to 2D-HPLC-MS/MS (Aggarwal et al. 2006; Zieske 2006).

6.3.4 Two-dimensional high-performance liquid chromatography-mass spectrometry analysis

Trypsinized peptides with or without iTRAQ label were separated in the first dimension using an Agilent 1100 Series HPLC system (Agilent Technologies, Wilmington, DE). Samples were injected onto a C18 X-Terra column (1 × 100 mm, 5 µm, 100 Å; Waters Corporation, Milford, MA, USA) and eluted with a linear water-acetonitrile gradient (20 mM ammonium formate, pH 10, in both eluents A and B, 1% acetonitrile/min, 150 µL/min flow rate). A concentrated 200 mM solution of ammonium formate at pH 10 was prepared as described by Gilar *et al*. (Gilar et al. 2005). Buffers A

and B for first-dimension separation were prepared by a 1/10 dilution of this concentrated buffer with water and acetonitrile, respectively. Sixty 1-min fractions were collected (roughly 6.6 μg/fraction). Samples were concatenated (fraction 1 and 31, 2 and 32, etc.) into a total of 30 fractions as described by Dwivedi *et al.* (Dwivedi et al. 2008). Each was lyophilized and re-suspended in 100 μL of 0.1% formic acid. A splitless nanoflow Tempo LC system (Eksigent, Dublin, CA, USA) with 20 μL sample injection via a 300 μm × 5 mm PepMap100 precolumn and a 100 μm × 150 mm analytical column packed with 5 μm Luna C18(2) (Phenomenex, Torrance, CA) was used in the second-dimension separation prior to tandem MS analysis. Both eluents A (2% acetonitrile in water) and B (98% acetonitrile) contained 0.1% formic acid as ion-pairing modifier. A 0.33% acetonitrile/min linear gradient (0-30% B) was used for peptide elution, providing a total 2 hour run time per fraction in the second dimension.

6.3.4.1 ***Mass spectrometry***

A QStar Elite mass spectrometer (Applied Biosystems, Foster City, CA) was used in standard MS/MS data-dependent acquisition mode with a nano-electrospray ionization source. The 1 s survey MS spectra were collected (*m*/*z* 400-1500) followed by three MS/MS measurements on the most intense parent ions (80 counts/s threshold, +2 to +4 charge state, *m*/*z* 100-1500 mass range for MS/MS), using the manufacturer's "smart exit" settings and iTRAQ settings. Previously targeted parent ions were excluded from repetitive MS/MS acquisition for 60 s (50 mDa mass tolerance).

6.3.4.2 *Database search, protein identification, and statistical analysis*

Raw spectra files of unlabeled peptides were treated using standard script (Analyst QS2.0) to generate text files in Mascot Generic File format (MGF) (Perkins et al. 1999). These files containing the MS/MS spectra information for all 30 fractions were concatenated and submitted for protein identification using an in-house GPU-based peptide identification engine described by McQueen *et al.* (McQueen et al. 2012). Standard QTOF settings were used for the search: 100 ppm and 0.4 Da mass tolerance for parent and fragment ions, respectively. Permitted amino acid modifications included constant carbamidomethylation of Cys. The "relative abundance index" (RAI) for each protein was calculated as the number of spectral counts (SpC) divided by molecular mass (Mr) of protein.

Spectra files of iTRAQ labelled peptides were also analyzed using ProteinPilot software version 2.0.1 (Applied Biosystems/MDS Sciex, Concord, ON, Canada). The search parameters were complete modifications of Cys alkylation with iodoacetic acid, and inbuilt iTRAQ analysis residue modifications settings were on. The reporter ion (iTRAQ tag) intensities for each tryptic peptide identified (with expectation values <-1.5) were histogrammed by the log2 of the ratios (Z0 = tag_{116}/tag_{114}, Z1 = tag_{117}/tag_{115}, Z2 = tag_{115}/tag_{114}, and Z3 = tag_{117}/tag_{116}) to build overall peptide population distributions, where exponential phase replicates were labelled with tags 114 and 115, respectively, and stationary phase replicates were labeled with tags 116 and 117, respectively. Peptide level Z-scores are mapped as the distance from the population mean in units of standard deviation; initial protein-level Z-scores are average of the member peptide Z-score

values. The Z-scores (Z2,Z3) contain information about the stability across biological replicates at the same growth state. We have devised a simple algorithm to combine these with the differential data in (Z0,Z1), expressed as the difference between the magnitudes of vectors from the origin to points (Z0,Z1) and (Z2,Z3), scaled by the widths of their peptide histogram distributions. The sign of the transformed value is determined by the angle subtended by a vector from the origin to the point (Z0,Z1). We denote this combined value as the vector difference (V_{diff})[23]. Z-scores were converted into fold-changes by taking 2 to the power of the Z-score.

---

[23] Vector difference was used to asses statistical significance of protein expression ratios between exponential and stationary phase samples. Z-scores derived from reporter ion ratios for each protein were converted into vectors, and the vector difference was calculated using the following formulas:

1a $$V_{116/114\ vs\ 117/115} = \sqrt{{Z_{116/114}}^2 + {Z_{117/115}}^2}$$

1b $$V_{115/116\ vs\ 117/114} = \sqrt{{Z_{115/116}}^2 + {Z_{117/114}}^2}$$

1c $$V_{115/114\ vs\ 117/116} = \sqrt{{Z_{115/114}}^2 + {Z_{117/116}}^2}$$

2a $$V_{\text{diff}} = V_{116/114\ \text{vs}\ 117/115} + V_{115/116\ \text{vs}\ 117/114} - 2(V_{115/114\ \text{vs}\ 117/116})$$

where Z is the calculated z-score of a particular ratio, V is the vector of a particular ratio, and V_{diff} is the vector difference of all calculated vectors deduced from z-scores.

6.4 Results and discussion

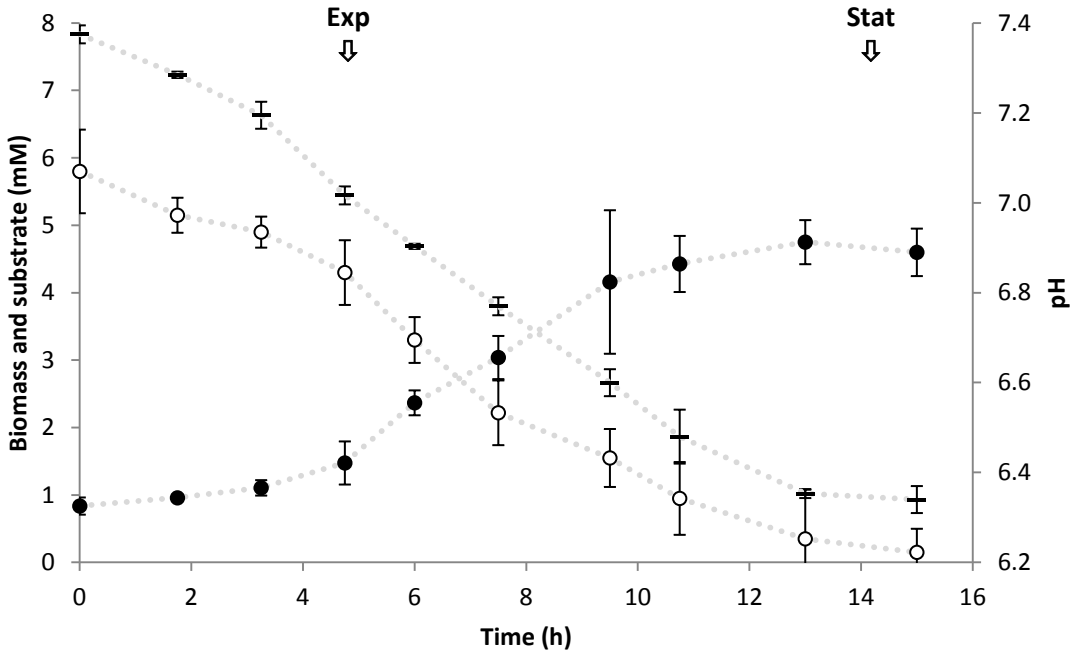
6.4.1 Growth and end-product synthesis

In this study, we investigated the relative abundance profiles (RAI) of core metabolic proteins in exponential phase cultures, and changes in protein expression in response to growth phase. All *C. thermocellum* DSM 1237 cultures were grown in complex 1191 medium closed-batch cultures with no pH control, on 2.2 g L^{-1} cellobiose. Cell growth (as indicated by biomass production), substrate consumption, change in pH, and end-product formation during growth are shown in Figure 6.1. Cultures reached stationary phase in ~14 h upon exhaustion of cellobiose, suggesting carbon limited growth, and had a final biomass and pH of 4.8 mM and 6.3, respectively. In agreement with previous reports (Lamed and Zeikus 1980a; Lynd et al. 1989; Strobel 1995; Stevenson and Weimer 2005; Islam et al. 2006; Rydzak et al. 2009) H_2, CO_2, ethanol, and acetate were major end-products and paralleled growth and cellobiose consumption. A slight inversion of acetate-to-ethanol ratio was observed during the transition to stationary phase. This was also observed by Raman *et al*. (Raman et al. 2011) and could be stimulated by H_2 build-up (Lamed and Zeikus 1980a; Ben-Bassat et al. 1981; Freier et al. 1988; Lamed et al. 1988; Soboh et al. 2004; Rydzak et al. 2011). Formate was also a major end-product in agreement with Sparling *et al*., Islam *et al*., and Rydzak *et al*. (Islam et al. 2006; Sparling et al. 2006; Rydzak et al. 2009; Rydzak et al. 2011). The lack of formate detection in some *C. thermocellum* studies could be attributed to HPLC detection methods or media composition (Magnusson et al. 2009). Lactate production

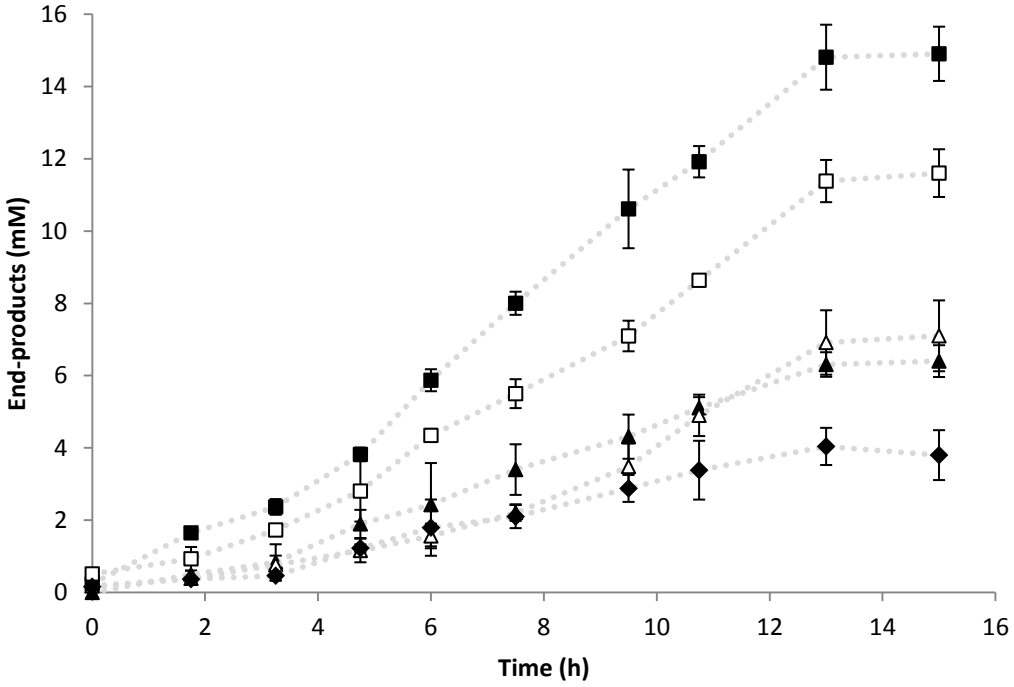
was below detectable limits as expected under carbon-limited conditions (Islam et al. 2006). Carbon recovery (91%) and O/R ratio (0.93) confirm that major end-products were accounted for.

Figure 6.1: Fermentation growth and metabolite production. (A) Cellobiose utilization, biomass production, pH change, and (B) metabolite production plots of *C. thermocellum* grown in 1191 medium batch cultures on 2 g L^{-1} cellobiose. Arrows indicate sampling points for exponential and stationary phase proteomic analysis. Biomass (black circle), cellobiose (white circles), pH (dash), H_2 (black square), CO_2 (white square), acetate (black triangel), ethanol (white triangle), formate (black diamond).

A)


Exp
Stat
Biomass and substrate (mM)
pH
Time (h)
0
1
2
3
4
5
6
7
8
6.2
6.4
6.6
6.8
7.0
7.2
7.4
0
2
4
6
8
10
12
14
16

B)


End-products (mM)
Time (h)
0
2
4
6
8
10
12
14
16

6.4.2 Relative protein abundance using shotgun and 4-plex 2D-HPLC-MS/MS

Two-dimensional high-performance liquid chromatography-tandem mass spectrometry detected (with a 99.9% confidence score and minimum peptide detection threshold of 2) a total of 1575 of 3236 predicted proteins, including 1468 proteins detected by shotgun 2D-HPLC-MS/MS in exponential phase cell lysates, and 1071 proteins detected by 4-plex 2D-HPLC-MS/MS of duplicate iTRAQ labelled exponential and stationary phase samples. We have currently focused strictly on core metabolic proteins that primarily dictate the majority of carbon and electron flux from cellulose and/or cellobiose to end-products. Putative proteins responsible for (i) carbohydrate hydrolysis, (ii) cellodextrin transport, (iii) glycolysis, (iv) energy storage, (v) pentose phosphate pathway, (vi) pyruvate catabolism, (vii) end-product synthesis, and (viii) energy generation and pyrophosphate metabolism are examined.

Determination of relative protein expression profiles is essential for targeted metabolic engineering strategies for strain improvement (i.e. optimization of product titres, increasing growth rates, preventing product inhibition). In recent years, spectral counts obtained from shotgun proteomic approaches have been shown to be a good estimation of protein abundance (Liu et al. 2004; Old et al. 2005; Zybailov et al. 2005; Zhu et al. 2010). Liu *et al*. demonstrated a linear correlation between spectral counts and relative protein abundance (R^2 = 0.9997) over 2 orders of magnitude (Liu et al. 2004). Quantitation via spectral counting shows strong correlation with isotopic label-based approaches (Zybailov et al. 2005), such as $^{14}N/^{15}N$ and precursor peak area intensity measurements (Old et al. 2005). Given that larger proteins generally give rise to a greater

number of peptides following digestion, and thus a greater number spectral counts, relative protein abundance is commonly standardized to account for protein size. Rappsilber *et al*. used "protein abundance index" (PAI), which represents the number of peptides identified divided by the number of theoretically available peptides, to quantify the relative abundance of proteins detected by MS analyses (Rappsilber et al. 2002). Zybailov *et al*. and Florens and Washburn used "normalized spectral abundance factor" (NSAF), which represents the number of spectral counts divided by protein length (Florens and Washburn 2006; Zybailov et al. 2006). In this study, we have quantified 2D-HPLC-MS/MS abundance profiles based on each proteins "relative abundance index" (RAI), calculated as the number of spectral counts (SpC) divided by molecular mass (Mr) of protein.

While the number of proteins detected by shotgun 2D-HPLC-MS/MS was greater than 4-plex 2D-HPLC-MS/MS, RAI values followed a similar trend, further verifying general protein abundance using both acquisition methods (Supplementary Figure 3). However, the RIA per a given protein was lower using the 4-plex versus shotgun acquisition method. This was expected given that the 4-plex run simultaneously measures four samples and associated labels, thus reducing available peptide acquisition time. Due to the increased sensitivity and deeper coverage, we use the RAI data of shotgun exponential phase samples when discussing relative protein expression profiles in the text.

6.4.3 Changes in stationary phase protein expression levels using iTRAQ 2D-HPLC-MS/MS

Understanding cellular responses to pH change, end-product accumulation, and substrate limitation may aid in improving strain growth through targeted deregulation of factors that limit growth and production of desired end-products. Comparison of expression levels of two biologically replicated iTRAQ-labelled exponential phase and stationary phase samples (tagged with reporter ions 114 & 115 and 116 & 117, respectively) was performed using 4-plex 2D-HPLC-MS/MS. Ratios of z-score values among exponential and stationary phase biological replicates (reporter ion ratios 115/114 vs 117/116) and between exponential phase *vs* stationary phase samples (reporter ion ratio 116/114 *vs* 117/115) are plotted in Supplementary Figures 4a and 4b, respectively, to illustrate correlation between biological replicates. While Supplementary Figure 4a shows good correlation between biological replicates (perfect correlation represented by coordinates 0,0), a number of proteins have poorer correlation between replicates. To determine the statistical significance of protein expression ratios between exponential and stationary phase samples when factoring in the deviation between biological replicates, z-scores ratios for each protein were converted into vectors, and the vector difference was calculated (see Methods). Exponential phase *vs* stationary phase z-scores ($Z_{116/114}$ and $Z_{117/115}$) are plotted in Supplementary Figure 4b and are color coordinated based on vector difference. Vector differences greater than 2 represent proteins with the highest change in expression, while vector differences less than 0.5 represent proteins with little statistical change in expression. This calculation

allowed us to eliminate values of high change between exponential and stationary phase samples when variation between replicates was higher than that of the change in exponential *vs* stationary phase samples. We propose that a vector difference of ≥ 0.5 as a confident change in expression between exponential and stationary phase proteins. Changes in protein expression levels were manually verified. Differences in protein expression between stationary and exponential phase cell lysates of core metabolic proteins are summarized in Table 6.1. A total of 166 of 252 encoded core metabolic proteins were detected using a combination of both shotgun and 4-plex acquisition methods. Twenty-four percent (24%) of proteins detected using 4-plex 2D-HPLC-MS/MS had a change in expression with a V_{diff} greater than 0.5. Nineteen percent (19%) of these proteins increased during the transition from exponential to stationary phase, while only 4% decreased in stationary phase, and 15% of these differentially expressed proteins changed by a magnitude greater than 1.

Table 6.1: Protein detection using shotgun (single-plex) and iTRAQ labelled 4-plex 2D-HPLC-MS/MS and relative changes in protein expression levels. Core metabolic proteins were classified into functional categories. The number of proteins that changed during transition from exponential to stationary phase were listed only when their vector difference (V_{diff}) was greater than 0.5.

Core Metabolic Protein Categories	Gene Count	Proteins Detected			Changes in Protein Levels (Stat/Exp)	
		1-Plex	4-Plex	Total	$V_{diff} \geq 0.5$ Increased	$V_{diff} \geq 0.5$ Decreased
Non-catalytic cellulosomal proteins	8	5	6	7	0	0
Cellulosomal glycosidase	73	29	26	31	2	1
Non-cellulosomal glycosidases	35	17	13	19	3	0
RsgI-like σ-factors and anti-σ^I factors	9	3	2	3	0	0
Cello-oligosaccharide ABC transporters	14	9	8	10	2	1
Glycolysis	20	15	15	15	3	1
Pentose phosphate pathway	6	4	3	5	1	0
Energy storage	13	11	11	13	3	0
Pyruvate formation from phosphoenolpyruvate	8	8	8	8	0	2
End-product synthesis from pyruvate	49	39	38	41	12	0
Energy generation	17	14	14	14	2	1
Total	252	154	144	166	28	6

6.4.4 Central carbohydrate metabolism

Global proteomic analysis is fundamental in verifying carbon utilization and end-product synthesis pathways. While mRNA expression profiles provide a great wealth of information with regards to transcriptional patterns, proteomics can rectify the discrepancy between transcription and translation. Relative protein expression profiles allow us to deduce which proteins, and therefore pathways, are utilized during carbohydrate metabolism. Furthermore, changes in protein levels in response to growth phase may help in hypothesizing regulatory elements that may be targeted for increasing product yields during monoculture and co-culture fermentation processes. Below we discuss key proteins involved in carbohydrate utilization and transport, glycolysis, energy storage, pentose phosphate production, pyruvate catabolism, end-product synthesis, and energy production.

6.4.4.1 *Proteins involved in cellulose and hemicellulose degradation and transport*

6.4.4.1.1 Cellulose hydrolysis

C. thermocellum encodes a number of carbohydrate active enzymes (CAZymes) allowing for efficient degradation of cellulose and associated polysaccharides (Carbohydrate Active Enzyme database; www.cazy.org/). These include (i) endo-β-glucanases, which cleave internal amorphous regions of the cellulose chain into shorter soluble oligosaccharides, (ii) exo-β-glucanases (cellodextrinases and cellobiohydrolases), which act in a possessive manner on reducing or non-reducing ends of the cellulose chain liberating shorter cellodextrins, and (iii) β-glucosidases (cellodextrin and cellobiose

phosphorylases), which hydrolyze soluble cellodextrins ultimately into glucose (Lynd et al. 2002). Other glycosidases that allow hydrolysis of lignocellulose include xylanases, lichenases, laminarinases, β-xylosidases, β-galactosidases, and β-mannosidases, while pectin processing is accomplished via pectin lyase, polygalacturonate hydrolase, and pectin methylesterase (Spinnler et al. 1986; Demain et al. 2005). These glycosidases may be secreted as free enzymes or may be assembled together into large, cell-surface anchored protein complexes ("cellulosomes") allowing for the synergistic breakdown of cellulosic material. The cellulosome consists of a scaffoldin protein (CipA) which contains (i) a cellulose binding motifs (CBM) allowing for the binding of the scaffoldin to the cellulose fiber, (ii) nine type I cohesion domains with that mediate binding of various glycosyl hydrolases via their type I dockerin domains, and (iii) a type II dockerin domain which mediates binding to the type II cohesion domain found on the cell-surface anchoring proteins. The cell-surface anchoring proteins are in turn noncovalently bound to the peptidoglycan cell wall via C-terminal surface-layer homology (SLH) repeats (Demain et al. 2005).

During growth on cellulose, the cellulosome is attached to the cell in early exponential phase, released during late exponential phase, and is found attached to cellulose during stationary phase (Demain et al. 2005). Cellulosome expression has been shown to be negatively regulated by cellobiose via a carbon catabolite repression mechanism (Mishra et al. 1991; Zhang and Lynd 2005b), and positively regulated through binding of cellulose and associated polysaccharides to anti-σ factors, allowing expression of cellulosomal components using alternative σ factors (Nataf et al. 2009;

Kahel-Raifer et al. 2010), suggesting that the cellulosome should not be expressed in cellobiose-grown cultures. The ability of *C. thermocellum* to control scaffoldin and cellulase mRNA (Mishra et al. 1991; Dror et al. 2003a; Dror et al. 2003b; Dror et al. 2005) and protein (Zhang and Lynd 2005b; Gold and Martin 2007; Zverlov and Schwarz 2008; Raman et al. 2009) levels in response to substrate type and growth rate has been extensively studied, and reveals that expression of cellulosomal enzymes is present in the absence of cellulose, albeit at lower levels. We detected expression of 7 cellulosomal structural proteins, 31 cellulosome-associated glycosidases, and 19 non-cellulosomal CAZymes on cellobiose using 2D-HPLC-MS/MS (Supplementary Table 2).

Of the 8 encoded non-catalytic cellulosomal proteins, 7 were detected using the combined acquisition methods (shotgun and 4-plex). SdbA (Cthe__1307) was the most abundant anchoring protein, and scaffoldin CipA (Cthe__3077) was found in the top 50% of total proteins detected (RAI = 0.42). OlpB, Orf2p, and OlpA located downstream of CipA (Cthe__3078-3080) were also detected, but at sequentially lower levels. Expression of cellulosomal anchoring proteins Cthe__0452 and Cthe_0736 was also detected, but only during 4-plex acquisition. Microarray studies revealed that transcription of *sdbA* was low compared to *cipA, olpB, orf2p,* and *olpA* on cellulose (Raman et al. 2011), while nano-LC-ESI-MS revealed that SdbA was only expressed in cellobiose-grown cultures (Gold and Martin 2007). This coincided with our high SdbA levels detected in cellobiose-grown cell lysates. On cellulose, Raman *et al*. found no change in *cipA* transcription and a 2-fold increase in *orf2p* transcription in stationary phase (Raman et al. 2011), while Dror *et al*. observed an increase in transcription of *orf2p*

as well as *cipA* and *olpB* with decreasing growth rate (Dror et al. 2003b). Alternatively, Gold *et al*. showed similar expression of Orf2p relative to CipA in both cellobiose and cellulose-grown samples and increased expression of OlpB in cellobiose-grown cultures (Gold and Martin 2007). We, however, did not observe any statistically relevant changes of cellulosomal proteins on cellobiose between exponential and stationary phase cells.

C. thermocellum encodes 73 glycosidases containing a type I dockerin, 65 of which have been detected and characterized at the protein level (Raman et al. 2011). 2D-HPLC-MS/MS of exponential phase cell extracts detected 31 cellulosomal glycosidases (Supplementary Table 2), 19 of which were in the top 90th percentile of total proteins detected (RAI > 0.1). In addition to high RAI levels of CelS, a cellulosomal subunit shown to be highly expressed (Dror et al. 2003a; Dror et al. 2005), XynC, CelA, XynA/U, CelG, and glycosidase Cthe_0821 were also detected in high amounts. Other characterized cellulosomal glycosidases detected included CelB, XynZ, XghA, CelR, CelK, and CelV. Proteomic analysis has shown that exoglucanases CelS and CelK, and endoglucanase CelJ are higher in cellulose versus cellobiose-grown cultures, while hemicellulases (XynZ, XynC, XynA/U, XghA, Cthe__0032) and endoglucanases belonging to family GH5 (CelB, CelG, Cthe__2193) and GH8 (CelA) were more abundant in cellobiose versus cellulose-grown cultures (Gold and Martin 2007). This agrees with our relative protein abundance profiles exhibiting high xylanase, GH5 family glycosidase, and CelA expression, and lower CelK and CelJ expression in exponential cellobiose-grown cell extracts. Interestingly, despite the presence of xylanases, sequence homology-based annotation has not revealed the presence of xylose reductase, xylitol

dehydrogenase, xylose isomerase, or xylulokinase required for xylose utilization. This suggests that, in the absence of cellulose, *C. thermocellum* may be predisposed to expressing xylanases, which typically degrade hemicellulosomal xylans, exposing buried cellulose fibres.

With the exception of a 2-fold increase in cellulosomal glycosidases Cthe__0821, Cthe__2761, and Cthe__0745, and a 1.6-fold decrease in XynD (Cthe__0625), no other statistically significant changes were observed in detected cellulosomal cellulases during transition from exponential to stationary phase. While this contradicted high variability in transcription of cellulosomal glycosidases of cellulose-grown cells (Raman et al. 2011), lack of variability in our experiment may have been attributed to differences in growth substrate used. In fact, Dror *et al*. found negligible changes in transcription of *celB, celG*, *celD,* and *celF* between exponential and stationary phase cellobiose-grown cultures (Dror et al. 2005). Alternatively, our processing method, which included several wash steps prior to lysing the cells, may have imposed bias and variability by potentially washing off weakly bound cellulosomal glycosidases.

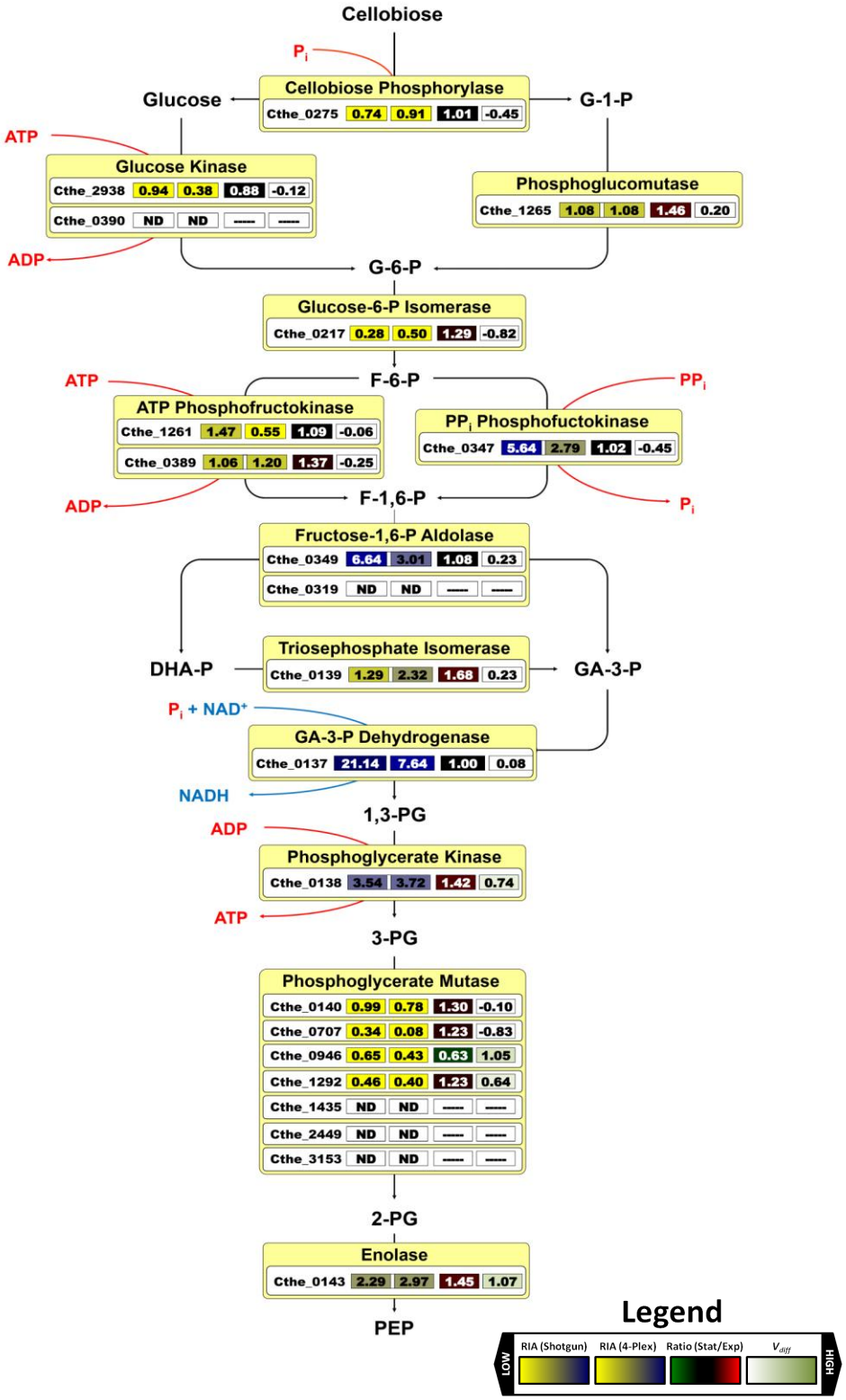
In addition to cellulosomal glycosidases, 35 non-cellulosomal CAZymes that do not have a dockerin domain are encoded in the genome. Of the 19 non-cellulosomal CAZymes detected in exponential phase cell extracts using 2D-HPLC-MS/MS, half had RAI ratios in the top 90% (RAI > 0.1) of total peptides detected. Not surprisingly, the most abundant CAZyme cellobiose phosphorylase Cthe__0275 (glycosyltransferase family 36), which is involved in intracellular phosphorylytic cleavage of cellobiose, fell within the top 25% of detected proteins. Cellobiose phosphorylase Cthe__2989 was also

found in high amounts (RAI = 0.23), whereas glycosyltransferase Cthe__1221, a putative cyclic β-1,2 glucan synthetase, was detected in the bottom 10% of all proteins detected (Figure 6.2a). CelI, an endo-1,4-β-glucanase (Cthe__0040) was not detected, consistent with growth on cellobiose. Other highly abundant non-cellulosomal CAZymes include amidohydrolase (Cthe__1777), glucoamylase (Cthe__1787), xylanase A precursor (Cthe__1911), α-N-arabinofuranosidas (Cthe__2548), CelC (Cthe__2807), and several less characterized glycosidases (Cthe__3163, Cthe__1911, Cthe__2989). While Raman *et al*. report decreased transcription of glycosyltransferases involved in intracellular phosphorolytic cleavage of cellodextrin and cellobiose (family 36), and increased transcription of a number of other CAZymes in response to decreased substrate availability in stationary phase (Raman et al. 2011), we saw no statistically significant changes in CAZyme expression with two exceptions: LicA (Cthe__2809) increased in stationary phase, consistent with reports by Newcomb and Wu (Newcomb et al. 2011) and Raman *et al*. (Raman et al. 2011), and acetyl xylan esterase (Cthe__3063) also increased contradicting previously reported microarray data (Raman et al. 2011). CelC expression (Cthe__2807), which is negatively regulated by the co-transcribed LacI family transcriptional regulator GlyR3 (Cthe__2808), has been consistently been shown to increase in the presence of laminaribiose (Newcomb et al. 2007) and in stationary phase on cellulose (Raman et al. 2011) and cellobiose (Mishra et al. 1991). While CelC expression was shown to have an overall increase in stationary phase among biological replicates, deviation between replicates makes it difficult to tell if this is simply an artifact. Finally, of the 7 membrane-associated RsgI-like anti-σ^I factors proposed to

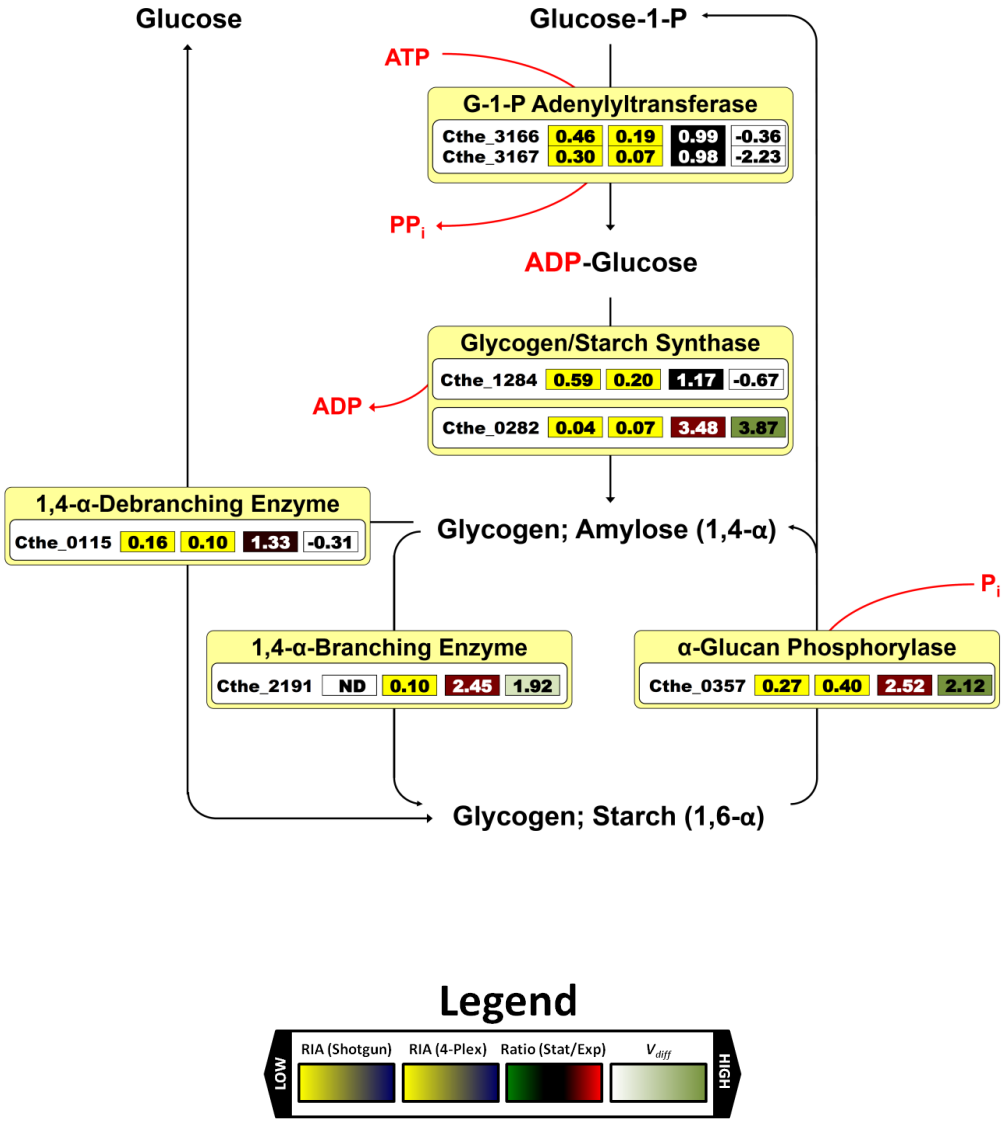
activate expression of different glycosidases in the presence of cellulose and other polysaccharides, three have been detected (Cthe__0059, Cthe__0267, and Cthe__2521). The binding of a particular polysaccharide to corresponding anti-σ^I factor N-terminal carbohydrate binding domains is proposed to promote the C-terminal release of putative alternative σ^I-factors (SigI) encoded upstream of these anti-σ^I factors, allowing for expression of cselect glycosidases, some of which (ex. CelA) are encoded downstream of the anti-σ^I factors that regulate their expression (Kahel-Raifer et al. 2010; Riederer et al. 2011).

Figure 6.2: Relative abundance indexes and changes in protein expression levels of protein involved in glycolysis, glycogen metabolism, and pentose phosphate pathway. Relative abundance indexes (values 1 and 2), changes in protein expression ratios (value 3), and associated V_{diff} values (value 4) indicating confidence levels of changes in expression ratios are indicated for enzymes involved in (A) glycolysis, (B) glycogen metabolism, and (C) pentose phosphate pathway. Given the absence of genes encoding transaldolase, we propose an alternative pathway for production of xylulose-5-phosphate and ribose-5-phosphate using fructose-1,6-P aldolase and PP_i phosphofructokinase. Metabolites shown in grey are those commonly metabolized by these enzymes. G-1-P, glucose-1-phosphate; G-6-P, glucose-6-phosphate; F-1-P, fructose-1-phosphate; F-1,6-P, fructose-1,6-bisphosphate; DHA-P, dihydroxyacetone phosphate; GA-3-P, glyceraldehydes-3-phosphate; PG, phosphoglycerate; PEP, phosphoenolpyruvate; X-5-P, xylulose-5-phosphate; E-4-P, erythrose-4-phosphate; S-7-P, sedoheptulose-7-phosphate; S-1,7-P, sedoheptulose-1,7-phosphate; R-5-P, ribose-5-phosphate; Ru-5-P, ribulose-5-phosphate; ND, not detected.

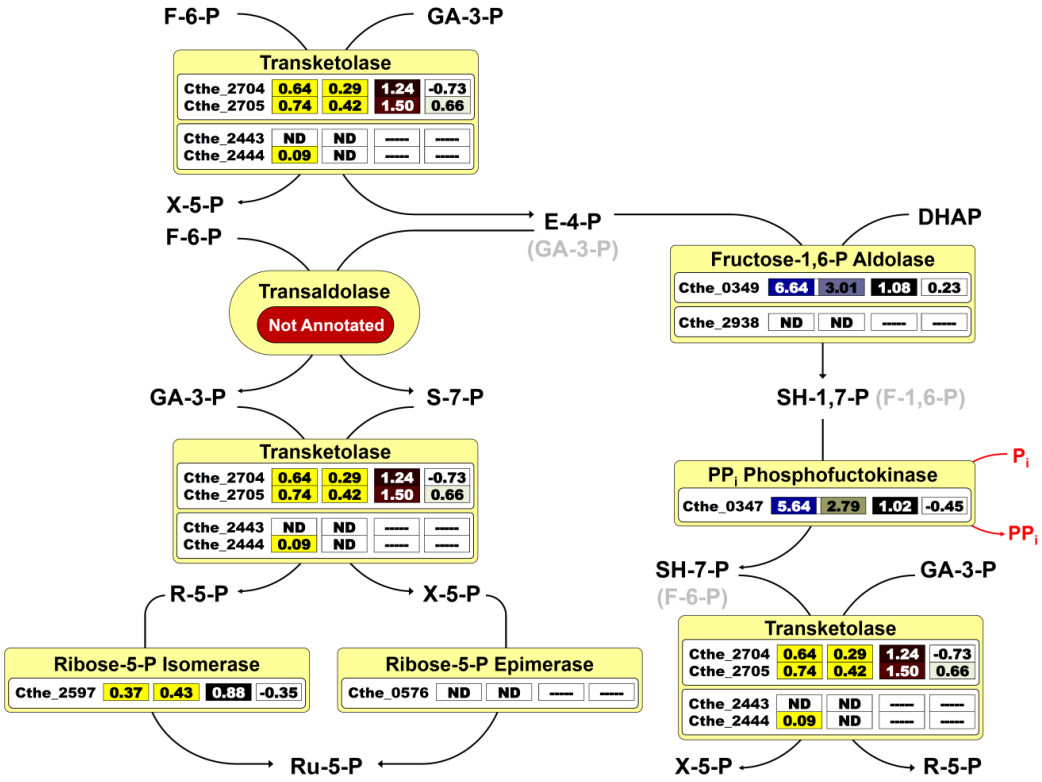
A)

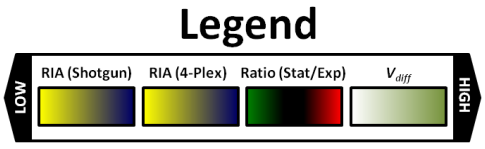

Cellobiose
Pi
Cellobiose Phosphorylase
Cthe_0275 0.74 0.91 1.01 -0.45
Glucose
G-1-P
ATP
Glucose Kinase
Cthe_2938 0.94 0.38 0.88 -0.12
Cthe_0390 ND ND ----- -----
Phosphoglucomutase
Cthe_1265 1.08 1.08 1.46 0.20
ADP
G-6-P
Glucose-6-P Isomerase
Cthe_0217 0.28 0.50 1.29 -0.82
F-6-P
ATP
PPi
ATP Phosphofructokinase
Cthe_1261 1.47 0.55 1.09 -0.06
Cthe_0389 1.06 1.20 1.37 -0.25
PPi Phosphofuctokinase
Cthe_0347 5.64 2.79 1.02 -0.45
ADP
F-1,6-P
Pi
Fructose-1,6-P Aldolase
Cthe_0349 6.64 3.01 1.08 0.23
Cthe_0319 ND ND ----- -----
DHA-P
Triosephosphate Isomerase
Cthe_0139 1.29 2.32 1.68 0.23
GA-3-P
Pi + NAD+
GA-3-P Dehydrogenase
Cthe_0137 21.14 7.64 1.00 0.08
NADH
1,3-PG
ADP
Phosphoglycerate Kinase
Cthe_0138 3.54 3.72 1.42 0.74
ATP
3-PG
Phosphoglycerate Mutase
Cthe_0140 0.99 0.78 1.30 -0.10
Cthe_0707 0.34 0.08 1.23 -0.83
Cthe_0946 0.65 0.43 0.63 1.05
Cthe_1292 0.46 0.40 1.23 0.64
Cthe_1435 ND ND ----- -----
Cthe_2449 ND ND ----- -----
Cthe_3153 ND ND ----- -----
2-PG
Enolase
Cthe_0143 2.29 2.97 1.45 1.07
PEP
Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
Vdiff
HIGH

B)


Glucose
Glucose-1-P
ATP
G-1-P Adenylyltransferase
Cthe_3166 0.46 0.19 0.99 -0.36
Cthe_3167 0.30 0.07 0.98 -2.23
PPi
ADP-Glucose
Glycogen/Starch Synthase
Cthe_1284 0.59 0.20 1.17 -0.67
ADP
Cthe_0282 0.04 0.07 3.48 3.87
1,4-α-Debranching Enzyme
Cthe_0115 0.16 0.10 1.33 -0.31
Glycogen; Amylose (1,4-α)
Pi
1,4-α-Branching Enzyme
Cthe_2191 ND 0.10 2.45 1.92
α-Glucan Phosphorylase
Cthe_0357 0.27 0.40 2.52 2.12
Glycogen; Starch (1,6-α)
Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
V_{diff}
HIGH

C)


F-6-P
GA-3-P
Transketolase
Cthe_2704 0.64 0.29 1.24 -0.73
Cthe_2705 0.74 0.42 1.50 0.66
Cthe_2443 ND ND ----- -----
Cthe_2444 0.09 ND ----- -----
X-5-P
E-4-P
(GA-3-P)
DHAP
F-6-P
Transaldolase
Not Annotated
Fructose-1,6-P Aldolase
Cthe_0349 6.64 3.01 1.08 0.23
Cthe_2938 ND ND ----- -----
GA-3-P
S-7-P
SH-1,7-P (F-1,6-P)
Transketolase
Cthe_2704 0.64 0.29 1.24 -0.73
Cthe_2705 0.74 0.42 1.50 0.66
Cthe_2443 ND ND ----- -----
Cthe_2444 0.09 ND ----- -----
PPi Phosphofuctokinase
Cthe_0347 5.64 2.79 1.02 -0.45
Pi
PPi
R-5-P
X-5-P
SH-7-P
(F-6-P)
GA-3-P
Ribose-5-P Isomerase
Cthe_2597 0.37 0.43 0.88 -0.35
Ribose-5-P Epimerase
Cthe_0576 ND ND ----- -----
Transketolase
Cthe_2704 0.64 0.29 1.24 -0.73
Cthe_2705 0.74 0.42 1.50 0.66
Cthe_2443 ND ND ----- -----
Cthe_2444 0.09 ND ----- -----
Ru-5-P
X-5-P
R-5-P


Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
V_{diff}
HIGH

6.4.4.1.2 Cellodextrin transport

Oligosaccharides derived from cellulose hydrolysis are actively transported via ATP-dependent cello-oligosaccharide ABC transporters (Strobel et al. 1995). Of the five encoded cello-oligosaccharide ABC transporters, only Cthe__0391-0393, Cthe__1018-1020, and Cthe__1862 were detected in significant amounts, consistent with mRNA expression levels reported by Raman *et al*. (Raman et al. 2011). While the RAI was low for membrane spanning domains of these transporters, cytoplasmic nucleotide binding domains and extracellular carbohydrate-binding domains (Cbp) had higher RAI values (Figure 6.2a, Supplementary Table 2). Characterization of Cbp subunits revealed that CbpA (Cthe__0393) binds only to cellotriose, CbpB (Cthe__1020) binds to cellodextrins of different lengths (G2-G5), while CbpC and CbpD (Cthe__2128 and Cthe__2446, respectively) preferentially bind to G3-G5 cellodextrins (Nataf et al. 2009). Given the absence of cellodextrins longer than cellobiose (G2) in our growth medium, the absence of the latter transporters Cthe__2125-2128 and Cthe__2446-2449 is not surprising. While high expression levels of cellotriose ABC transporter were a bit surprising given the cells were grown on cellobiose, studies have shown that *C. thermocellum* and other cellulolytic bacteria (i.e. *Fibrobacter succinogenes*) are capable of producing cellotriose during growth on cellobiose via reversible cellodextrin phosphorylases (Wells et al. 1995; Zhang and Lynd 2005a) as verified by enzyme assay. While the 2.8-fold increase in Cthe__1020 expression and 2.6-fold decrease in Cthe__0391 expression in stationary phase was statistically significant ($V_{diff} > 1$), the other subunits of these transporters did not follow suit.

6.4.4.2 *Conversion of cellobiose to end-products*

6.4.4.2.1 Glycolysis

In *C. thermocellum*, conversion of glucose to phosphoenolpyruvate (PEP) occurs via the Embden-Meyerhoff-Parnas pathway (Figure 6.2a, Supplementary Table 3). All glycolytic proteins were detected in the top 20% (RAI > 0.83) of total proteins detected by 2D-HPLC-MS/MS, with a few exceptions. Glucose-6-P isomerase (Cthe__0217) had a RAI = 0.28, and one of the two encoded glucose kinases (Cthe__0390) was not detected. While glyceraldehyde-3-P dehydrogenase was the most highly expressed protein (RAI = 21.1) of all proteins detected, expression of subsequent proteins encoded in the predicted operon (Cthe__0137-0140) decreased respectively with increasing gene distance from glyceraldehydes-3-P dehydrogenase, suggesting transcriptional and/or post-transcriptional regulation of the operon. Protein expression profiles show that interconversion of fructose-1-P to fructose-1,6-bisphosphate can occur via pyrophosphate (PP_i)-dependent 6-P-fructokinase (RAI = 5.64), which was detected at higher levels than ATP-dependent 6-P-fructokinases Cthe__1261 and Cthe__0389 (RAI = 1.47 and 1.06, respectively). Of the two encoded fructose-1,6-P aldolases (Cthe__0349 and Cthe__2938), only Cthe__0349 was detected. While seven copies of putative phosphoglycerate mutase are encoded, Cthe__0140, which is encoded in a predicted operon containing glyceraldehydes-3-P dehydrogenase, phosphoglycerate kinase, and triosephosphate isomerise (Cthe__0137-0139) shows maximal expression throughout fermentations, consistent with mRNA expression profiles on cellulose (Raman et al. 2011). Expression of phosphoglycerate mutase Cthe__0946, Cthe__1292, and

Cthe__0707 were also detected, while Cthe__1435, Cthe__2449, and Cthe__3153 were not.

While the majority of glycolytic proteins did not change during transition to stationary phase, phosphoglycerate kinase and enolase increased by ~1.4-fold with a V_{diff} confidence score of >0.7, while phosphoglycerate mutase and triosephosphate isomerase increased by ~1.4-fold, but only with a V_{diff} confidence score of >0.2. While Raman et al. (2011) observed a decrease in mRNA expression of ATP-dependent phosphofructokinase Cthe__1261 and PP_i-dependent phosphofructokinase Cthe__0389 during transition to stationary phase, we did not observe any changes in protein levels. However, we did observe a decrease in phosphoglycerate mutase Cthe__0946 and an increase in Cthe__1292, consistent with cellulose grown *C. thermocellum* mRNA profiles (Raman et al. 2011).

6.4.4.2.2 Energy storage

Glycogen, an energy and carbon storage compound, is commonly synthesized during periods of slow or no growth, especially in carbon excess, and is often associated with sporulation (Preiss 1984; Preiss and Romeo 1989). Glucose-1-P adenylyltransferase (Cthe__3166 and Cthe__3167), involved in the synthesis of the primary glucosyl donor ADP-glucose, was detected in exponential phase cell extracts using shotgun 2D-HPLC-MS/MS (Figure 6.2b, Supplementary Table 3). Of the two genes encoding glycogen synthase (Cthe__1284 and Cthe__0282), which catalyzes α-1,4-glucosyl linkages to a pre-existing α-1,4-glucan, levels of Cthe__1284 were ~15-fold higher than that of Cthe__0282, suggesting it is the primary glycogen synthase in *C. thermocellum*. While

the level of 1,4-α-glucan branching enzyme, required for catalyzing α-1,6-glucosyl linkages, was below our threshold cutoff in shotgun analysis, it was detected in 4-plex analysis. A putative 1,4-α-glycogen debranching enzyme and α-glucan phosphorylase, required for glycogen breakdown, was also detected in exponential phase cultures. On the basis of simultaneous glucose-1-P adenylyltransferase, glycogen synthase, and glycogen phosphorylase activities in *C. cellulolyticum* cell extracts, Guedon *et al*. have proposed that glycogen synthesis and glycogenolysis can occur simultaneously (Guedon et al. 2000a). However, the effect of allosteric regulators on these enzymes was not studied in *C. cellulolyticum*.

We observed a 3.5-fold increase in glycogen synthase Cthe__0282 and a 2.5-fold increase in 1,4-α-branching enzyme in stationary phase, suggesting that glycogen synthesis is favoured during stationary phase. While glucose-1-P adenylyltransferase expression did not change, its activity has been shown to be allosterically activated via glycolytic intermediates and inhibited via AMP, ADP, P_i, and PP_i (Preiss 1984; Preiss and Romeo 1989). While enzyme assays show that levels of glucose-1-P adenelylytransferase and glycogen synthase increase with decreasing growth rate during transition to stationary phase in most organisms (Preiss 1984), catalytic activities of these enzymes, as well as α-glucan phosphorylase activity, increased with higher growth rates in *C. cellulolyticum* (Guedon et al. 2000a). Furthermore, in contrast to many bacterial species, which produce glycogen during the onset of stationary phase, glycogen synthesis reached a maximum in exponential phase and was utilized during transition to stationary phase in batch *C. cellulolyticum* cultures (Guedon et al. 2000a). Interestingly, expression

of α-glucan phosphorylase also increased 2.5-fold, which may help the cell utilize glycogen in the absence of an external carbon source.

6.4.4.2.3 Pentose phosphate pathway

The oxidative branch of the pentose phosphate pathway (PPP) generates reducing equivalents (NADPH) for biosynthesis, whereas the non-oxidative branch produces key intermediates, namely ribose-5-P and erythrose-4-P, required for the synthesis of nucleotides and aromatic amino acids, respectively. The absence of genes encoding glucose-6-P dehydrogenase, gluconolactonase, and 6-P-gluconate dehydrogenase of the oxidative PPP branch suggests that an alternative NADPH generation system must exist and that glycolytic intermediates (glyceraldehydes-3-phosphate and fructose-6-phosphate) must feed the non-oxidative branch of the PPP (Figure 6.2c, Supplementary Table 3). Furthermore, homology-based annotation suggests that the non-oxidative branch of the PPP is incomplete. While *C. thermocellum* encodes ribulose-5-P isomerase, ribulose-5-P epimerase, and two transketolases (Cthe__2443-2444 and Cthe__2704-2705), no gene encoding a transaldolase has been identified. 2D-HPLC-MS/MS expression profiles reveal that transketolase Cthe__2704-2705 is highly expressed throughout fermentation (RAI ~ 0.7), while Cthe__2443 is not detected and Cthe__2444 is found only in low amounts (RAI = 0.09). While ribose-5-P isomerase was detected (RAI = 0.37), ribose-5-P epimerase was not. Given the absence of transaldolase, ribose-5-phosphate must be synthesized using an alternative pathway.

A novel mechanism of non-oxidative hexose-to-pentose conversion that does not require transaldolase has been demonstrated in *Entamoeba histolytica* and other parasitic

protists (Susskind et al. 1982; Mertens 1993; Mertens et al. 1993). This system employs transketolase, aldolase, and PP_i-dependent 6-phosphofructokinase (Figure 6.2c). Susskind *et al*. have shown that fructose-1,6-bisphosphate aldolase, which typically converts glyceraldehyde-3-P and dihydroxyacetone-P into fructose-1,6-bisphosphate, is capable of converting dihyroxyacetone-P and erythrose-4-P into sedoheptulose-1,7-bisphosphate (Susskind et al. 1982). PP_i-dependent phosphofructokinase, which commonly catalyzes the reversible interconversion of fructose-6-P to fructose-1,6-bisphosphate, can then produce sedoheptulose-7-bisphosphate from sedoheptulose-1,7-bisphosphate. Finally sedoheptulose-7-bisphosphate and glyceraldehyde-3-P can be converted to ribose-5-P and xylose-5-P using transketolase again. While enzyme assays have not been carried out to determine the substrate specificity of fructose-1,6-bisphosphate aldolase and PP_i-dependent 6-phosphofructokinase in *C. thermocellum*, it is tempting to propose a similar hexose-to-pentose conversion mechanism.

6.4.4.2.4 Pyruvate formation from phosphoenolpyruvate

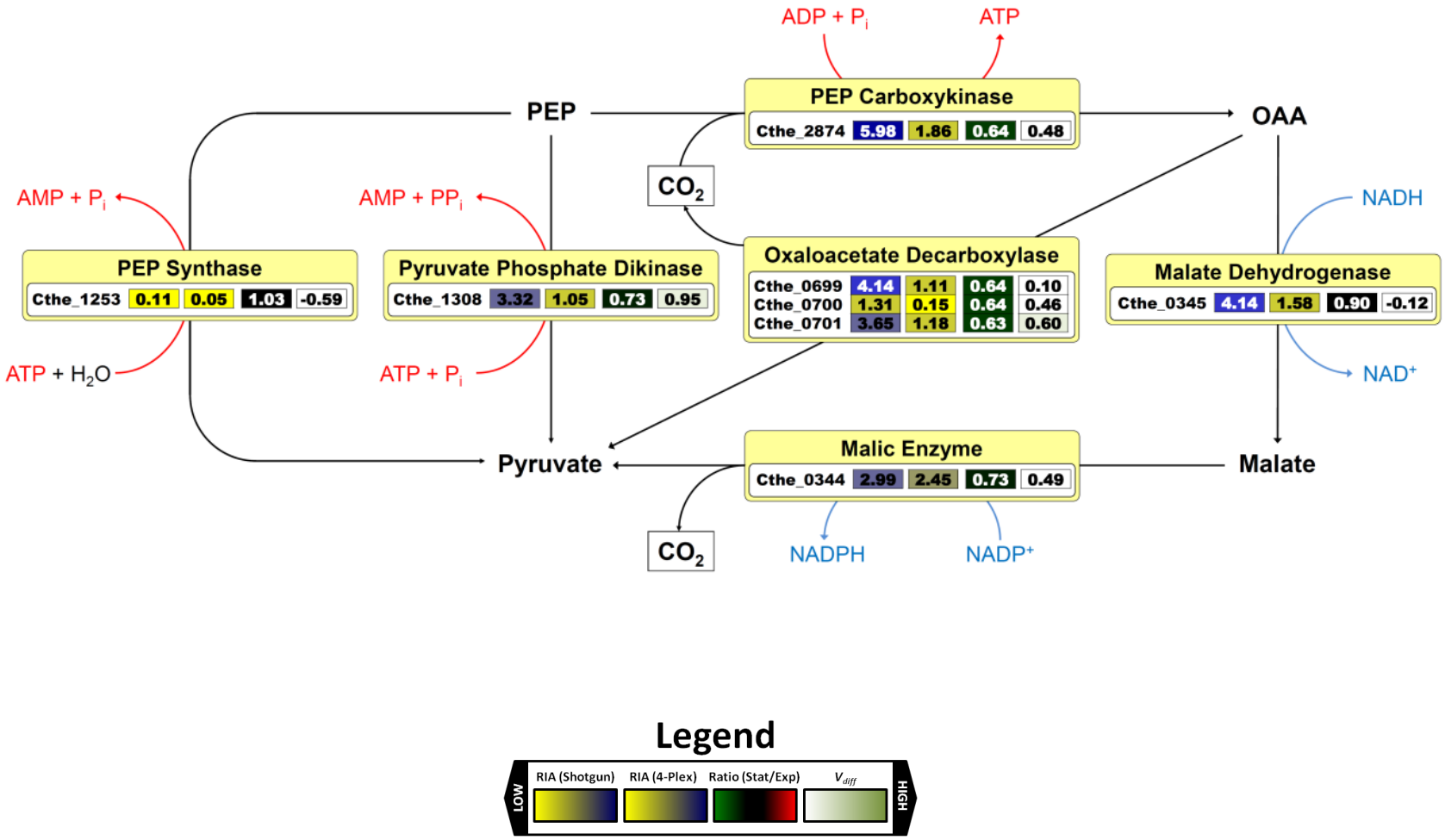
While most organisms convert phosphoenolpyruvate (PEP) to pyruvate via pyruvate kinase, producing ATP from ADP (Sparling et al. 2012), sequence homology-based annotation has not revealed the presence of a pyruvate kinase in *C. thermocellum*. However, several alternative proteins are expressed that may result in a tightly regulated pathway node (Figure 6.3a, Supplementary Table 3) leading to pyruvate synthesis. Phosphoenolpyruvate can be reversibly converted to pyruvate via pyruvate phosphate dikinase (PPDK), producing ATP and P_i from AMP, and PP_i, or using PEP synthase (PEPS) which produces ATP and H_2O from AMP, and P_i. While PPDK was expressed at

high levels in exponential phase, PEPS was not (RAI = 3.32 *vs* 0.11). Alternatively, PEP carboxykinase (PEPCK), which was also highly expressed (RAI = 5.98), can convert PEP to oxaloacetate while generating ATP. Oxaloacetate can subsequently be converted either directly to pyruvate via oxaloacetate decarboxylase (OAADC), or indirectly through malate via malate dehydrogenase (MDH) and malic enzyme (MalE), all of which were also highly expressed. High NADH-dependent MDH and $NADP^+$-dependent MalE activities (Rydzak *et al*., *unpublished*) suggest that MDH/MalE facilitate transhydrogenation from NADH to $NADP^+$, resulting in NADPH for biosynthesis, or potential H_2 or ethanol synthesis (Rydzak et al. 2011). Interestingly, all the enzymes in this node, with the exception of PEPS and MDH, decrease ~1.4 to 1.6-fold during stationary phase, generally consistent with reported mRNA profiles of cellulose grown cells (Raman et al. 2011). Regulation of carbon flux through this node cannot be simply attributed to changes in protein expression level since MalE has been shown to be regulated allosterically. Ammonia has been reported as an activator of MalE in *C. thermocellum*, and thus, transhydrogenation of NADH to $NADP^+$ via MDH and MalE is only allowed when sufficient NH_4^+ is present for biosynthesis (Lamed and Zeikus 1981). More recently, PP_i inhibition of MalE has been demonstrated (Taillefer and Sparling, *unpublished*). While this may be counterintuitive given that high levels of PP_i are present in the cell during rapid growth and biosynthesis, which in turn increases the demand for NADPH, the regulatory aspects with MDH and MalE are tightly knit with PPDK, which either uses PP_i during glycolysis, allowing for NADPH formation using MDH and MalE,

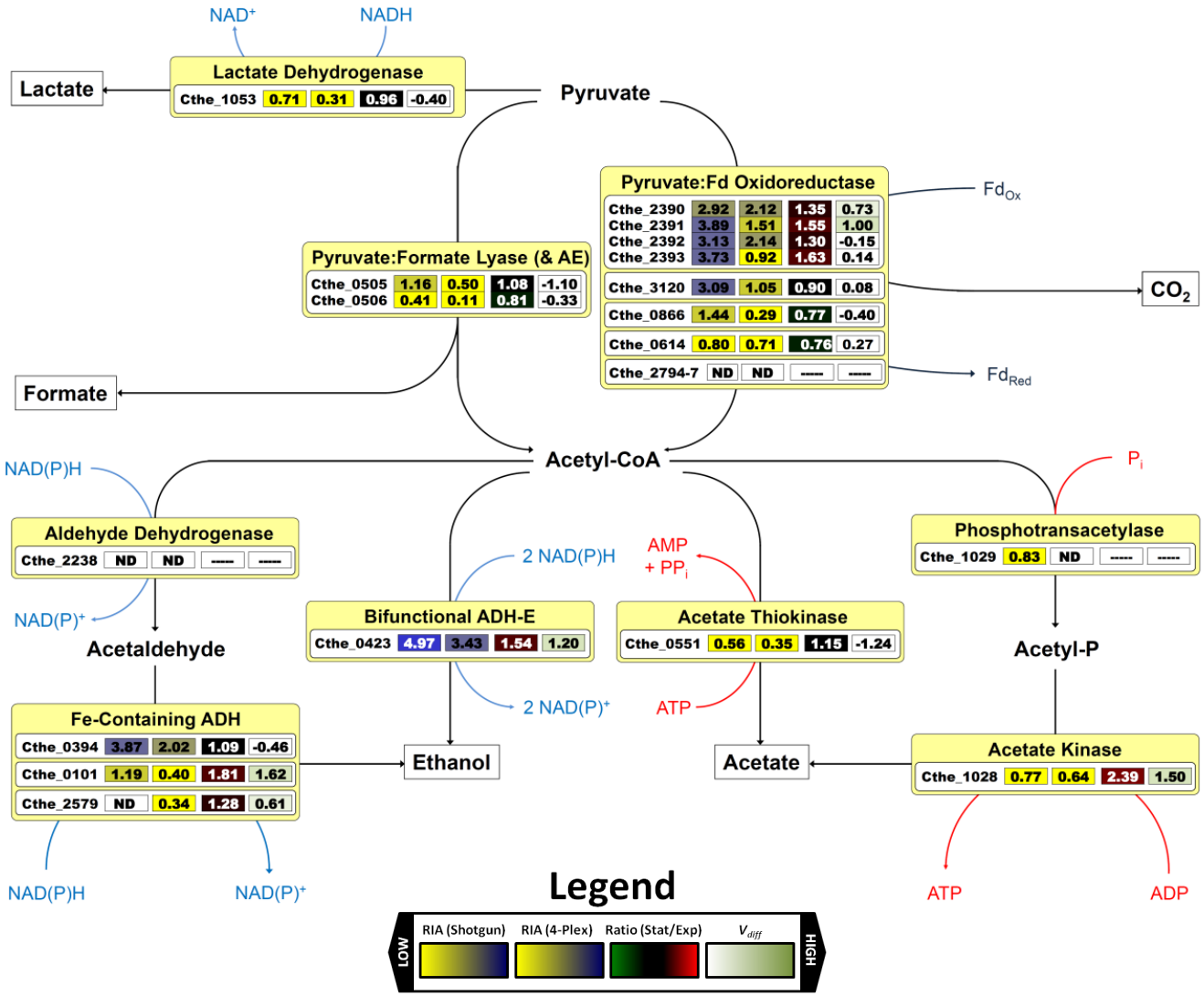
or produces PP_i during carbon starvation and gluconeogenesis, inhibiting the MDH/MalE pathway accordingly to the cell's NADPH demand.

Figure 6.3: Relative abundance indexes and changes in protein expression levels of proteins involved in conversion of phosphoenolpyruvate to end-products. Relative abundance indexes (values 1 and 2), changes in protein expression ratios (value 3), and associated V_{diff} values (value 4) indicating confidence levels of changes in expression ratios for enzymes involved in (A) conversion of phosphoenolpyruvate to pyruvate (B) catabolism of pyruvate into end-products, and (C) electron transfer pathways between ferredoxin (Fd), NAD(P)H, and H_2. PEP, phosphoenol pyruvate; OAA, oxaloacetate; Fd, ferredoxin; ND, not detected.

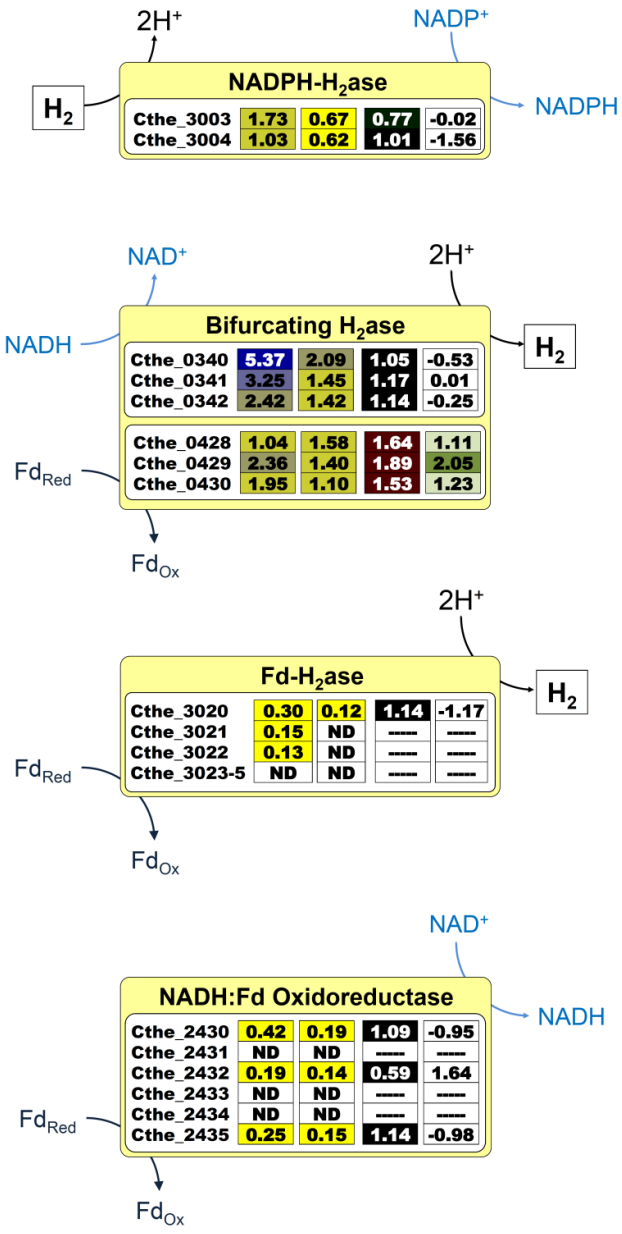
A)


ADP + Pi
ATP
PEP
PEP Carboxykinase
Cthe_2874 5.98 1.86 0.64 0.48
OAA
CO2
AMP + Pi
AMP + PPi
NADH
PEP Synthase
Cthe_1253 0.11 0.05 1.03 -0.59
Pyruvate Phosphate Dikinase
Cthe_1308 3.32 1.05 0.73 0.95
Oxaloacetate Decarboxylase
Cthe_0699 4.14 1.11 0.64 0.10
Cthe_0700 1.31 0.15 0.64 0.46
Cthe_0701 3.65 1.18 0.63 0.60
Malate Dehydrogenase
Cthe_0345 4.14 1.58 0.90 -0.12
ATP + H2O
ATP + Pi
NAD+
Malic Enzyme
Cthe_0344 2.99 2.45 0.73 0.49
Pyruvate
Malate
CO2
NADPH
NADP+
Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
Vdiff
HIGH

B)


NAD+
NADH
Lactate
Lactate Dehydrogenase
Cthe_1053 0.71 0.31 0.96 -0.40
Pyruvate
Pyruvate:Fd Oxidoreductase
Cthe_2390 2.92 2.12 1.35 0.73
Cthe_2391 3.89 1.51 1.55 1.00
Cthe_2392 3.13 2.14 1.30 -0.15
Cthe_2393 3.73 0.92 1.63 0.14
Fd_Ox
Pyruvate:Formate Lyase (& AE)
Cthe_0505 1.16 0.50 1.08 -1.10
Cthe_0506 0.41 0.11 0.81 -0.33
Cthe_3120 3.09 1.05 0.90 0.08
CO2
Cthe_0866 1.44 0.29 0.77 -0.40
Cthe_0614 0.80 0.71 0.76 0.27
Cthe_2794-7 ND ND ----- -----
Fd_Red
Formate
Acetyl-CoA
NAD(P)H
Pi
Aldehyde Dehydrogenase
Cthe_2238 ND ND ----- -----
2 NAD(P)H
AMP + PPi
Phosphotransacetylase
Cthe_1029 0.83 ND ----- -----
NAD(P)+
Acetaldehyde
Bifunctional ADH-E
Cthe_0423 4.97 3.43 1.54 1.20
Acetate Thiokinase
Cthe_0551 0.56 0.35 1.15 -1.24
Acetyl-P
2 NAD(P)+
ATP
Fe-Containing ADH
Cthe_0394 3.87 2.02 1.09 -0.46
Cthe_0101 1.19 0.40 1.81 1.62
Cthe_2579 ND 0.34 1.28 0.61
Ethanol
Acetate
Acetate Kinase
Cthe_1028 0.77 0.64 2.39 1.50
NAD(P)H
NAD(P)+
ATP
ADP
Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
V_diff
HIGH

C)


2H+
NADP+
NADPH-H2ase
H2
Cthe_3003 1.73 0.67 0.77 -0.02
Cthe_3004 1.03 0.62 1.01 -1.56
NADPH
NAD+
2H+
Bifurcating H2ase
NADH
H2
Cthe_0340 5.37 2.09 1.05 -0.53
Cthe_0341 3.25 1.45 1.17 0.01
Cthe_0342 2.42 1.42 1.14 -0.25
Cthe_0428 1.04 1.58 1.64 1.11
Cthe_0429 2.36 1.40 1.89 2.05
Cthe_0430 1.95 1.10 1.53 1.23
FdRed
FdOx
2H+
Fd-H2ase
H2
Cthe_3020 0.30 0.12 1.14 -1.17
Cthe_3021 0.15 ND ----- -----
Cthe_3022 0.13 ND ----- -----
Cthe_3023-5 ND ND ----- -----
FdRed
FdOx
NAD+
NADH:Fd Oxidoreductase
NADH
Cthe_2430 0.42 0.19 1.09 -0.95
Cthe_2431 ND ND ----- -----
Cthe_2432 0.19 0.14 0.59 1.64
Cthe_2433 ND ND ----- -----
Cthe_2434 ND ND ----- -----
Cthe_2435 0.25 0.15 1.14 -0.98
FdRed
FdOx
Legend
LOW
RIA (Shotgun)
RIA (4-Plex)
Ratio (Stat/Exp)
Vdiff
HIGH

6.4.4.2.5 Pyruvate catabolism and end-product synthesis

Synthesis of organic end-products from pyruvate is mediated by enzymes comprising two major branchpoints, namely the *pyruvate/acetyl-CoA/lactate* branchpoint and the *acetyl-CoA/ethanol/acetate* branchpoint, while H_2 can be generated from reduced ferredoxin (Fd_r), NADH, or NADPH using multiple H_2ase complexes (Figure 6.3). While the functionality of these pathways has been verified using enzyme assays (Rydzak et al. 2009; Rydzak et al. 2011), and transcriptional expression of the genes involved in these pathways has recently been elucidated (Carere et al. 2008a; Raman et al. 2011; Riederer et al. 2011), there have been no reports regarding the expression levels of these genes at the protein level. Given that there are apparent redundancies in genes encoding enzymes with analogous functions (e.g. PORs, ADHs, H_2ases) according to the current annotation, it is important that protein abundances and their expression profiles under physiological conditions be determined for the effective application of metabolic engineering strategies to improve rates and/or yields of H_2, ethanol, and other desired end-products.

6.4.4.2.5.1 Pyruvate/acetyl-CoA/lactate branchpoint

C. thermocellum may convert pyruvate into (i) CO_2, Fd_r, and acetyl-CoA, (ii) formate and acetyl-CoA, and (iii) lactate via pyruvate:ferredoxin oxidoreductase (PFOR), pyruvate:formate lyase (PFL), and lactate dehydrogenase (LDH), respectively (Rydzak et al. 2009). Based on end-product profiles (Figure 6.1), carbon flux is preferentially channelled through PFOR. *C. thermocellum* encodes two 4-subunit PORs. While the γ,

δ, α, and β subunits encoded by the gene cluster Cthe__2390-2393 are highly expressed, proteins encoded by Cthe__2794-2797 are not detected by 2D-HPLC-MS/MS, in agreement with mRNA profiles reported by Raman *et al*. (Raman et al. 2011) and Fong *et al*. (Gowen and Fong 2010). This contrasted with RT-PCR experiments by Carere *et al*., which reported high expression of subunit Cthe__2796 and low expression of subunit Cthe__2392 during exponential phase cultures grown on cellulose (Carere et al. 2008a).

Three putative single subunit PFOR-like oxidoreductases, including Cthe__3120 putatively involved in methionine degradation, Cthe__0866, a putative 2-oxogluterate synthase, and Cthe__0614, a putative indolepyruvate:fd oxidoreuctase, were also detected at high levels using 2D-HPLC-MS/MS. In agreement with our relative protein abundance profiles, RT-PCR expression experiments have confirmed high expression levels of Cthe__3120 (Carere et al. 2008a). Given that BLAST analysis of Cthe__3120 revealed homology to a characterized pyruvate-dependent PFOR isolated from *C. acetobutylicum* (Meinecke et al. 1989), also found in a number of *Thermoanaerobacter* species, these oxidoreductases may also be capable of converting pyruvate into acetyl-CoA.

Formate production was consistent with the presence of PFL (Cthe__0505). While a number of studies have reported formate production (Stevenson and Weimer 2005; Islam et al. 2006; Sparling et al. 2006; Rydzak et al. 2009; Rydzak et al. 2011), others have not (Lamed and Zeikus 1980a; Meinecke et al. 1989; Strobel et al. 1995). These discrepancies may be a result of the use of different detection methods (gas chromatography *vs* high pressure liquid chromatography), fermentation conditions (batch

with no pH control *vs* bioreactor with pH control), or media composition (complex *vs* minimal). Expression levels of PFL were lower than that of PFOR Cthe__2390-2393, in agreement with end-product accumulation rates and previously reported enzyme activities (Rydzak et al. 2009). Of the four putative PFL-activating enzymes (Cthe__0506, Cthe__0647, Cthe__1167, Cthe__1578) required for glycyl radical formation on the C-terminal portion of PFL (Sawers and Bock 1988; Vey et al. 2008), only Cthe__0506 was detected. While this agreed with high mRNA levels in cellobiose (Carere et al. 2008a) and cellulose grown batch cultures (Raman et al. 2011), Raman *et al*. also reported high expression levels of Cthe__0647 during fermentation. While PFL and PFL-activating enzyme Cthe__0506 are encoded next to each other, the 3-fold difference in expression levels supports that they are transcribed independently (Asanuma et al. 2004).

While LDH was expressed, albeit at lower levels than detected PORs and PFL, lactate production was not detected under the conditions tested. In *C. thermocellum* LDH has been shown to be allosterically activated by fructose-1,6-bisphosphate (FDP), (Özkan 2004) while in *Caldicellulosiruptor saccharolyticus*, a close relative to *C. thermocellum*, LDH is activated by FDP and ATP, and inhibited by NAD^+ and PP_i (Willquist and van Niel 2010). While lactate production in *C. thermocellum* was observed in batch cultures under carbon excess (Islam et al. 2006) and low culture pH (Rydzak et al. *unpublished*), this may be due to high intracellular FDP, concentrations, high $NADH/NAD^+$ ratios, and/or high ATP/PP_i ratios during transition to stationary phase (Willquist and van Niel 2010), which may have not been reached under our growth conditions.

6.4.4.2.5.2 Acetyl-CoA/ethanol/acetate branchpoint

Catabolism of acetyl-CoA into ethanol and acetate plays an important role in NADH reoxidation and energy conservation, respectively. Acetyl-CoA can be converted into ethanol directly using a bi-functional acetaldehyde/alcohol dehydrogenase (Cthe__0423, *adhE*), or indirectly via an NADH-dependant aldehyde dehydrogenase (Cthe__2238) and a number of iron containing alcohol dehydrogenases (Cthe__0101, Cthe__0394, Cthe__2579). Expression of Cthe__2238 (*aldH*), Cthe__0394 (*adhY*), and Cthe__2579 (*adhZ*) has been confirmed by real-time PCR (Stevenson and Weimer 2005). Of these ADHs, AdhE was the most abundant ADH detected (Figure 6.3b). Low expression levels of AldH suggest that AdhE is the primary protein involved in acetyl-CoA reduction and ethanol formation. ADHs Cthe__0394, Cthe__0101, and Cthe__2579 were expressed at 78%, 24%, and 9% of the levels of AdhE, respectively, suggesting that they may also be involved in formation of ethanol from acetaldehyde, albeit at lower levels. Two other zinc-containing ADH GroES-like heat shock proteins, Cthe__0388 and Cthe__2445, were also detected, the former being more highly expressed (Supplementary Table 3). While crude cell extract enzyme activities have shown the presence of both NADH and NADPH-dependent ADH activities, sequence analysis could not verify the substrate specificities of these enzymes.

Acetyl-CoA can be converted into acetate directly via acetate thiokinase (ATK) or indirectly through an acetyl phosphate intermediate using contiguously encoded phosphotransacetylase (PTA) and acetate kinase (ACK), which has been purified and characterized (Lin et al. 1998b). Although both reactions generate ATP, ATK does so

using AMP and PP_i, whereas PTA and ACK use ADP and P_i. This in turn has an impact on the thermodynamics of each reaction. The free energy of acetate production using PTA and ACK is more thermodynamically favourable than using ATK ($\Delta G^{\circ\prime} = -4$ kJ mol^{-1} *vs* +9 kJ mol^{-1}), and thus PTA and ACK are proposed to favour acetate production from acetyl-CoA, while ATK favours acetyl-CoA production from acetate. While Raman *et al*. report low mRNA levels of *pta* and *ack* and higher levels of *atk* (Raman et al. 2011), 2D-HPLC-MS/MS showed that all three proteins were detected at comparable levels (Figure 6.3b). Expression of all three enzymes remained constant throughout fermentation.

6.4.4.2.5.3 H_2 generation pathways

The genome of *C. thermocellum* encodes four putative H_2ase, including an Ech-like Fd-dependent [NiFe]-H_2ase (Cthe__3019-3024) and 3 Fe-only H_2ase catalytic subunits (Cthe__0342, Cthe__0430, Cthe__3003). Transcription of all of these subunits has been confirmed using RT-PCR (Carere et al. 2008a). Enzyme assays have shown that NADPH-dependent H_2ase activity is 5 to 10-fold higher than Fd and NADH-dependant H_2ase activities (Rydzak et al. 2009; Rydzak et al. 2011). The presence of a gene similar to the NADPH-binding subunit of glutamate synthase (Cthe__3004) adjacent to Cthe__3003 suggests that it may form a dimer with Cthe__3003 capable of generating NADPH from H_2 (Calusinska et al. 2010). 2D-HPLC-MS/MS reveals that both subunits are highly expressed, while subunits comprising both Fd-dependent [NiFe]-H_2ase and RNF-like NADH:Fd oxidoreductase were detected in low amounts or not at all

(Figure 6.3c), consistent with enzyme activity profiles (Rydzak et al. 2009; Rydzak et al. 2011) and mRNA profiles (Raman et al. 2011).

This leads to the question of how reduced Fd, formed by PFOR, is reoxidized. Recently a heterotrimeric bifurcating H_2ase, which utilizes reduced Fd and NADH synergistically to overcome the thermodynamic barrier of NADH oxidation for H_2 production, has been purified and characterized in *Thermotoga maritima* (Schut and Adams 2009). Genomic organization of Fe-only H_2ases Cthe__0342 and Cthe__0430 suggests that they may form bifurcating heterotrimers with neighbouring Nuo-like gene products Cthe__0340/0341 and Cthe__0428/0429, respectively. Both Cthe__0340-0342 and Cthe__0428-0430 were detected in high amounts, providing a probable method for Fd reoxidation. The fact that these putative NADH-dependent hydrogenases are in fact bifurcating H_2ases may explain the low NADH-dependent H_2ase activities detected in cell extracts. In fact, Schut *et al*. claim that "the trimeric [FeFe]-type hydrogenase purified from *T. maritima* does not use either reduced ferredoxin or NADH as a sole electron donor" (Schut and Adams 2009). While NADH-dependent H_2ase activities would likely be higher in the presence of reduced Fd, bifurcating H_2ase activities have not been assayed in crude cell extracts. Ideally, purification of these H_2ases is required for validation of bifurcating activity.

Interestingly, the genomic organization of *C. thermocellum* H_2ase subunits and upstream regulatory elements (see below) of Cthe__0428-0430, Cthe__0340-0342, and Cthe__3019-3014 reveal high similarity to that of *Thermoanaerobacterum saccharolyticus* (a.k.a. *T. thermosaccharolyticus*) gene clusters *hfs*, *hyd*, *ech*,

respectively. While all three H_2ases were expressed in wild-type *T. saccharolyticus*, as demonstrated by real-time PCR, gene knockout studies revealed that: i) *hfs* was the primary H_2ase responsible for H_2 production as its deletion drastically decreased H_2 production; ii) *hyd* knockouts had no effect on H_2 yields in batch fermentations, but decreased total methyl viologen-dependent H_2ase activity compared to wild type cells; and iii) *ech* knockouts had no effect on H_2 production or methyl viologen-dependent H_2ase activity (Shaw et al. 2009). This demonstrates the importance of mutational studies to determine the physiological role of H_2ases.

6.4.4.2.6 Changes in expression of enzymes involved in pyruvate catabolism and end-product synthesis

The subtle decrease in formate production rate and inversion of acetate-to-ethanol ratio during transition from exponential to stationary phase are consistent with previous studies (Raman et al. 2011). Transition from early to late log phase in pH regulated batch cultures (Payot et al. 1999), decreasing pH in steady state continuous cultures (Desvaux et al. 2001c), and increasing dilution rates (Guedon et al. 2000a) have all resulted in a shift from acetate to lactate and/or ethanol production mediated by an increase in $NADH/NAD^+$ ratios in *C. cellulolyticum*. Similarly, pH controlled batch cultures of *Caldicellulosiruptor saccharolyticus* exhibited increased $NADH/NAD^+$ ratios as cells approached mid to late-log phase, which subsequently triggered lactate production thus rebalancing $NADH/NAD^+$ ratios in late log and stationary phase (Willquist and van Niel 2010). These changes were also accompanied by an increase in LDH and ADH activity, despite the absence of ethanol production. While these studies were performed under

carbon excess conditions resulting in prolonged growth and more pronounced changes in end-product ratios, parallels can be drawn with our carbon limited *C. thermocellum* studies.

The ~1.4-fold increase in PFOR Cthe__2390-2393 and constant expression of PFL during stationary phase (Figure 6.3c) may divert carbon and electron flux away from PFL explaining the decrease in formate production rates as cells enter stationary phase. The additional reduced Fd produced via PFOR must then be reoxidized using Fd-dependant or bifurcating H_2ases. Accordingly, expression of bifurcating H_2ase Cthe__0428-0430 increases >1.5-fold in stationary phase. While both bifurcating H_2ases (Cthe__0428-0430 and Cthe__0340-342) contain various upstream regulatory elements including phosphatases, kinases, and/or PAS/PAC sensors potentially capable of regulating transcription in response to H_2 levels or redox changes via a two-component regulatory system as in *Ralstonia eutropha* (Kleihues et al. 2000; Buhrke et al. 2004; Friedrich et al. 2005), only Cthe__0428-0430 expression changed under the conditions tested. Regulation of a NAD(H)-dependent Fe-only H_2ase containing an upstream histidine and serine/threonine protein kinase has also been reported in *Ta. tencongensis*, in which a fourfold decrease in NAD(H)-dependent H_2ase activity was accompanied by an increase in AldH and ADH activities in response to high H_2 partial pressures (Soboh et al. 2004).

Providing that NADH/NAD^+ ratios increase during transition from exponential to stationary phase as in *C. cellulolyticum* and *Ca. saccharolyticus*, the observed increase in select ADHs [AdhE (Cthe__0423), Cthe__0101, glutamyl reductase (Cthe__1863), and

groES (Cthe__0388)] during stationary phase may help *C. thermocellum* reoxidize NADH and concomitantly produce ethanol, which explains the observed inversion of acetate-to-ethanol ratio. A similar mechanism of increasing expression of select ADHs to dispose of reducing equivalents during growth and ethanol accumulation is employed by *Thermoanaerobacter* species (Pei et al. 2010). Surprisingly, we observed a 2.4-fold increase in acetate kinase expression in stationary phase despite having lower acetate to ethanol ratios. This differs from the mRNA expression profiles on cellulose reported by Raman *et al*. (Raman et al. 2011). However, 4-plex 2D-HPLC-MS/MS did not detect the presence of PTA required for production of acetyl-P, and thus changes in expression profiles of PTA in response to growth phase could not be determined.

6.4.4.3 ***Energy generation and pyrophosphate (PP_i) metabolism***

In addition to substrate level phosphorylation mediated by 1,3-phosphoglycerate kinase, pyruvate phosphate dikinase, phosphoenolpyruvate carboxykinase, acetate kinase, and acetate thiokinase (see above), ATP can also be generated using ATP synthase powered by a proton motive force (PMF). While two types of ATP synthases were detected, including the F-type (Cthe__2602-2609) commonly involved in ATP synthesis, and the V-type (Cthe__2262-2269) commonly involved in generation of proton gradients, overall expression of the latter was higher (Supplementary Table 3). Expression of both ATP synthases was generally consistent throughout growth. Given the low expression of both ech H_2ase and NADH:Fd oxidoreductase, which pump out H^+ and Na^+, respectively, across the cell membrane during oxidation of reduced Fd, PMF generation for ATP synthesis using these enzymes seems unlikely. While, PMF may be generated through

PP_i hydrolysis using a membrane bound proton-translocating pyrophosphatase (PPase), the directionality of this PPase is unknown, and may in fact use PMF for PP_i synthesis.

PP_i is a by-product of various endergonic biosynthetic reactions, including poly-nucleic acid synthesis from (deoxy)nucleotide triphosphates and activation of amino acids, carbohydrates, and fatty acids for protein, polysaccharide, and lipid synthesis (Bielen et al. 2010). Thus, the effective removal of PP_i improves the thermodynamic feasibility of these reactions. Concentrations as low as 2 mM PP_i have shown to inhibit growth of some bacteria (Blumenthal et al. 1967). In addition to serving as a central energy carrier, PP_i serves to regulate key enzymes in carbohydrate metabolism including LDH in *Ca. saccharolyticus* (Willquist and van Niel 2010), malic enzyme in *C. thermocellum* (Taillefer and Sparling, *unpublished*), ATP-dependent PFK in *T. maritima* (Ding et al. 2001), and PTA in *C. acidiurici* (Robinson and Sagers 1972).

As mentioned above, PP_i can be utilized in the glycolytic direction by (i) PP_i-dependent 6-P-fructokinase, (ii) PPDK, and (iii) acetate thiokinase. Alternatively, hydrolysis of PP_i via a membrane-bound PPase (Cthe__1425) can be coupled to PMF generation that could be utilized for transport of nutrients, motility, and ATP synthesis. The PP_i-dependent enzymes used by *C. thermocellum* have remarkable similarities to that of parasitic protists (i.e. *Trichomonas foetus*, *Entamoeba histolytica*; (Mertens 1993)) and other bacteria such as *Ca. saccharolyticus* (Willquist et al. 2010). PP_i levels in *Ca. saccharolyticus* have been shown to be elevated (4 ± 2 mM) during exponential phase and lower during transition to stationary phase (Willquist et al. 2010), consistent with other organisms that do not contain a cystolic PPase (*C. thermoaceticum* and *C.*

pasteuranum; (Heinonen and Drake 1988)). Conversely, PP_i levels in *E. coli*, which possesses a cystolic PPase, were low (0.3 mM) and did not fluctuate during growth (Heinonen and Drake 1988). We observed a 1.9-fold increase in membrane-bound PPase expression in stationary phase cells.

6.5 Conclusions

A unified understanding of how gene and gene-product expression, stability, and regulation, in conjunction with intracellular metabolic profiling and thermodynamics of product formation, are key elements for targeted metabolic engineering strategies and fermentation optimization for the economic feasibility of biofuels production via consolidated bioprocessing. *Clostridium thermocellum*, like many cellulolytic, fermentative, biofuel producing organisms, has multiple enzymes capable of catalyzing parallel reactions and branched product pathways. Understanding protein expression profiles may provide genetic engineering strategies targeted at redirecting carbon and electron flux for the optimization of end-product production. Furthermore, responses of protein expression in response to physiological conditions (i.e. change in pH, end-product accumulation, carbon limitation) are essential in optimization of growth parameters during fermentation. In this study we performed proteomic analysis of core metabolic proteins involved in hemicellulose degradation and conversion of cellobiose into end-products in order to determine relative expression profiles of key enzyme dictating these pathways, and their changes in expression during their transition from exponential and stationary phase under closed-batch cellobiose-limited conditions.

Using shotgun 2D-HPLC-MS/MS, we determined relative protein expression profiles based on peptide spectral counts. As expected, the majority of annotated proteins predicted to be involved in conversion of cellobiose to end-products had relative abundance indices in the top 20^{th} percentile of all proteins detected. We demonstrated that many cellulosomal proteins were constitutively expressed despite the absence of cellulose, and relative expression profiles of cellodextrin ABC transporters reflected growth substrate and were in agreement with previously published mRNA profiles (Raman et al. 2011). Using relative expression profiles, we identified which proteins and metabolic networks are likely to be utilized during conversion of cellobiose to end-products. While there are multiple proteins with the same putative function encoded in the genome (i.e. glucose kinases, ATP-dependent phosphofructokinases, phospoglycerate mutases, pyruvate:ferredoxin oxidoreductases, alcohol dehydrogenases, bifurcating H_2ases), many of these proteins were differentially expressed. Similarly, we observed differential expression of proteins capable of parallel reactions that can interconvert one metabolite into another while using different cofactors. For example, PP_i-dependent phosphofructokinase was detected at 3-fold higher levels than ATP-dependent phosphofructokinases, highlighting the importance of PP_i as an energy currency during glycolysis. Similarly, several pyruvate:ferredoxin oxidoreductases were detected at 3-fold higher levels than pyruvate:formate lyase, corresponding with high CO_2 and low formate production. Genome analyses propose the presence of an aldehyde dehydrogenase, multiple alcohol dehydrogenases, and a bifunctional aldehyde/alcohol dehydrogenase (AdhE). While proteomic analyses revealed high AdhE levels, aldehyde

dehydrogenase was not detected, suggesting that ethanol synthesis from acetyl-CoA occurs primarily via AdhE.

C. thermocellum encodes four putative H_2ases, including an Ech-like Fd-dependent [NiFe]-H_2ase and 3 Fe-only H_2ases. While all three Fe-only hydrogenases were detected at high amounts, subunits of the Fd-dependent hydrogenase, as well as NADH:ferredoxin oxidoreductase, were detected at low levels, or not at all. Although relative abundance profiles do not necessarily reflect the catabolic activity of proteins, their low levels/absence of detection suggests that they do not play a major role in electron-transfer reactions. Given the high flux through pyruvate:ferredoxin oxidoreductase, as demonstrated by high rates of CO_2 production, reduced ferredoxin must be re-oxidized in some way. Genomic organization of Fe-only hydrogenases Cthe__0342 and Cthe__0430 suggests that they may form bifurcating heterotrimers, capable of oxidizing NADH and ferredoxin simultaneously for H_2 production. High NADPH-dependent hydrogenase protein and activity levels, along with high malate dehydrogenase and malic enzyme levels, which perform transhydrogenation of NADH to $NADP^+$ during conversion of phosphoenolpyruvate to pyruvate, show the importance in NADPH with regards to H_2 synthesis. Thus, knowledge of protein production profiles can provide a more targeted approach when attempting to redirect metabolic pathways through genetic engineering.

To determine changes in protein expression during transition from exponential phase to stationary phase we performed 4-plex 2D-HPLC-MS/MS on duplicate exponential and stationary phase samples labelled with different isobaric tags for relative

and absolute quantification. Cells were grown in closed-batch cultures grown under cellobiose-limited conditions, and were thus subjected to decreasing pH, end-product accumulation, and carbon limitation during fermentation. While end-product formation paralleled growth, we observed a slight inversion in acetate-to-ethanol ratio as cells approached stationary phase. Of the 144 core metabolic proteins detected using 4-plex 2D-HPLC-MS/MS, 24% showed statistically significant differential expression during transition from exponential to stationary phase. However, only 15% of these differentially expressed proteins changed by a magnitude greater than 2. The majority of proteins involved in hemicellulose hydrolysis, cello-oligosaccharide ABC-dependent transport, glycolysis, and the pentose phosphate pathway did not change in response to growth phase on cellobiose. Several low abundance proteins involved in glycogen synthesis, including glycogen/starch synthase, 1,4-α-branching enzyme, and α-glucan phosphorylase increased by more than 2-fold in stationary phase, suggesting in increased in glycogen metabolism. While the majority of proteins involved in interconversion of phosphoenol pyruvate and pyruvate decrease in stationary phase, proteins involved in conversion of pyruvate to end-products increase in stationary phase. Changes in expression profiles during transition from exponential to stationary phase in proteins involved in pyruvate catabolism and end-product synthesis may help explain the subtle changes in end-product formation rates. The ~1.4-fold increase in pyruvate:ferredoxin oxidoreductase Cthe__2390-2393 may divert carbon and electron flux away from constitutively pyruvate:formate lyase explaining the decrease in formate production rates during stationary phase. In turn, increased levels of reduced ferredoxin may be

reoxidized by bifurcating hydrogenase Cthe__0428-0430, which increased by >1.5-fold in stationary phase. If NADH/NAD^+ ratios increase during transition from exponential to stationary phase in *C. thermocellum*, the increase in alcohol dehydrogenase expression may aid in NADH reoxidition via ethanol production over lactate production, coinciding with increased ethanol to acetate ratios and absence of lactate production.

While we must further examine the physiological stimuli dictating not only gene and protein expression, but intracellular metabolite levels that may regulate carbon and electron flux via allosteric regulation and thermodynamic efficiencies, we have shown that differential protein expression levels under the conditions tested can influence end-product synthesis. Combined knowledge of relative protein expression levels and their changes in response to physiological conditions may aid in targeted metabolic engineering strategies and optimization of fermentation condition for improvement of biofuels production.

6.6 Authors` contributions

Thomas Rydzak, John A. Wilkins, David B. Levin, Oleg V. Krokhin, and Richard Sparling conceived and designed the study. Thomas Rydzak performed growth studies, end-product analysis, processed samples for proteomic analysis, analyzed proteomic data, and drafted the manuscript with input from David B Levin and Richard Sparling. Oleg V. Krokhin, Peter D. McQueen, Ravi C. Dwivedi, and Demitri Shamshurin aided in sample processing for proteomic analysis. Peyman Ezzati and Oleg V. Krokhin performed MS runs. Vic Spicer performed statistical analysis on MS data.

6.7 Acknowledgements

This work was supported by funds provided by Genome Canada, the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944-04) and the BIOCAP Canada Foundation and Microbial Genomics for Biofuels and Co-Products from Biorefining Processes ($MGCB^2$).

Chapter 7: Redirection of Pyruvate Flux in *Clostridium thermocellum* through Manipulation of Substrate Oxidation States and Use of Enzyme Inhibitors[24]

7.1 Abstract

The effects of substrate oxidation state and enzyme inhibitors on growth, end-product synthesis, and enzyme expression were studied in *Clostridium thermocellum*. While pyruvate alone did not sustain growth, it was metabolized in the presence of cellobiose. In the presence of cellobiose, exogenous pyruvate (i) increased acetate, formate, CO_2, and H_2 by as much as 4.2, 2.2, 1.8 and 1.6-fold, respectively, (ii) triggered lactate production, and (iii) and marginally decreased ethanol production. Specific enzyme inhibitors redirected flux through the various fermentation branches. Inhibition of pyruvate:formate lyase using hypophosphite redirected carbon flux through pyruvate:ferredoxin oxidoreductase, increasing CO_2 and H_2, and triggered lactate production during exponential phase, but had little impact on acetate and ethanol production. In the presence of 10% carbon monoxide, a known hydrogenase inhibitor, H_2 production decreased 6.4-fold, and flux was directed away from CO_2 and acetate towards

[24] Contributing authors: Thomas Rydzak, Marina Grigoryan, Zack J. Cunningham, Oleg V. Krokhin, Peyman Ezzati, Nazim Cicek, David B. Levin, John A. Wilkins, Richard Sparling. *Applied Microbiology and Biotechnology* (Submitted).

ethanol and formate. Proteomic analyses, using multiple reaction monitoring-mass spectrometry (MRM), revealed only minor (<2-fold) changes in pyruvate:ferredoxin oxidoreductase and putative bifurcating hydrogenase subunit Cthe_0430 in response to pyruvate and hypophosphite, respectively. The majority of peptides belonging to subunits of ferredoxin-dependent (Ech) hydrogenase and NADH:ferredoxin oxidoreductase (NFOR) were below MRM detection limits under the conditions tested, suggesting that they do not play a major role in electron transfer reactions. We discuss the implications of enzyme expression profiles on carbon and electron flux in response to selected growth conditions, and propose how electron flux is channeled using various enzymes.

7.2 Introduction

The modulation of *C. thermocellum* end-product yields in response to fermentation conditions including (i) growth phase (Rydzak et al. 2009; Raman et al. 2011; Rydzak et al. 2012) (ii) substrate loading/dilution rates (Stevenson and Weimer 2005; Islam et al. 2006; Magnusson et al. 2009), (iii) substrate composition (Islam et al. 2006; Levin et al. 2006; Sparling et al. 2006), (iv) medium composition (Johnson et al. 1981; Holwerda et al. 2012), (v) H_2 partial pressures (Weimer and Zeikus 1977; Lamed and Zeikus 1980a; Lamed and Zeikus 1980b; Ben-Bassat et al. 1981; Ng et al. 1981; Freier et al. 1988; Lamed et al. 1988; Lovitt et al. 1988; Bothun et al. 2004; Soboh et al. 2004), and (vi) end-product loading (Rydzak et al. 2011) has been well documented. Observed changes in end-product yields in *C. thermocellum* and related organisms have been linked to changes in mRNA and/or protein expression (Soboh et al. 2004; Stevenson and Weimer

2005; Rydzak et al. 2009; Raman et al. 2011; Rydzak et al. 2012), allosteric regulation (Özkan 2004), and NADH/NAD^+ and ATP/ADP ratios (Guedon et al. 1999b; Payot et al. 1999; Guedon et al. 2000a; Guedon et al. 2000b; Desvaux et al. 2001c; Desvaux and Petitdemange 2001; Desvaux et al. 2001d).

A number of substrates with different oxidation states, including glycerol (-2) and pyruvate (+1), have been used to study the effect of substrate oxidation state on end-product yields and enzyme expression levels in different fermentative organisms. The addition of glycerol to glucose-grown *C. acetobutylicum* and *C. butyricum* was shown to increase NADH/NAD^+ ratios and redirected carbon and electron flux from acids to alcohols (Vasconcelos et al. 1994; Saint-Amans et al. 2001). Alternatively, the addition of pyruvate to glucose grown *C. acetobutylicum* has been shown to promote acetogenesis over ethanologenesis, and was accompanied by a decrease in acetoacetate decarboxylase activities, while acetate kinase and butyrate kinase activities were unchanged (Ballongue et al. 1986; Junelles et al. 1987).

Redirection of carbon and electron flux has also been demonstrated through the use of enzyme inhibitors. Given the inherent difficulty and low transformation efficiencies of *C. thermocellum* strain ATCC 27405 (Tyurin et al. 2004; Tyurin et al. 2005), protein inhibitors, including hypophosphite (HPP) and carbon monoxide (CO), may be used to mimic gene knockouts/knockdowns of PFL and H_2ases, respectively. Preliminary studies by Maness *et al.* demonstrated that HPP increased methyl viologen-dependent H_2ase activity and H_2 production in *C. thermocellum* (Maness and Thammannagowda 2008). Conversely, addition of CO to the headspace of *C. acetobutylicum* redirected carbon and

electrons away from H_2, CO_2, acetate, and butyrate towards ethanol and butanol (Kim et al. 1984; Datta and Zeikus 1985). While inhibition of H_2ases using CO in *C. thermocellum* has been shown to increase ethanol/acetate ratios as well as lactate production (Li et al. 2012a), H_2, CO_2, and formate concentrations were not reported and thus redistribution of flux through end-product synthesis pathways could not be determined and carbon and electron balances could not be calculated.

In order to gain a better understanding of flux distribution towards end products, *C. thermocellum* was challenged with pyruvate to assess how substrate oxidation state can affect end-product yields. The cells were also grown in the presence of enzyme inhibitors (HPP and CO) to mimic PFL and H_2ase knockouts/knockdowns, respectively. Manipulation of carbon and electron fluxes provided insights into the *in vivo* activities of putative enzymes involved in pyruvate metabolism, end-product formation, and electron-transfer reactions. Multiple reaction monitoring (MRM) mass spectrometry was used to confirm expression of these proteins and to determine changes in relative protein expression profiles on cellobiose grown cells supplemented with pyruvate or HPP versus cells grown on cellobiose alone.

7.3 Materials and methods

7.3.1 Organism, media, and experimental design

Clostridium thermocellum DSM 1237 (equivalent to ATCC 27405) was employed for all growth experiments. Culture maintenance, media composition and preparation, and growth experiments were performed as previously described[25] by Rydzak *et al*. (Rydzak et al. 2009) with the following exceptions. To determine substrate utilization, the following combinations of substrates were used: (i) 1.1 g L^{-1} (3.2 mM) cellobiose, (ii) pyruvate (40, 60, 80 mM) or (iii) cellobiose (1.1 g L^{-1}) with pyruvate (40, 60, 80 mM). While glycerol (40, 60, 80 mM) was also tested as a potential substrate with and without cellobiose, cells did not grow on glycerol alone and glycerol was not metabolized in the presence of cellobiose. To test the effect of protein inhibitors, cells were grown on (i) 2.2 g L^{-1} cellobiose in the presence of hypophosphite (3, 6, 9 mM) and (ii) 1.1 g L^{-1} cellobiose in the presence of carbon monoxide (10% headspace). All substrate experiments were performed under a N_2 atmosphere unless otherwise specified. Samples for end-product, cell biomass, and pH measurements were taken after entry into stationary phase (OD_{600} ~ 0.7), while samples for proteomic analysis were taken in exponential phase (OD_{600} ~ 0.33).

7.3.2 Cell growth, pH, and end-product analysis

Cell growth measurements, sample processing, pH measurement, product gas, protein, sugar, and end-product analyses, and biomass and gas calculations were

[25] Refer to section 2.3.1 for details

performed as previously described[26] (Rydzak et al. 2009). Data represents the average of a minimum of three biological replicates.

[26] Refer to section 2.3.2 for details

7.3.3 Preparation of cell-free extracts for MRM mass spectrometry

Exponential phase cultures (10.5 mL) grown on cellobiose (1.1 and 2.2 g L^{-1}), cellobiose (1.1 g L^{-1}) plus pyruvate (60 mM), or cellobiose (2.2 g L^{-1}) plus hypophosphite (9 mM) were processed as described[27] by Rydzak *et al*. (Rydzak et al. 2012). In short, cells were washed 3 times in 1×phosphate buffered saline (PBS) buffer, resuspended in lysis buffer, and sonicated 5 times for 15 seconds. Protein concentrations of lysate were determined using a Bicinchononic Acid (BCA) Protein Assay Kit (Pierce Biotechnology, Rockford, IL) as outlined by the manufacturer. Proteins were reduced with dithiothreitol and alkylated with iodoacetamide. Excess iodoacetamide was quenched with dithiothreitol. Peptides were digested in a 1:50 trypsin/protein ratio (Promega, Madison, WI), acidified with trifluoroacetic acid (TFA), lyophilized, and suspended in 0.1% TFA. Samples were desalted using a C18 X-Terra column (1 × 100 mm, 5 μm, 100 Å; Waters Corporation, Milford, MA, USA) with 0.1% TFA. Peptides were eluted with 50% acetonitrile. Desalted samples lyophilized and resuspended in 0.1% formic acid. Samples were spiked with a mixture of six synthetic peptides (50 fM; LGGGGGGDGSR (P1), LGGGGGGDFR (P2), LLGGGGDFR (P3), LLLGGDFR (P4), LLLLDFR (P5), LLLLLDFR (P6); BioSynthesis Inc, Lewsville, TX), customized by Krokhin and Spicer (Krokhin and Spicer 2009), for prediction of retention times of selected peptides and MS/MS signal normalization (see below). Two biological

[27] Refer to section 6.3.3 for details

replicates grown under each condition were used for MRM. A minimum of two technical replicates were run for each biological replicate.

7.3.4 High-performance liquid chromatography

Peptide samples (0.2 and/or 0.09 μg μL^{-1}) were separated using a splitless nanoflow Tempo LC 215 system (Eksigent, Dublin, CA, USA) with 10 μL sample injection via a 300 μm × 5 mm PepMap100 precolumn and a 100 μm× 200 mm analytical column packed with 5 μm Luna C18(2) (Phenomenex, Torrance, CA) using a 4000 QTRAP (AB Sciex, Foster City, CA). Both eluents A (2% acetonitrile in water) and B (98% acetonitrile) contained 0.1% formic acid as ion-pairing modifier. A 1.03% acetonitrile/min linear gradient (0-40% B) was used for peptide elution, providing a total 38.8 minute run time per sample.

7.3.5 Tandem mass spectrometry

MS/MS was performed on a 4000 QTRAP hybrid triple quadrupole/ion trap mass spectrometer equipped with a nano-electrospray ion source (AB Sciex, Foster City, CA) using the following settings: ion spray potential (3050 V for positive mode), nebulizer gas (40 psi), bath gas (0 psi), curtain gas (20 psi), and collision gas (high), capillary temperature (160^0C). The optimal collision energy for precursor ions was at 37, with a duration time 48 min and dwell time for each transition 10 msec. MRM transitions were acquired at "unit" resolution in both the Q1 and Q3 quadrupoles to maximize specificity.

7.3.6 Design of MRM method and retention time prediction using MRMR in-house developed software

MRM method development software using a minimum redundancy maximum relevance (MRMR) algorithm was used to select most-likely reaction peaks for a protein sequence to be monitored (Han and Higgs 2008; Lange et al. 2008). In short, *in silico* tryptic digests were performed to determine most probable transition ions detected based on amino acid sequences (i.e. those greater than 7 amino acids long and least likely to be subject to post-translational modifications). Peptide retention times (RT) were computed based on hydrophobicity index using SSRCalc (Formic Acid model) and mapped into actual RT based on a linear regression of the observed RT of the calibrating peptides P2-P5 (specified above) and their precisely known hydrophobic indices (Krokhin and Spicer 2009). Selected peptides are provided in Additional Tables 4 and 5.

7.3.7 Data processing.

The spectral database obtained was processed by MultiQuant™ Software (AB Sciex, Foster City, CA). Spectra of peptides that were not detected via the MultiQuant software were manually inspected to verify their presence/absence based on the following criteria: i) presence of a minimum of 3 y-ions per peptide and ii) a minimum of signal to noise ratio of 3:1 as outlined by Keshishian et al (2007). Peptides that did not adhere to these criteria were reported as not quantifiable (NQ). We further verified if this was due to interfering peptides or if the peaks were unaccounted for due to low abundance. Peptide retention times were verified by synthetic peptide standards (Krokhin and Spicer

2009) and previous shotgun 2D-HPLC-MS/MS *C. thermocellum* runs (Rydzak et al. 2012). Standard synthetic peptides of equal and known concentrations were used for peak area normalization. Integrated peak areas acquired from all MRM runs (a minimum of two technical replicates per each of the two biological replicates per condition, with the exception of Ech H_2ase subunits and NFOR subunits) were used for protein ratio calculations. Areas of a minimum of three different transition states per peptide were acquired for each protein subunit (Supplementary Tables 4 and 5). The calculated ratio of each area (cellobiose + pyruvate / cellobiose or cellobiose + hypophosphite / cellobiose) was standardized using the average area ratios of the 4 synthetic peptide standards (P2-P5) for each run. The final protein subunit ratios were calculated by averaging all standardized peptide area ratios for all replicates performed (Supplementary Tables 4 and 5).

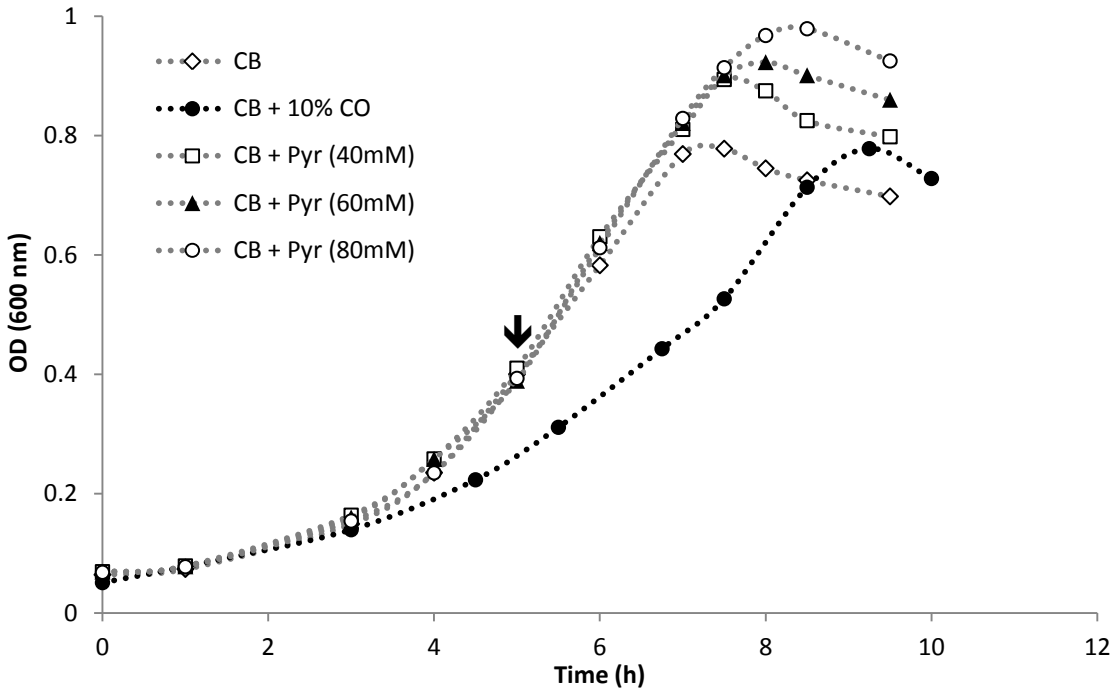
7.4 Results

7.4.1 Growth, substrate utilization, and end-product synthesis

Growth of *C. thermocellum* on carbon limited concentrations of cellobiose (1.1 g L^{-1}; Figure 7.1) was consistent with previous studies (Rydzak et al. 2009; Rydzak et al. 2011). Pyruvate or glycerol alone could not sustain growth and was not metabolized. We observed no changes in growth or final end-product concentrations when glycerol was added to cellobiose-grown cells suggesting that glycerol was not metabolized. Pyruvate, however, was metabolized in the presence of cellobiose. Addition of exogenous pyruvate to cellobiose-grown cells increased final biomass yields, and, to a

lesser extent, growth rates (Figure 7.1). Furthermore, a portion of the added pyruvate (~13% of initial pyruvate concentration added) was metabolized into end-products, corresponding to 31, 40, and 50% of total carbon recovered at the end of growth when cells were grown on 1.1 g L^{-1} cellobiose with 40, 60, and 80 mM pyruvate, respectively. The addition of various concentrations (3, 6, 9 mM) of HPP to cellobiose-grown cells had no significant effect on growth (data not shown). Addition of 10% CO (with 90% N_2) to cellobiose-grown cells increased generation time from 1.7 to 2.5 hours (Figure 7.1).

Figure 7.1: Effect of exogenous addition of pyruvate or CO on growth of *C. thermocellum* in the presence of cellobiose. Arrow indicates sampling point for MRM analysis. Addition of HPP to cell cultures had no effect on growth and is thus not shown.



Addition of pyruvate increased acetate, formate, CO_2, and H_2 by as much as 4.2, 2.2, 1.8 and 1.6-fold, respectively (at 80 mM pyruvate), triggered lactate production, and decreased ethanol production by 25% (Figure 7.2). Inhibition of PFL using HPP eliminated formate production and increased CO_2 production by 66%, regardless of HPP concentrations used. The increase in CO_2 production was proportional to the amount of formate formed in the absence HPP. The increase in flux through PFOR increased H_2 production by only 26%. Lactate production was also stimulated in the presence of HPP while ethanol and acetate production did not change significantly (Figure 7.3). While addition of 10% carbon monoxide decreased H_2 production by 6.4-fold, a small amount of H_2 (1.0 mM) was still produced, suggesting that H_2ase activity was not completely inhibited. In the presence of CO, CO_2 and acetate production decreased by 1.3 and 1.4-fold, respectively, while formate and ethanol production increases by 1.4 and 1.3-fold respectively (Figure 7.4). For all experiments, carbon balances were ~90%, and O/R ratios and C1/C2 ratios ranged from 0.9 to 1.1, verifying that all major end-products were accounted for.

Figure 7.2: Effect of co-metabolism of pyruvate and cellobiose versus metabolism of cellobiose alone on final end-product concentrations.

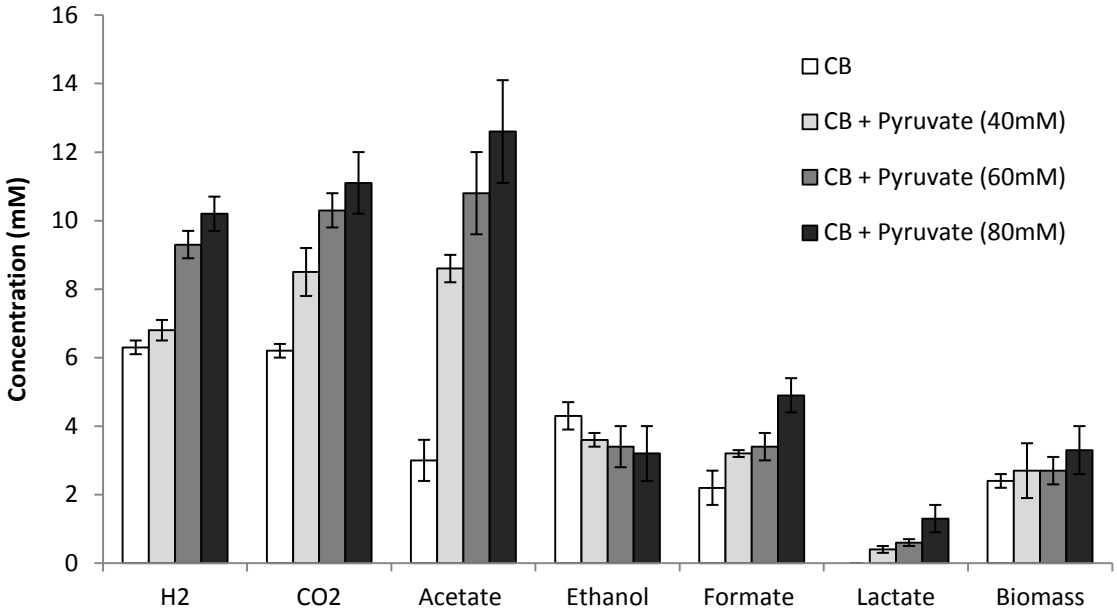


Figure 7.3: Effect of hypophosphite (HPP) on final end-product concentrations.

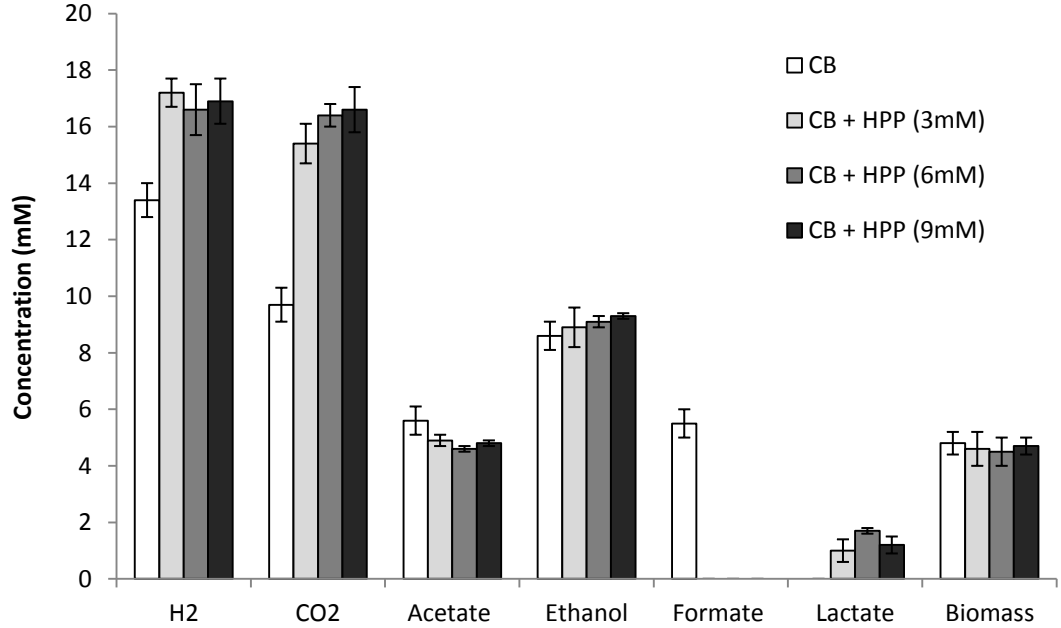
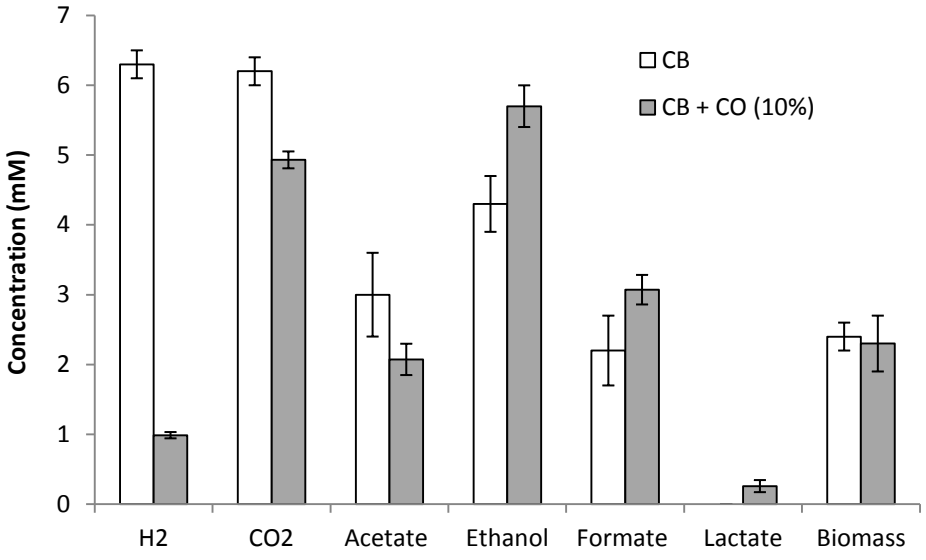


Figure 7.4: Effect of 10% carbon monoxide (CO) headspace on final end-product concentrations.



7.4.2 Detection of core metabolic proteins and changes in their expression in response to addition of exogenous pyruvate or HPP using MRM

We developed a MRM method to detect select peptides of proteins involved in glycolysis, pyruvate metabolism, and end-product synthesis (Supplementary Tables 4 and 5). All selected peptides were detected in high amounts and were used for comparative quantitation between conditions with the following exceptions: selected peptides belonging to Fd-dependent Ech-type H_2ase subunits Cthe_3021, 3023, and 3024, NFOR subunits Cthe_2431, 2433, and 2434 aldehyde dehydrogenase (Cthe_2238), and PFOR subunits Cthe_2794-2797 had poor signal to noise ratios (< 3:1) and did not generate a minimum of 3 y-ions, and were thus reported as not detected (Table 7.1, Supplementary

Tables 4 and 5). Interestingly, all Ech-type H_2ase subunits (Cthe_3020-3024) and NFOR subunits (Cthe_2430-2435) were not detected in MRM runs where other, high abundance, peptides were searched for due to high signal to noise ratios and overlapping peptide peaks. Therefore, we performed MRM runs in which only Ech-type H_2ase and NFOR peptides were monitored. While the majority of these selected peptides were below threshold values, several of the peptides belonging to subunits Cthe_2430, 2432, 2435, 3020, and 3022 were detected above threshold sensitivity (Table 7.1, Supplementary Tables 4 and 5). Peptides belonging to Cthe_2430 were only detected in cells grown on 1 g L^{-1} cellobiose (with and without pyruvate), but not 2 g L^{-1} cellobiose (with and without HPP).

Despite the significant changes in final end-product concentrations observed in response to exogenous pyruvate or HPP additions, expression levels of proteins involved in glycolysis, pyruvate metabolism, and end-product synthesis either did not change significantly or changed significantly by a factor of less than 2-fold (Table 7.1). Of note, PFOR (subunits Cthe_2390 and Cthe_2391), increased 1.9-fold, and phosphotransacetylase and acetate kinase increased 1.6-fold in response to exogenous pyruvate, consistent with increased CO_2 and acetate production. Furthermore, both selected subunits of oxaloacetate decarboxylase, Cthe_0699 and Cthe_0701 increased 1.5-fold whereas PEP carboxykinase decreased 1.4-fold in the presence of exogenous pyruvate. Bifurcating H_2ase subunit Cthe_0342 also increased 1.4-fold in the presence of exogenous pyruvate. No significant changes in enzyme expression profiles were observed in cells grown in the presence of HPP with the exception of a putative

bifurcating H_2ase subunit α (Cthe_0430), which increased 1.7-fold. However, the other subunits of this H_2ase (Cthe_0428-0429) did not follow this pattern.

Table 7.1: Changes in protein ratios between cells grown on cellobiose plus pyruvate versus cellobiose alone as determined by MRM mass spectrometry. Ratios (cellobiose plus pyruvate/cellobiose alone of cellobiose plus hypophosphite/cellobiose alone) represent average changes of a minimum of 2 representative peptides for each protein. Values of statistical significance are indicated in bold.

Locus Tag	Description	Average Peptide Area Ratios[1,2]			
		CB+Pyr/CB		CB+HPP/CB	
		Avg	Stdev	Avg	Stdev
Glycolysis					
Cthe_2938	Glucokinase	0.9	0.2	NA	-
Cthe_0217	glucose-6-isomerase	1.0	0.2	NA	-
Cthe_1261	6-phosphofructokinase (ATP-dependent)	1.7	0.8	NA	-
Cthe_0389	6-phosphofructokinase (ATP-dependent)	1.0	0.2	NA	-
Cthe_0347	6-phosphofructokinase (PPi-dependent)	0.9	0.1	NA	-
Cthe_0349	fructose-bisphosphate aldolase	0.7	0.1	NA	-
Cthe_0139	triosephosphate isomerase	1.0	0.2	NA	-
Cthe_0137	glyceraldehyde-3-phosphate dehydrogenase	1.0	0.3	NA	-
Cthe_0138	phosphoglycerol kinase	1.1	0.1	NA	-
Cthe_0140	Phosphoglycerate mutase	0.8	0.1	NA	-
Cthe_0143	Enolase	0.9	0.4	NA	-
Pyruvate Formation					
Cthe_1308	pyruvate phosphate dikinase	1.2	0.3	NA	-
Cthe_0699	oxaloacetate decarboxylase	**1.5**	0.4	NA	-
Cthe_0701	oxaloacetate decarboxylase	**1.5**	0.4	NA	-
Cthe_2874	PEP carboxykinase	0.7	0.2	NA	-
Cthe_0344	malic enzyme	0.9	0.2	1.0	0.2
Cthe_0345	malate dehydrogenase	NA	-	1.1	0.2

Pyruvate Catabolism					
Cthe_1053	lactate dehydrogenase	0.9[3]	0.1	1.2	0.4
Cthe_0505	pyruvate:formate lyase	1.4	0.4	1.0	0.2
Cthe_0506	pyruvate:formate lyase activating enzyme	NA	-	1.1	0.2
Cthe_2390	pyruvate:ferredoxin oxidoreductase (γ)	**1.9**	0.6	1.1	0.2
Cthe_2391	pyruvate:ferredoxin oxidoreductase (δ)	**1.9**	0.6	1.1	0.2
Cthe_2392	pyruvate:ferredoxin oxidoreductase (α)	NA	-	1.1	0.2
Cthe_2393	pyruvate:ferredoxin oxidoreductase (β)	NA	-	1.0	0.2
Cthe_2794	pyruvate:ferredoxin oxidoreductase (γ)	**ND**	-	**ND**	-
Cthe_2795	pyruvate:ferredoxin oxidoreductase (δ)	NA	-	NA	-
Cthe_2796	pyruvate:ferredoxin oxidoreductase (α)	**ND**	-	**ND**	-
Cthe_2797	pyruvate:ferredoxin oxidoreductase (β)	**ND**	-	**ND**	-
Cthe_0614	pyruvate:ferredoxin oxidoreductase	NA	-	1.3	0.3
Cthe_0866	pyruvate:ferredoxin oxidoreductase	1.1	0.3	1.0	0.2
Cthe_3120	pyruvate:ferredoxin oxidoreductase	1.0	0.5	1.4	0.4
Acetate Production					
Cthe_1028	acetate kinase	**1.6**	0.5	0.9	0.2
Cthe_1029	Phosphotransacetylase	**1.6**	0.5	1.0	0.1
Cthe_0551	acetate thiokinase	1.3	0.4	0.9	0.1
Ethanol Production					
Cthe_0423	bifunctionalaldehyde/alcohol dehydrogenase (AdhE)	1.3	0.3	1.1	0.2
Cthe_2238	aldehyde dehydrogenase	**ND**	-	**ND**	-
Cthe_0101	alcohol dehydrogenase	1.2	0.3	0.9	0.2
Cthe_0388	alcohol dehydrogenase	0.9	0.3	0.9	0.2
Cthe_0394	alcohol dehydrogenase	1.2	0.2	0.9	0.2
Hydrogen Production					
Cthe_0340	bifurcating hydrogenase (subunit γ)	1.3	0.3	1.0	0.1
Cthe_0341	bifurcating hydrogenase (subunit β)	1.1	0.4	1.2	0.4
Cthe_0342	bifurcating hydrogenase (subunit α)	**1.4**	0.3	1.0	0.1
Cthe_0428	bifurcating hydrogenase (subunit γ)	0.9	0.2	1.4	0.4
Cthe_0429	bifurcating hydrogenase (subunit β)	1.4	0.4	1.2	0.1

Cthe_0430	bifurcating hydrogenase (subunit α)	1.0	0.4	**1.7**	0.4
Cthe_3003	NADPH-dependent hydrogenase	1.0	0.2	1.1	0.2
Cthe_3004	NADPH-dependent hydrogenase	1.1	0.2	1.0	0.2
Cthe_3020	Fd-dependent Ech hydrogenase (subunit E)	1.2[3]	0.3	1.0	0.2
Cthe_3021	Fd-dependent Ech hydrogenase (subunit D)	**ND**	-	**ND**	-
Cthe_3022	Fd-dependent Ech hydrogenase (subunit C)	**1.7[3]**	0.2	1.1	0.2
Cthe_3023	Fd-dependent Ech hydrogenase (subunit B)	**ND**	-	**ND**	-
Cthe_3024	Fd-dependent Ech hydrogenase (subunit A)	**ND**	-	**ND**	-
Cthe_3025	Fd-dependent Ech hydrogenase (hypothetical)	NA	-	0.9	0.3
Cthe_0372	NfnA	NA	-	NA	-
Cthe_0373	NfnB	1.2[3]	0.3	**ND**	-
Cthe_2430	NADH:fd oxidoreductase (subunit C)	**ND**	-	**ND**	-
Cthe_2431	NADH:fd oxidoreductase (subunit D)	1.2[3]	0.3	1.0	0.1
Cthe_2432	NADH:fd oxidoreductase (subunit G)	**ND**	-	**ND**	-
Cthe_2433	NADH:fd oxidoreductase (subunit E)	**ND**	-	**ND**	-
Cthe_2434	NADH:fd oxidoreductase (subunit A)	1.0[3]	0.2	1.0	0.2
Cthe_2435	NADH:fd oxidoreductase (subunit B)	1.3	0.3	1.0	0.1

[1]NA: Not assesed
[2]ND: Not detected
[3]Only one biological replicate of cells grown in the presence of pyruvate plus cellobiose was used for ratio calculation

7.5 Discussion

7.5.1 Redistribution of metabolic fluxes through manipulation of growth parameters: A way of verifying *in vivo* protein activity

Understanding the elements that dictate distribution of carbon and electron flux towards end-products is crucial in optimization ethanol and/or H_2 yields. While elucidation and expression of metabolic pathways provides insights into how flux can be distributed, it does not necessarily indicate how flux is distributed *in vivo* under a particular growth condition. This is due to a number of elements that may dictate flux

through select pathways including enzyme kinetics, protein allostery, intracellular metabolite levels, and the cell's energy and redox requirements. A better understanding of how *in vivo* flux is regulated can be attained through the manipulation of growth conditions that lead to differences in measureable end-product yields. The modulation of *C. thermocellum* end-product yields in response to fermentation conditions has been well documented (Weimer and Zeikus 1977; Lamed and Zeikus 1980a; Lamed and Zeikus 1980b; Ben-Bassat et al. 1981; Johnson et al. 1981; Ng et al. 1981; Freier et al. 1988; Lamed et al. 1988; Lovitt et al. 1988; Bothun et al. 2004; Soboh et al. 2004; Stevenson and Weimer 2005; Islam et al. 2006; Levin et al. 2006; Sparling et al. 2006; Magnusson et al. 2009; Rydzak et al. 2009; Raman et al. 2011; Rydzak et al. 2011; Holwerda et al. 2012; Rydzak et al. 2012). However, most of these studies have not looked at the entire suite of end-products produced by *C. thermocellum*, and thus carbon and electron balances, or flux distribution through the branched end-product pathways, could not be elucidated. By determining how the entire suite of end-products is modulated in response to enzyme inhibitors, as well as the co-metabolism of pyruvate and cellobiose versus cellobiose alone, we have elucidated how flux is redistributed towards end-products and to what extent certain enzymes participate in balancing reducing equivalents during fermentation.

7.5.2 Co-metabolism of pyruvate and cellobiose promotes acetogenesis over ethanologenesis

To test the effects of metabolism of a more oxidized substrate, pyruvate was added to cellobiose-grown cells. Co-metabolism of pyruvate and cellobiose resulted in a

higher increase in total oxidized end-products (CO_2 + formate) than reduced end-products (H_2 + ethanol) when compared to growth on cellobiose alone (Figure 7.2, Figure 7.5). This was expected given that the oxidation value of pyruvate is +1. During growth on cellobiose, PFOR catabolizes pyruvate at 2.8 times the rate of PFL as demonstrated by CO_2 and formate product concentrations at the end of growth. However, the increase in formate was greater than the increase in CO_2 in the presence of pyruvate (2.2 vs 1.8-fold, respectively, at 80 mM pyruvate), suggesting that additional pyruvate preferentially fluxes through PFL rather than PFOR. Accordingly, the total molar ratio of CO_2 to formate was 1.2 fold lower during co-metabolism of 80 mM pyruvate and cellobiose compared to growth on cellobiose alone. The addition of pyruvate also triggered lactate production in exponential phase. Given that increasing *C. cellulolyticum* feeding rates were shown to increase lactate and extracellular pyruvate production (Guedon et al. 1999b; Desvaux et al. 2001d), it is likely that influx and buildup of intracellular pyruvate may be linked to increased lactate production.

Addition of pyruvate drastically affected metabolic flux at the acetyl-coA/acetate/ethanol branchpoint. Co-metabolism of pyruvate and cellobiose increased acetate production by up to 4.2-fold (at 80 mM pyruvate) when compared to growth on cellobiose alone, whereas overall ethanol production decreased by up to 25% (Figure 7.2). Accordingly, the addition of pyruvate to glucose grown *C. acetobutylicum* has been shown to promote acetogenesis over ethanologenesis (Ballongue et al. 1986; Junelles et al. 1987). The preferential increase in flux through acetate-producing pathways versus ethanol-producing pathways in *C. thermocellum* results in increased ATP production,

which is reflected by higher biomass yields. Furthermore, metabolism of additional pyruvate towards end-products does not generate additional NADH and thus does not require additional ethanol production. While MRM reveals no significant changes in alcohol dehydrogenases in response to addition of pyruvate, expression of both phosphotransacetylase and acetate kinase increased 1.6-fold during co-fermentation of pyruvate and cellobiose compared to growth on cellobiose alone (Table 7.1), consistent with an increase in acetate production. Similarly, the addition of pyruvate to glucose grown *C. acetobutylicum* elevated levels of acetate kinase and butyrate kinase levels compared to those of cells grown on glucose alone (Ballongue et al. 1986; Junelles et al. 1987).

Despite the oxidation state of pyruvate, an increase in H_2 production was also observed (1.6-fold at 80 mM pyruvate). This likely stems from the nearly proportional increase in flux through PFOR (1.8-fold), which in turn generates additional reducing equivalents in the form of reduced Fd. While the additional reduced Fd could be reoxidized via a Fd-dependent Ech H_2ase to produce H_2, the absence of detectable levels of most Ech H_2ase subunits using MRM suggests that it is not a major contributor to H_2 production. Alternatively, one of the two putative bifurcating H_2ases can be used to reoxidize Fd to produce H_2. Given that bifurcating H_2ases oxidize Fd and NADH in a 1:1 ratio (Wang et al. 2010), and that no additional NADH is produced through metabolism of pyruvate, it is not surprising to see a decrease in NADH-dependent ethanol production. However, based on the increase in H_2 production, we would expect to see an even greater decrease in ethanol production if H_2 was produced only via a bifurcating H_2ase. While

NFOR can potentially be used to replenish some NADH for ethanol production from reduced Fd, the majority of NFOR peptides were either not detected or detected at low levels using MRM. Alternatively, H_2 produced via the bifurcating H_2ase can be reoxidized to NADPH via a putative NADPH-dependent hydrogenase, which can in turn be used to produce ethanol as well. The implications of the absence of Fd-dependent H_2ase and RNF on electron flux are discussed below.

7.5.3 Hypophosphite redirects flux from PFL to PFOR and increases H_2 production

HPP is a formate analogue which can inhibit PFL by covalently binding its active site via a free radical dependent mechanism (Brush et al. 1988; Knappe and Sawers 1990). Addition of various concentrations of HPP to cellobiose grown cells ceased formate production and increased pyruvate flux through PFOR (Figure 7.2). The increase in CO_2 production (1.6-fold) in the presence of HPP was proportional to the amount of formate produced in the absence of HPP. The inhibition of PFL also triggered lactate production. The elimination of PFL activity in this pathway branch, responsible for catabolism of pyruvate, likely results in buildup of intracellular pyruvate pools. Accumulation of pyruvate and/or pyruvate leakage has in fact been linked to increased lactate production (Guedon et al. 1999b; Desvaux et al. 2000; Guedon et al. 2000a; Guedon et al. 2002).

The inhibition of PFL, and thus redirection of flux through PFOR, which generates additional reducing equivalents in the form of reduced Fd, has major implications on reduced end-product synthesis (Figure 7.3). In the presence of HPP, the

total amount of H_2 produced increased by 26% while the increase in ethanol production was not statistically significant when compared to growth in the absence of HPP. This increase in H_2 was consistent with previous reports by Mannes *et al*., who also showed an increase in methyl viologen-dependent H_2ase activity in response to HPP in *C. thermocellum* (Maness and Thammannagowda 2008). MRM revealed that increases in most H_2ase subunits in response to HPP were of little statistical significance with the exception of bifurcating H_2ase subunit Cthe_0430, which increased 1.7-fold. The inability to detect any Fd-dependent H_2ase subunits suggests that H_2 is produced using a one of two bifurcating H_2ases. The fact that a much greater increase in H_2 versus ethanol production was observed suggests that NFOR activity is negligible *in vivo*. This is also supported by the inability to detect NFOR using MRM. Given the thermodynamic favorability of NADH production from reduced Fd (Rydzak et al. 2009), we would expect an increase in ethanol production if NFOR was present and functional *in vivo*.

7.5.4 Carbon monoxide redirects flux from PFOR to PFL and increases ethanol production

Given that we were able to redirect flux towards H_2 production via inhibition of PFL using HPP, we attempted to redirect flux towards ethanol, rather than H_2, using CO. CO has been shown to inhibit H_2ases by binding to vacant [FeFe] or [NiFe] coordination sites (Hallahan et al. 1986; Fauque et al. 1988; De Lacey et al. 2002), and was therefore used to inhibit H_2 production. While H_2 production was not completely inhibited in the presence of 10% headspace CO, it did decrease by 6.4-fold (Figure 7.4). The decreased ability to reoxidize reduced Fd through H_2 production redirected pyruvate flux from

PFOR to PFL as demonstrated by the proportional decrease in CO_2 and increase in formate production. This was also accompanied by a decrease in acetate production proportional to the increase in ethanol production, as well as stimulation of lactate production. The increase in ethanol/acetate ratios and lactate production we observed was consistent with previous *C. thermocellum* studies, which did not report H_2, CO_2, or formate yields (Li et al. 2012a). Similarly, in *C. acetobutylicum*, addition of CO to the headspace redirected carbon and electrons away from H_2, CO_2, acetate, and butyrate towards ethanol and butanol (Kim et al. 1984; Datta and Zeikus 1985). Interestingly, in our studies, the observed decrease in CO_2 production, and thus reduced Fd production, was also proportional to the increase in ethanol production. However, the decrease in H_2 production (which requires 2 electrons per H_2 from reduced ferredoxin and/or NAD(P)H) was >2-fold greater that the decrease in CO_2 production or the increase in ethanol production (which requires 4 electrons from NAD(P)H), suggesting that a bifurcating H_2ase and/or NAD(P)H-dependent system is used for H_2 production. Otherwise, given that H_2 was produced strictly from reduced ferredoxin, we would expect to see a decrease in H_2 production that was proportional to the decrease in CO_2 (and reduced Fd). Instead we see that the change in CO_2 (and reduced Fd) accounts for less than half of the change in H_2. The fact that the decrease in H_2 production was greater than the increase in ethanol production can be explained by the observed increase in lactate production, which siphons off electrons in the form of NADH.

7.5.5 Proposed electron flux pathways based on end-product synthesis patterns and protein expression levels

Knowledge of expressed metabolic pathways, in conjunction with measured changes in end-product synthesis in response to substrates with (i) different oxidation states (i.e. pyruvate) and (ii) enzyme inhibitors (HPP and CO) can be used to model potential carbon and electron flux pathways involved in conversion of substrate into end-products. While flux through many of the end-product pathways can be evaluated by measuring concentrations of end products, determining electron flux through the various enzymes that can transfer electrons between NADH, NADPH, ferredoxin, and H_2 is more challenging. While a traditional understanding of these pathways has highlighted the importance of NFOR and Fd-dependent H_2ases in transferring electrons between electron carriers and H_2, MRM has revealed that these enzymes are found in low amounts when compared to putative bifurcating hydrogenases and NADPH-dependent hydrogenases. To further complicate deduction of electron-flow pathways, a recently characterized *Clostridium kluyveri* NADH-dependent reduced ferredoxin:$NADP^+$ oxidoreductase (NfnAB) complex, which simultaneously uses a reduced Fd and NADH to reduce 2 $NADP^+$ (Wang et al. 2010), was also encoded by (Cthe_0372-0373) and expressed in *C. thermocellum* (Rydzak et al. 2012)

While the inability to detecte most NFOR and the Fd-dependent Ech H_2ase subunits in *C. thermocellum* using MRM does not ascertain that they are in fact not expressed or functional *in vivo*, previous studies have also supported that these enzymes and/or their subunits are expressed at low levels or not detected at all. Although Fd-dependent H_2ase activities were detected by Rydzak *et al*., they were roughly 5-10 fold lower than NADPH-dependent H_2ase activities (Rydzak et al. 2009; Rydzak et al. 2011).

While these studies also reported low NADH-dependent H_2ase activities that were comparable to Fd-dependent H_2ase activities, bifurcating H_2ase activities have not been assessed in crude cell extracts. Microarray studies performed by Raman *et al*. demonstrated that transcription of Ech H_2ase subunits were low when compared to expression levels of other genes encoding catabolic pathways including bifurcating and NADPH-dependent H_2ases (Raman et al. 2011). Furthermore, their expression levels decreased as growth approached stationary phase. Similarly, shotgun proteomic analysis demonstrated that Ech H_2ase subunits Cthe_3020-3022 were at very low amounts, whereas subunits Cthe_3023 and 3024 were not detected at all (Rydzak et al. 2011). This was consistent with our MRM results, with the exception that Cthe_3021 was not detected using MRM. Interestingly, deletion of [FeFe]-H_2ase maturase gene *hydG* eliminated H_2 production in *C. thermocellum*, suggesting that the Fd-dependent [NiFe]-H_2ase, which uses different maturases, does not play a significant role in H_2 production (Guss and Lynd 2012).

While Lamed and Zeikus reported high NFOR activities in *C. thermocellum* LQRI and AS39 were more than 6-fold lower than that of *Thermoanaerobacter brocii* (Lamed and Zeikus 1980a), their assay conditions could not ensure that this activity was strictly derived from the presence of NFOR alone or from H_2ases that can use NAD(P)H as a cofactor[28]. Both microarray and shotgun proteomic studies have shown that expression of NFOR subunits Cthe_2430, 2432, and 2435 was low, whereas expression of subunits Cthe_2431, 2433, and 2434 was not observed at either the mRNA or protein level. Again, this was consistent with our MRM results. While the latter three subunits are typically embedded in the membrane, other membrane embedded proteins were detected using previous shotgun proteomic studies (Rydzak et al. 2011), demonstrating that lack of detection of these subunits is not necessarily due to membrane association. The absence of these subunits together with the identifiable, but low expression of the

---

[28] NFOR activities were measured in both directions using the following assay conditions: Reduction of Fd by NADH was measured by monitoring H_2 production in the presence of an NADH regenerating system (galactose, galactose dehydrogenase, NADH), acetyl-coA regeneration system (required for NFOR activation; acetyl phosphate, PTA, CoA), and Fd. However, given the presence of NADH-dependent H_2ase activity, H_2 can be produced directly from NADH, negating the need for NFOR mediated Fd reduction and subsequent Fd-dependent H_2ase mediated H_2 production. Reduction of NADH by Fd_{red} was measured by monitoring H_2 consumption in the presence of a NAD^+ regenerating system (pyruvate, LDH, NAD^+), Fd_{red} regenerating system (H_2, Fd, H_2ase(s) of the lysate). However, again, given the presence of NADH-dependent H_2ase activity, H_2 can be consumed directly using this H_2ase rather than Fd-dependent H_2ase, thus negating the need for NFOR-dependent oxidation of Fd_{red}.

cytoplasmic subunits, suggests that the transfer of electrons from reduced ferredoxin to NADH via NFOR may not necessarily be coupled to Na^+ and/or H^+ export in *C. thermocellum*.

We were initially surprised that the more targeted MRM approach did not detect NFOR and Ech H_2ase peptides that were previously detected using shotgun 2D-HPLC-MS/MS. However, this likely stems from one of two reasons. First and foremost, samples were fractionated using HPLC in two dimensions prior to loading each fraction into the mass spectrometer during shotgun proteomic experiments, whereas protein samples were separated only in one dimension during MRM. This caused more overlapping peptide peaks due to a more complex peptide composition in each sample, making peptide identification and quantitation more difficult. Second, the concentrations of less complex peptide samples loaded into the mass spectrometer were higher (~6.6 μg/fraction) during shotgun proteomics versus MRM proteomics (≤4.0 μg/fraction) which decreased detection sensitivity.

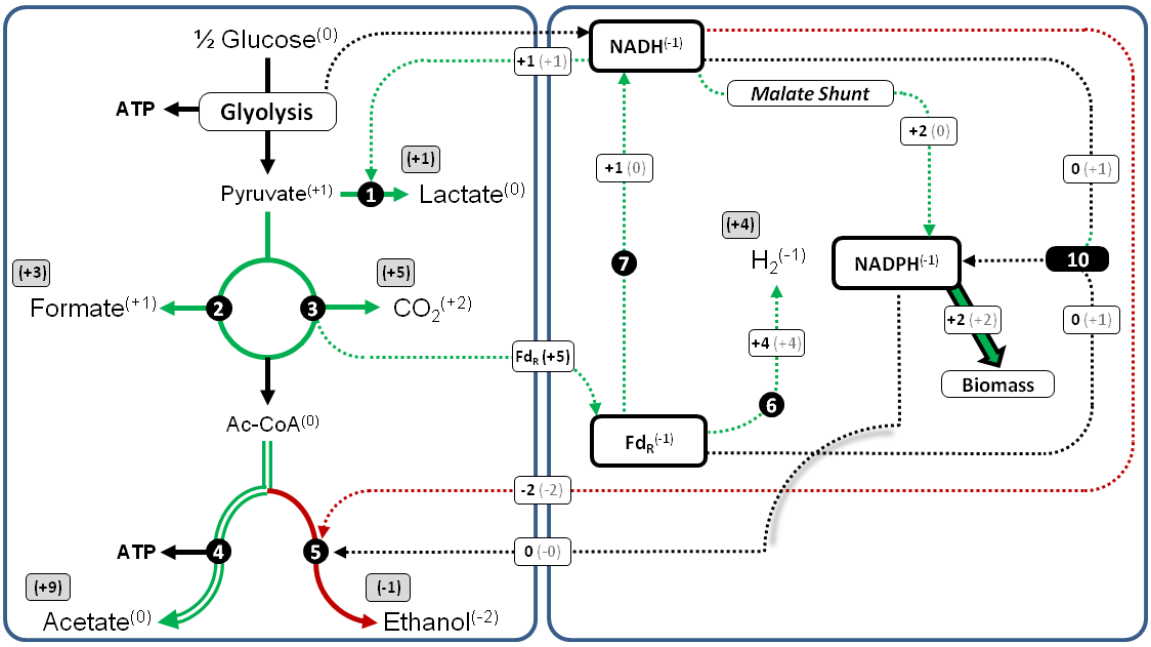
Regardless of detection sensitivity, parallel experiments have demonstrated that NFOR and Fd-dependent Ech H_2ase activity and/or expression is considerably lower than bifurcating and NADPH-dependent H_2ases (Rydzak et al. 2009; Raman et al. 2011; Rydzak et al. 2011; Rydzak et al. 2012). Therefore we have proposed how carbon and electron flux can change in response to changes in substrate oxidation state and inhibition of PFL using HPP through deduced metabolic pathways (Figures 7.5 and 7.6) using the '*traditional*' and '*novel*' electron transferring pathways, which employ either NFOR and Fd-dependent H_2ase, or bifurcating and NADPH-dependent pathways, respectively.

These are based on the assumptions that (i) CO_2 production is equal to reduced Fd production via PFOR, (ii) NADH production is constant during glycolysis, and (iii) NAD(P)H consumption/production can be determined from lactate, ethanol, and H_2 production measurements.

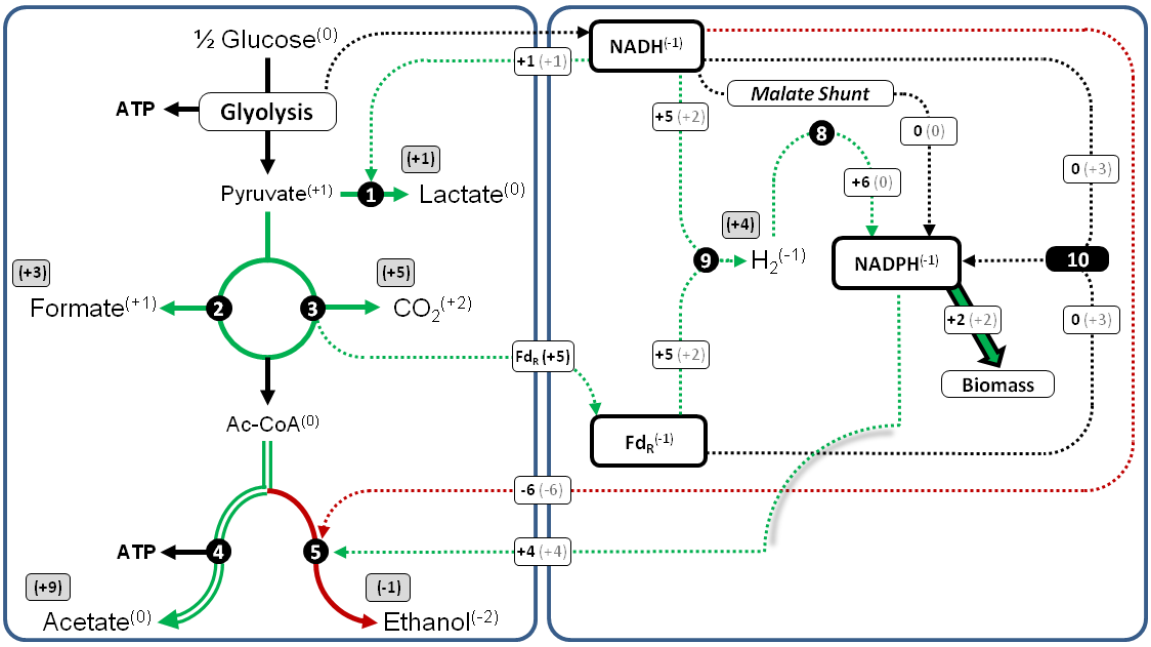
Figure 7.5 demonstrates how carbon and electron flux is altered in response to addition of pyruvate (80mM) to cellobiose grown *C. thermocellum* cultures. The co-metabolism of exogenous pyruvate and cellobiose versus cellobiose alone results in an increase in lactate, formate, CO_2, acetate, and H_2 production (~1, 3, 5, and 9 mM, respectively) and a decrease in ethanol production (~1 mM) (left panel in Figure 7.5a and b). The right panel demonstrates how electrons can flow between Fd, NADH, NADPH, and H_2 using *'malate shunt'* and NfnAB and either Fd-dependent H_2ase and NFOR (Figure 7.5a), or bifurcating H_2ases and NADPH-dependent H_2ase, (Figure 7.5b), respectively. Note that electron carriers coupling carbon and electron flow reactions (left panel) to carbonless electron-flow reactions (right panel) are depicted in the figure linking the panels.

Figure 7.5: Proposed redirection of carbon and electron flux in response to addition of exogenous pyruvate to cellobiose-grown cells. Carbon and electron fluxes based on changes in measured end products (shown in grey boxes) are depicted in the left panel. End-product redox values are provided in superscript next to corresponding end-product. Right panel depicts predicted changes in electron flux using only (a) NFOR and Fd-dependent H_2ase (black integers in boxes) or NFOR, Fd-dependent H_2ase, and NfnAB (grey integers in boxes) and (b) bifurcating H_2ase and NADPH-dependent H_2ase (black integers in boxes) or bifurcating H_2ase, NADPH-dependent H_2ase, and NfnAB (grey integers in boxes). Note that proposed electron flow through the malate shunt and towards biomass production is also indicated in the right panel. Increased (green) and decreased (red) flux (in direction depicted) is indicated in boxes. Color scheme represents fluxes only in the absence of NfnAB. Solid lines reflect carbon and electron flux and broken lines reflect electron flux only. 1, LDH; 2, PFL; 3, PFOR; 4, PTA/ACK; 5, AldH/ADH; 6, Fd-dependent H_2ase; 7, NFOR; 8, NADPH-dependent H_2ase; 9, bifurcating H_2ase; 10, NfnAB.

A)


½ Glucose(0)
ATP
Glyolysis
Pyruvate(+1)
Lactate(0)
Formate(+1)
CO2(+2)
Ac-CoA(0)
ATP
Acetate(0)
Ethanol(-2)
NADH(-1)
Malate Shunt
H2(-1)
NADPH(-1)
FdR(-1)
Biomass
+1 (+1)
(+1)
+1 (0)
+2 (0)
0 (+1)
(+4)
10
(+3)
(+5)
+2 (+2)
FdR (+5)
+4 (+4)
-2 (-2)
0 (-0)
(+9)
(-1)

B)


½ Glucose(0)
ATP
Glyolysis
Pyruvate(+1)
Lactate(0)
Formate(+1)
CO2(+2)
Ac-CoA(0)
ATP
Acetate(0)
Ethanol(-2)
NADH(-1)
Malate Shunt
H2(-1)
NADPH(-1)
FdR(-1)
Biomass
+1 (+1)
(+1)
+5 (+2)
0 (0)
+6 (0)
0 (+3)
(+4)
10
(+3)
(+5)
+2 (+2)
FdR (+5)
-6 (-6)
+4 (+4)
(+9)
(-1)

Using the '*traditional*' pathway, additional reduced Fd produced by metabolism of pyruvate into acetyl-CoA via PFOR, can be converted into H_2 using Fd-dependent H_2ase (Figure 7.5a). However, given that production of reduced Fd is greater than H_2 production (by 1 mM) suggests that excess reduced Fd produced must be utilized by alternative pathways. This can be accomplished by NFOR. Given that less ethanol is produced, and thus less total reducing equivalents are used, even when accounting for increased lactate production, we expect that electrons can be transferred from NADH to NADPH, which can subsequently be used for biomass production. This is consistent with the observed increase in biomass production in the presence of exogenous pyruvate. Transhydrogenation can be accomplished by increasing flux through the malate branch (Figure 7.5a; black integers), or via NfnAB, which simultaneously uses reduced Fd and NADH to produce 2 NADPH (Figure 7.5a; grey integers). Interestingly, despite the thermodynamic favorability of both (i) NAD^+ reduction using reduced Fd ($\triangle G°'$ = -15 kJ mol^{-1}) and (ii) ethanol production versus acetate plus H_2 production ($\triangle G°'$ = -5 kJ mol^{-1}) (Rydzak et al. 2009), H_2 production increased whereas ethanol production decreased (marginally), suggesting that *in vivo* NFOR activity is negligible.

Alternatively, the *'novel'* electron-transferring pathway can oxidize all of the additional reduced Fd produced (5 mM), with equimolar amounts of NADH, via highly expressed bifurcating H_2ases (Figure 7.5b; black integers). While this would lead to 10 mM additional H_2 produced, 6 mM of H_2 could be converted to NADPH via NADPH-dependent H_2ase, which can then be used for biomass production (2 mM) as well as NADPH-dependent ethanol synthesis (4 mM) as verified by enzyme activities by Rydzak

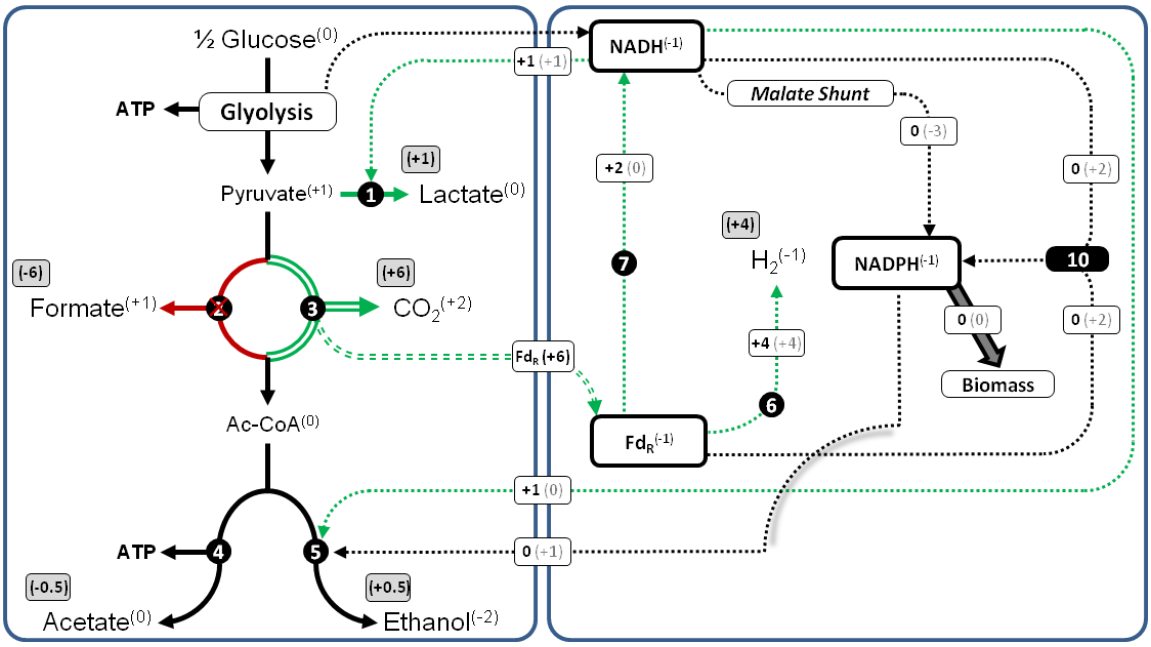
et al. (Rydzak et al. 2009; Rydzak et al. 2011). Given that (i) reduction of NADPH using H_2 is thermodynamically favorable under standard conditions (Rydzak et al. 2009; Rydzak et al. 2011; Carere et al. 2012) and (ii) intracellular ratios of NADPH/$NADP^+$ ratios are typically low (i.e. 0.5 in *C. acetobutylicum*; (Vasconcelos et al. 1994)), H_2 uptake via NADPH-dependent H_2ase may be possible. Alternatively, bifurcating H_2ases may oxidize only a portion (2 mM) of the additional reduced Fd produced, with the remainder catalyzed to NADPH via NfnAB (Figure 7.5b).

A similar analysis of potential electron flow was assessed in response to inhibition of PFL using HPP (Figure 7.6). Addition of HPP increased CO_2, H_2, lactate, and acetate production, (~6, 4, 1 and 0.5 mM, respectively) and decreased formate and ethanol production (~ 6, 0.5 mM, respectively). Using the *'traditional'* pathway, the additional reduced Fd produced by redirecting flux through PFOR can be used to generate H_2 and NADH via Fd-dependent H_2ase and NFOR, respectively (Figure 7.6a; black integers). Given that growth is unaffected by addition of HPP, we would not expect an increase in NADPH utilization for biomass production. Instead, in the absence of NFOR, excess reduced Fd can be oxidized via NfnAB, which would trigger the transition from NADH to NADPH-dependent ethanol production, and would be accompanied by a decrease in transhydrogenation via the malate shunt (Figure 7.6a; grey integers). Alternatively, the *'novel'* pathway would cause a major shift in NADH to NADP-dependent ethanol production, regardless of NfnAB activity (Figure 7.6b). In the absence of NfnAB, excess H_2 produced via bifurcating H_2ases would be, again, converted into NADPH via

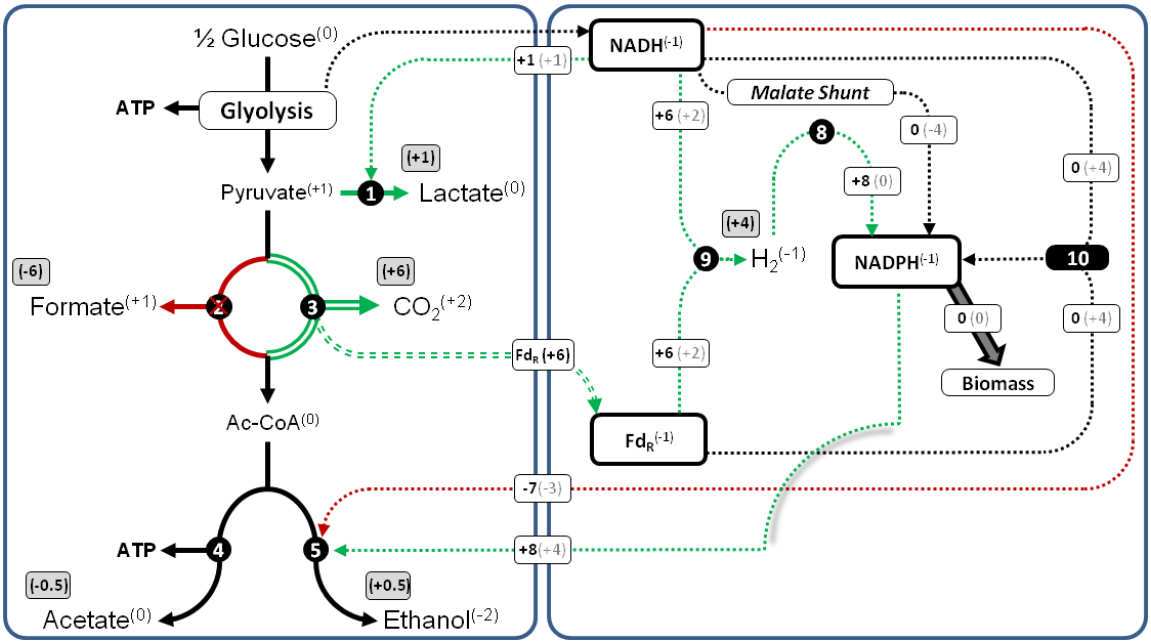
NADPH-dependent H_2ases, whereas in the presence of NfnAB, excess reduced Fd would be converted into NADPH at the expense of malate shunt flux.

Figure 7.6: Proposed redirection of carbon and electron flux in response to inhibition of PFL via HPP. Carbon and electron fluxes based on changes in measured end products (shown in grey boxes) are depicted in the left panel. End-product redox values are provided in superscript next to corresponding end-product. Right panel depicts predicted changes in electron flux using only (a) NFOR and Fd-dependent H_2ase (black integers in boxes) or NFOR, Fd-dependent H_2ase, and NfnAB (grey integers in boxes) and (b) bifurcating H_2ase and NADPH-dependent H_2ase (black integers in boxes) or bifurcating H_2ase, NADPH-dependent H_2ase, and NfnAB (grey integers in boxes). Note that proposed electron flow through the malate shunt and towards biomass production is also indicated in the right panel. Increased (green) and decreased (red) flux (in direction depicted) is indicated in boxes. Color scheme represents fluxes only in the absence of NfnAB. Solid lines reflect carbon and electron flux and broken lines reflect electron flux only. 1, LDH; 2, PFL; 3, PFOR; 4, PTA/ACK; 5, AldH/ADH; 6, Fd-dependent H_2ase; 7, NFOR; 8, NADPH-dependent H_2ase; 9, bifurcating H_2ase; 10, NfnAB.

A)


½ Glucose(0)
ATP
Glyolysis
Pyruvate(+1)
Lactate(0)
Formate(+1)
CO2(+2)
Ac-CoA(0)
ATP
Acetate(0)
Ethanol(-2)
NADH(-1)
Malate Shunt
H2(-1)
NADPH(-1)
Biomass
FdR(-1)
FdR (+6)
+1 (+1)
(+1)
+2 (0)
0 (-3)
0 (+2)
(+4)
(-6)
(+6)
+4 (+4)
0 (0)
10
+1 (0)
0 (+1)
(-0.5)
(+0.5)

B)


½ Glucose(0)
ATP
Glyolysis
Pyruvate(+1)
Lactate(0)
Formate(+1)
CO2(+2)
Ac-CoA(0)
ATP
Acetate(0)
Ethanol(-2)
NADH(-1)
Malate Shunt
H2(-1)
NADPH(-1)
Biomass
FdR(-1)
FdR (+6)
+1 (+1)
(+1)
+6 (+2)
0 (-4)
0 (+4)
+8 (0)
(+4)
(-6)
(+6)
0 (0)
10
-7(-3)
+8(+4)
(-0.5)
(+0.5)

7.6 Conclusions

While genomic and proteomic expression profiles can help predict how carbon and electron flux is distributed throughout pathways leading to end-product synthesis (Gowen and Fong 2010), actual *in vivo* flux distribution can only be validated by measuring intracellular metabolite levels and/or end-products during growth under specific conditions. Although flux analysis requires a full profile of intracellular metabolites and end-product synthesis patterns over time, the well deduced branched end-product pathways of *C. thermocellum* have allowed us to construe how fluxes may be redistributed among these pathways in response to fermentation conditions using end-product concentrations over a defined period of growth. While electron transfer between NAD(P)H, Fd, and H_2 has traditionally been assumed to be catalyzed by NFOR and Fd-dependent H_2ases, more recent studies have demonstrated the importance of bifurcating H_2ases (Schut and Adams 2009; Calusinska et al. 2010), NADPH-dependent H_2ases (Calusinska et al. 2010), the malate shunt (Li et al. 2012b), and NfnAB (Wang et al. 2010) in electron shuttling. Furthermore, electron carrier specificity required for H_2 and ethanol producing enzymes is not as strict as once thought. Elucidation of electron flux through encoded, and seemingly redundant, pathways may bring forth alternative metabolic engineering strategies to direct carbon and electron flux towards desired pathways through manipulation of $NADH/NAD^+$, NADPH/NADH, and Fd_{red}/Fd_{ox} ratios (Veit et al. 2008; Cho et al. 2011).

7.7 Authors' contributions

Thomas Rydzak, John A. Wilkins, David B. Levin, Oleg V. Krokhin, Nazim Cicek, and Richard Sparling conceived and designed the study. Thomas Rydzak performed growth studies, end-product analysis, processed samples for MRM analysis, selected proteins for MRM analysis, analyzed proteomic data, and drafted the manuscript with input from David B. Levin and Richard Sparling. Zack J. Cunninghm helped with growth studies and end-product analysis. Marina Grigoryan performed MRM runs with help from Peyman Ezzati and Oleg V. Krokhin.

7.8 Acknowledgements

This work was supported by funds provided by Genome Canada (to MGCB2 project), the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944–04) and the BIOCAP Canada Foundation.

Chapter 8: Thesis Conclusions and Future Directions

8.1 Thesis conclusions

The transition from fossil fuels to sustainable, environmentally friendly biofuels can lower greenhouse gas emissions and promote energy security. Consolidated bioprocessing is an attractive processing strategy for production of biofuels. *C. thermocellum*, given its high cellulose degradation rates and ability to produce biofuels such as ethanol and H_2, serves as an excellent candidate for biofuel production via consolidated bioprocessing. Furthermore, these inherent qualities have made it the subject of numerous biochemical, physiological, and molecular studies, resulting in a good characterization of its bioprocessing capabilities. Therefore, it is an excellent platform organism to further study how genome content, gene product expression, and enzyme activities can influence metabolic flux towards end products.

Despite its appealing characteristics, *C. thermocellum* produces poor ethanol and H_2 yields. The aim of this thesis was to (i) demonstrate that end-product yields can be altered by manipulation of fermentation conditions and (ii) to help refine how genome content, gene product expression, and enzyme activities can influence metabolic flux towards end products, which could be used to apply targeted metabolic engineering strategies for improvement of product yields in the future. We address the contributions made to the specific objectives of this thesis below.

Objective 1: Elucidation of enzymes activities involved in pyruvate catabolism and end-product synthesis, with particular emphasis on alcohol dehydrogenases and hydrogenases (Chapters 2, 3, and 4).

1) We detected activities of LDH, PFL, PFOR, H_2ases, ADHs, AldH, MalE, and MDH. FHL and PPK activities were not detected. Accurate NFOR activities could not be assessed in crude cell extracts.
2) NADPH-dependent H_2ase activities were higher than Fd-dependent and NADH dependent. Bifurcating H_2ase activities have not yet been assessed in crude cell extracts.
3) Comparable NADH and NADPH-dependent ADH activities were detected.
4) MDH activity was both NADH and NADPH-dependent, albeit NADH-dependent was greater that NADPH-dependent activity. MalE activity was strictly NADPH-dependent.

This provided a baseline of metabolic pathways potentially used *in vivo* for the production of end-products and demonstrated that both NADH and NADPH can be used as cofactors for reduction of intermediates during end-product synthesis.

Objective 2: We hypothesized that end-product yields are affected by fermentation conditions. We provide evidence consistent with this hypothesis by demonstrating that (i) addition of exogenous end products prior to fermentation (Chapter 3), (ii) fermenter sparging with H_2 or N_2 at different rates (Chapter 4), (iii) manipulation of substrate oxidation state (addition of pyruvate) (Chapter 7), and (iv) use of protein inhibitors

including hypophosphite (PFL inhibitor) of carbon monoxide (H_2ase inhibitor) (Chapter 7) can influence end-product yields. Specifically:

1) Exogenous H_2 increased ethanol and formate yields, and decreased CO_2 yields.
2) Exogenous acetate increased acetate yields.
3) Exogenous ethanol increased H_2 and acetate yields.
4) Exogenous formate increased H_2 yields.
5) Exogenous lactate appeared to be metabolized and increased all end products with the exception of formate, which decreased in response to exogenous lactate.
6) Sparging with N_2 at high rates increased H_2, CO_2, and acetate yields and decreased formate and ethanol yields marginally when compared to sparging with N_2 at low rates.
7) Sparging with H_2 at high rates increased ethanol and formate yields and decreased H_2, CO_2, and acetate yields when compared to sparging with N_2 at low rates.
8) In the presence of cellobiose, exogenous pyruvate was metabolized and (i) increased acetate, formate, CO_2, and H_2 yields, (ii) triggered lactate production, and (iii) and marginally decreased ethanol yields.
9) Inhibition of PFL using HPP increased CO_2, H_2 yields, triggered lactate production, marginally decreased acetate yields, and had little impact on ethanol yields.

10) Inhibition of H_2ases using CO increased ethanol and formate yields, triggered lactate production, and decreased H_2, CO_2, and acetate yields.

This demonstrated that manipulation of fermentation conditions through changing either end-product residence times, gas sparging rates or gasses, or substrate oxidation states can alter end-product yields. Furthermore, the use of enzyme inhibitors, which mimic gene knockouts/knockdowns, can also redirect bacterial metabolism.

Objective 3: We hypothesized that genome content can dictate end-product synthesis patterns. Our meta-analysis of select H_2 and/or ethanol producing bacteria linking genome content to end-product synthesis patterns (Chapter 5) only partially supported this hypothesis. Specifically:

1) Despite the presence of multiple ADH encoding genes, absence of AldH and/or AdhE encoding genes coincide with high H_2 yields and low/no ethanol production
2) Presence of both H_2ases and AdhE encoding resulted in relatively uniform mixed end-product patterns.

However:

3) While presence of LDH and PFL is not conducive for H_2 or ethanol production, genes encoding LDH or PFL did not necessarily confer high levels of lactate or formate production, respectively
4) Absence of all H_2ase encoding genes did not necessarily confer high ethanol production, but rather high lactate production

5) The types of H_2ases encoded do not necessarily appear to have an impact on H_2 production. Indeed the Fd-dependent [NiFe]-Ech H_2ase seemed to play a minor role, if any, in the production of hydrogen in *C. thermocellum* under the growth conditions tested (Chapters 2, 3, 4, 6, 7) on MRM

Therefore, while the genome can be used to (i) deduce potential end-product synthesis patterns, (ii) identify suitable biomarkers for screening a microorganism's potential of H_2 and/or ethanol production capabilities, and (iii) identify targets for metabolic engineering to improve biofuel yields, other factors including gene and gene product expression and pathway flux regulation through allostery can influence end-product yields and must be elucidated.

Objective 4: In order to validate our hypothesis that genes encoding proteins with analogous functions are differentially expressed, we measured protein expression profiles of proteins involved in core metabolism using shotgun proteomics (Chapters 6) and MRM proteomics (Chapter 7). Expression profiles demonstrate that:

1) Encoded proteins involved in different pathway branches as well as proteins with analogous functions and are in fact differentially expressed

Specifically:

2) The majority of proteins involved in pyruvate catabolism and end-product synthesis were detected at high levels with the exception of AldH, Fd-dependent Ech-type H_2ase, and NFOR.

3) Several putative proteins with analogous functions including glucose kinase Cthe_0390 (vs. Cthe_2938), fructose-1,6-bisphosphate aldolase Cthe_0319 (vs. Cthe_0349), phosphoglycerate mutase Cthe_1425, 2449, and 3153 (vs. Cthe_0140, 0707, 0946, and 1292), and PFOR Cthe_2794-2797 (vs. Cthe_2390-2393, 3120, 0866, and 0614) were not detected.
4) Many enzymes involved in direct and indirect interconversion of PEP and pyruvate were equally expressed at high levels. Specifically, PPDK, enzymes involved in the malate shunt (PEPDK, MDH, MalE) used for transhydrogenation, and OAADC were detected at high levels whereas PEP synthase was not.
5) Expression of a number of putative PP_i-dependent enzymes was comparable (if not higher) to that of ATP-dependent enzymes involved in analogous reactions. These include PP_i-dependent phosphofructokinase Cthe_0347 (vs. ATP-dependent phosphofructokinase Cthe_1261/Cthe_0389), PPDK (vs. PEP synthase or malate shunt enzymes), and ATK (vs PTA and ACK).

This not only has allowed refinement of metabolic flux distribution predictions through core metabolic pathways, but has also provided specific targets (i.e. genes encoding highly expressed proteins) for metabolic engineering that would most likely cause significant and/or desirable effects. Of note, expression profiles revealed that the malate shunt may play an important role in transhydrogenation between NADH and NADPH, which may in turn affect reduced product yields. Furthermore, expression profiles have

also revealed the significance of PP_i-dependent pathways, demonstrating that PP_i, in addition to ATP, may serve as an important energy carrier.

Objective 5: We hypothesized that changes in gene/gene-product expression can dictate carbon and electron flux, and, in turn, end-product yields, in response to different fermentation conditions. We therefore measured changes in (i) enzyme activities in response to addition of exogenous end products (Chapter 3) and gas sparging (Chapter 4), (ii) mRNA levels (using RT-qPCR) in response to gas sparging (Chapter 4), (iii) protein levels (using 2D-HPLC-MS/MS with iTRAQ labeled peptides) in response to growth phase (Chapter 6), and (iv) protein levels (using MRM mass spectrometry) in response to changes in substrate oxidation state (using pyruvate) and the use of protein inhibitors (HPP and CO; Chapter 7). While changes in expression profiles were observed in some proteins, they were not as pronounced as expected and did not always reflect changes in end-product yields. Specifically:

1) Enzyme activities did not change in response to addition of end products with the following exceptions: PFL and Fd-H_2ase activity increased 2-fold in response to added ethanol and PFOR activity decreased by ~50% in response to added formate. While increase in Fd-H_2ase correlated with an increase in H_2 yields, there was no correlation between changes in PFL and PFOR with changes in formate and CO_2 production. Furthermore, changes in end-product yields were generally not accompanied by changes in enzyme activities

suggesting that thermodynamically driven inhibition of select end-product pathways may be responsible for these observed changes.

2) Enzyme activities or mRNA levels did not change significantly in response to different gas sparging conditions with the exception of ADH activity and transcription. NADH-dependent ADH activity increased 2-fold in response to high H_2 versus high N_2 sparging, and was accompanied by a 10-fold increase in *adhE* transcription and a 3-fold decrease in *adhZ* transcription. While the increase in ADH activity and *adhE* transcription correlated well with increased ethanol yields, the downregulation of *adhZ* suggests that ethanol production is tightly regulated through gene expression.
3) While differential protein expression during transition from exponential to stationary phase was observed in 24% of the core metabolic proteins involved in cellodextrin breakdown and transport, glycolysis, pentose phosphate pathway, glycerol metabolism, pyruvate metabolism, and end-product synthesis, most of these changes, while statistically significant, were less than 2-fold. This was consistent with only minor changes in end-product kinetics (i.e. slight inversion of acetate to ethanol ratio).
4) Protein expression again did not change significantly (>2-fold) in response to addition of pyruvate to cellobiose grown cells or in response to enzyme inhibitors, despite the significant differences observed in end-product yields.

The general lack of significant changes in protein expression, regardless of assay methods used, did not correlate well with the significant changes in end-product synthesis

patterns observed. This was unexpected given that previous studies on other organisms suggested otherwise (Goodwin and Zeikus 1987; Junelles et al. 1988; Grupe and Gottschalk 1992; Girbal and Soucaille 1994; Girbal et al. 1995; Guedon et al. 1999b; Guedon et al. 2000b; Soboh et al. 2004; Stevenson and Weimer 2005; Maness and Thammannagowda 2008). This suggests that changes in end-product synthesis in *C. thermocellum* are not primarily dictated by changes in enzyme expression, but rather intracellular metabolite levels, cofactor ratios (i.e. NAD(P)H/NAD(P)$^+$, Fd_{red}/Fd_{ox}), or allosteric regulation of enzymes. Furthermore, the highly branched core metabolic pathways expressed in *C. thermocellum*, in addition to the multiple proteins capable of electron-transfer reactions between different cofactors (NADH, NADPH, Fd), may negate the need to alter expression of enzymes typically required to maintain carbon and electron balance in response to changing conditions.

8.2 Future Perspectives

While the research outlined in this thesis (and corresponding publications) has significantly improved our understanding of fermentative metabolism of *C. thermocellum*, which can also be extended to other fermentative bacteria, a number of strategies must still be explored in order to enhance the economic viability of lignocellulosic biofuel production. These include (i) improvement of desired biofuel yields, (ii) improvement of biofuel production rates, (iii) improvement of strain tolerance to high levels of end-products, (iv) co-production of low-volume, high-value products to offset the costs of producing high-volume, low value biofuels, and (v) the eventual

transition from ethanol and/or H_2 to novel biofuels that can be more readily used with current technologies. These strategies will still require a better understanding of bacterial metabolism which can be attained through (i) enzyme characterization, (ii) discovery of regulatory elements and networks involved in gene and protein expression, (iii) metabolomic profiling of intracellular metabolites, (iv) metabolic flux modelling, and (v) improved genetic engineering technologies.

Optimization of fermentation conditions is one strategy that can be employed to improve biofuel yields. Studies have shown that changing media composition (Junelles et al. 1988; Dabrock et al. 1992; Guedon et al. 1999a), pH (Goodwin and Zeikus 1987; Grupe and Gottschalk 1992; Desvaux et al. 2001c), and dilution rates (Guedon et al. 1999b; Guedon et al. 2000a; Guedon et al. 2000b; Desvaux et al. 2001b; Desvaux et al. 2001a; Desvaux and Petitdemange 2001; Desvaux et al. 2001d) can all influence product yields. We have also demonstrated how addition of individual end-products, fermenter sparging, and substrate oxidation states can influence end-product yields. While it may be impractical to intentionally add exogenous end products to redirect metabolism during high scale fermentations, manipulating specific end-product residence times in a fermenter may be a novel approach in redirecting metabolic flux. Furthermore, while sparging a fermenter at high gas-flow rates or using costly gasses such as H_2 is also impractical, variable stirring rates (Lamed et al. 1988), as well as novel membrane technologies (Liang et al. 2002), which can selectively remove gaseous end-products, can also be utilized to manipulate product yields. Finally, the use of low-cost industrial by-products with a low redox values (ex. glycerol) (Girbal and Soucaille 1994; Vasconcelos

et al. 1994; Saint-Amans et al. 2001; da Silva et al. 2009) can also be used to increase reduced end-product yields, providing that they are metabolized.

Genetic engineering of strains capable of lignocellulosic biofuels production is, perhaps, the best way to improve product yields and other desirable characteristics. However, a thorough understanding of bacterial metabolism is required for effective strain manipulation. Our elucidation of expressed metabolic pathways (using RT-qPCR and proteomics) involved in end-product synthesis has provided specific targets for genetic engineering attempts. For example, our hypothesis that the Fd-dependent, Ech-type, [NiFe]-H_2ase does not play a substantial role in H_2 production in *C. thermocellum* has been supported by colleagues performing parallel characterization studies, who demonstrated that genetic knockouts of the *hydG* gene, the maturation factor specific for [FeFe]-H_2ases (but not [NiFe] H_2ases), prevents H_2 production and increases ethanol yields in strain DSM 1313 (Guss and Lynd 2012). Other successful genetic engineering strategies have been demonstrated in *C. cellulolyticum*, where heterologous expression of PDC and ADH alleviated metabolic buildup at the pyruvate branchpoint, which increased growth and cellulose consumption, and decreased lactate and extracellular pyruvate production (Guedon et al. 2002). Surprisingly, the increase in ethanol production was not favored over the increase in H_2 production. Furthermore, deletion of *pta* in *C. thermocellum* DSM 1313, and subsequent knockouts of *pta* and *ldh*, did not increase ethanol production until these strains were 'evolved' on increasing cellulose concentrations over time (Argyros et al. 2011; Tripathi et al. 2012). Thus, seemingly targeted engineering strategies do not always result in rational results, demonstrating the

complexity of bacterial metabolism. Metabolomic profiling and metabolic flux modelling, which can be improved by omics data as presented here, may provide additional insights into regulation of metabolic flux and provide more suitable targets for genetic engineering.

While genetic engineering of end-product synthesis pathways is one approach of improving desirable end-product yields, enhancing substrate utilization pathways can also be targeted through genetic engineering. For example, while *C. thermocellum* can break down the xylan portion of hemicelluloses via encoded xylanases (Hayashi et al. 1997), it cannot utilize xylose, presumably due to the absence xylose reductase, xylitol dehydrogenase, xylose isomerase, or xylulokinase, which are required for xylose utilization. Thus heterologous expression of these enzymes may confer xylose utilization and, in turn, increase product yields on raw lignocellulosic substrates. Furthermore, while xylose-specific transporters may be present in *C. thermocellum*, their identification is difficult verify from sequence alignments alone, and thus their presence must be verified experimentally. Expression of other glucosidases and pathways capable of breakdown and utilization of other components of lignocelluloses can also increase fermentation yields on raw substrates.

In addition to optimal product yields, a number of other strain characteristics must be considered when attempting to minimize fermentation costs. Given that fermenter operation can be energy intensive and costly, high biofuels production rates are required to reduce operational costs. This may potentially be achieved by increasing substrate uptake rates or by increasing flux through desired fermentation pathways by expressing

enzymes with improved kinetic properties. These enzymes can be either naturally encoded by other microorganisms or tailored via recombinant mutagenesis or directed evolution (McLaughlin et al. 2009). Tolerance to high concentrations of end products can also decrease operation costs through allowing increased bacterial residence times and decreasing dilution rates during continuous fermentations. Ethanol, specifically, can cause leakage of the plasma membrane adversely affecting bacterial growth, viability, and metabolism. Thus, improved ethanol tolerance, typically achieved through strain evolution, is commonly linked to changes in cell membrane composition (Ingram 1990). More recently, ethanol tolerance in an evolved *C. thermocellum* strain was attributed to a change in AdhE cofactor specificity (from NADH to NADPH), suggesting that rational engineering strategies may be employed to improve ethanol tolerance (Brown et al. 2011).

Finally, in order to decrease production costs and improve long-tem sustainability of biofuels production at the industrial scale, optimization of all processes including biomass conversion, biofuels production, production value-added chemicals, and power generation is necessary. A biorefinery, based on the intensive fractionation scheme of fossil fuel refineries, integrates all of these processes (Bioenergy and Energy Planning Research Group 2010). One of the key features in improving the economic viability of lignocellulosic biofuels is to utilize all possible components of lignocellulose and to produce low-volume, high-value chemicals to offset production costs of high-volume, low-value biofuels. While fermentable sugars can be used for biofuels production, other accompanying byproducts can be used directly or converted into value-added products.

For example, xylitol or sorbitol can be used directly as artificial sweeteners. Lactate can be used as a building block for polylactides, which can then be used for biodegradable packaging material. Succinate can be used as a building block for butanediol or tetrahydrofuran, which can be used as industrial solvents or used for manufacturing of polyurethane fibers such as Lycra or Spandex. Furthermore, non-fermentable lignocellulosic components such as lignin, can be used as a solid fuel for heat and power generation, or can be used as building blocks for vanillin, a flavoring agent. Hence, through optimization of all aspects of biofuels production, there is a great potential for biofuels production to be economically viable sustainable in the long run.

Literature Cited

Abbad-Andaloussi S, Manginot-Durr C, Amine J, Petitdemange E, Petitdemange H (1995) Isolation and characterization of *Clostridium butyricum* DSM 5431 mutants with Increased resistance to 1,3-propanediol and altered production of acids. Appl Environ Microbiol 61:4413-4417

Adams MW, Mortenson LE, Chen J-S (1981) Hydrogenase. Biochim Biophys Acta 594:105-176

Aggarwal K, Choe LH, Lee KH (2006) Shotgun proteomics using the iTRAQ isobaric tags. Brief Funct Genomic Proteomic 5:112-120

Allen L, Cohen MJ, Abelson D, Miller B (2011) Fossil Fuels and Water Quality In: Gleick PH (ed) The World's Water. Island Press/Center for Resource Economics, pp 73-96

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410

Angenent LT, Karim K, Al-Dahhan MH, Wrenn BA, Domiguez-Espinosa R (2004) Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends Biotechnol 22:477-485

Argyros DA et al. (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Appl Environ Microbiol 77:8288-8294

Aristidou A, Penttila M (2000) Metabolic engineering applications to renewable resource utilization. Curr Opin Biotechnol 11:187-198

Asanuma N, Hino T (2000) Effects of pH and energy supply on activity and amount of pyruvate formate-lyase in *Streptococcus bovis*. Appl Environ Microbiol 66:3773-3777

Asanuma N, Yoshii T, Hino T (2004) Molecular characteristics and transcription of the gene encoding a multifunctional alcohol dehydrogenase in relation to the deactivation of pyruvate formate-lyase in the ruminal bacterium *Streptococcus bovis*. Arch Microbiol 181:122-128

Axley MJ, Grahame, D. A., Stadtman, T. C. (1990) *Escherichia coli* formate-hydrogen lyase. Purification and properties of the selenium-dependent formate dehydrogenase component. Journal of Biological Chemistry 265:18213-18218

Bagramyan K, Trchounian A (2003) Structural and functional features of formate hydrogen lyase, an enzyme of mixed-acid fermentation from Escherichia coli. Biochemistry (Mosc) 68:1159-1170

Ballongue J, Amine J, Petitdemange E, Gay R (1986) Enhancement of solvents production by *Clostridium acetobutylicum* cultivated on reducing compounds depletive medium. Biomass 10:121-129

Bapteste E, Moreira D, Philippe H (2003) Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. Gene 318:185-191

Barquera B et al. (2002) Purification and characterization of the recombinant Na(+)-translocating NADH:quinone oxidoreductase from *Vibrio cholerae*. Biochemistry 41:3781-3789

Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521-554

Ben-Bassat A, Lamed R, Zeikus JG (1981) Ethanol production by thermophilic bacteria: metabolic control of end product formation in *Thermoanaerobium brockii*. J Bacteriol 146:192-199

Benemann JR (1990) The future of microalgae biotechnology In: Cresswell RC, Rees TAV (eds) Algal biotechnology. Longman, London, pp 317-337

Bernard N, Johnsen K, Holbrook JJ, Delcour J (1995) D175 discriminates between NADH and NADPH in the coenzyme binding site of *Lactobacillus delbrueckii* subsp. *bulgaricus* D-lactate dehydrogenase. Biochem Biophys Res Commun 208:895-900

Biebl H (1991) Glycerol fermentation of 1,3-propanediol by *Clostridium butyricum*. Measurement of product inhibition by use of a pH-auxostat. Appl Microbiol Biotechnol 35:701-705

Biebl H (2001) Fermentation of glycerol by *Clostridium pasteurianum*--batch and continuous culture studies. J Ind Microbiol Biotechnol 27:18-26

Bielen AAM, Willquist K, Engman J, Van Der Oost J, Van Niel EWJ, Kengen SWM (2010) Pyrophosphate as a central energy carrier in the hydrogen-producing extremely thermophilic *Caldicellulosiruptor saccharolyticus*. FEMS Microbiology Letters 307:48-54

Bioenergy and Energy Planning Research Group (2010) Biorefineries. http://bpe.epfl.ch/page-34016-en.html (accessed June 2013)

Blumenthal M, Johnson MK, Johnson EJ (1967) Distribution of heat labile and heat stable inorganic pyrophosphatase. Can J Microbiol 13:1695-1699

Bothun GD, Knutson BL, Berberich JA, Strobel HJ, Nokes SE (2004) Metabolic selectivity and growth of *Clostridium thermocellum* in continuous culture under elevated hydrostatic pressure. Appl Microbiol Biotechnol 65:149-157

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254

Bredholt S, Sonne-Hansen J, Nielsen P, Mathrani IM, Ahring BK (1999) *Caldicellulosiruptor kristjanssonii* sp nov., a cellulolytic extremely thermophilic, anaerobic bacterium. International Journal of Systematic Bacteriology 49:991-996

Brekasis D, Paget MS (2003) A novel sensor of NADH/NAD+ redox poise in Streptomyces coelicolor A3(2). EMBO J 22:4856-4865

Breznak JA, Brune A (1994) Role of microorganisms in the digestion of lignocellulose by termites. Annu Rev Entomol 39:453-487

Brown MA, Levine MD, Romm JPRAH, Koomey JH (1998) Engineering-economic studies of energy technologies to reduce greenhouse gas emissions: opportunites and challenges. Annual review of energy environment 23:31-39

Brown SD et al. (2011) Mutant alcohol dehydrogenase leads to improved ethanol tolerance in *Clostridium thermocellum*. Proc Natl Acad Sci U S A 108:13752-13757

Brown TD, Pereira CR, Stormer FC (1972) Studies of the acetate kinase-phosphotransacetylase and the butanediol-forming systems in *Aerobacter aerogenes*. J Bacteriol 112:1106-1111

Bruggemann H, Gottschalk G (2008) Comparative genomics of clostridia: link between the ecological niche and cell surface properties. Ann N Y Acad Sci 1125:73-81

Brush EJ, Lipsett KA, Kozarich JW (1988) Inactivation of *Escherichia coli* pyruvate formate-lyase by hypophosphite: evidence for a rate-limiting phosphorus-hydrogen bond cleavage. Biochemistry 27:2217-2222

Buhrke T, Lenz O, Porthun A, Friedrich B (2004) The H_2-sensing complex of *Ralstonia eutropha*: interaction between a regulatory [NiFe] hydrogenase and a histidine protein kinase. Molecular Microbiology 51:1677-1689

Burdette D, Zeikus JG (1994) Purification of acetaldehyde dehydrogenase and alcohol dehydrogenases from *Thermoanaerobacter ethanolicus* 39E and characterization of the secondary-alcohol dehydrogenase (2 degrees Adh) as a bifunctional alcohol dehydrogenase--acetyl-CoA reductive thioesterase. Biochem J 302 (Pt 1):163-170

Calusinska M, Happe T, Joris B, Wilmotte A (2010) The surprising diversity of clostridial hydrogenases: a comparative genomic perspective. Microbiology 156:1575-1588

Carere C, Kalia V, Sparling R, Cicek N, Levin D (2008a) Pyruvate catabolism and hydrogen synthesis pathway genes of *Clostridium thermocellum* ATCC 27405. Indian Journal of Microbiology 48:252-266

Carere CR, Rydzak T, Verbeke TJ, Cicek N, Levin DB, Sparling R (2012) Linking genome content to biofuel production yields: a meta-analysis of major catabolic pathways among select H^2 and ethanol-producing bacteria. BMC Microbiol 12:295

Carere CR, Sparling R, Cicek N, Levin DB (2008b) Third Generation Biofuels via Direct Cellulose Fermentation. Int J Mol Sci 9:1342-1360

Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol 60:165-182

Chang T, Yao S (2011) Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. Appl Microbiol Biotechnol 92:13-27

Chemical Rubber Company. (1977) CRC handbook of chemistry and physics. In. CRC Press, Cleveland, OH

Chittibabu G, Nath K, Das D (2006) Feasibility studies on the fermentative hydrogen production by recombinant *Escherichia coli* BL-21. Process Biochem 41:682-688

Cho H-S, Kim YM, Min BE, Jung GY, Park JM (2011) Improvement of hydrogen production yield by rebalancing NADPH/NADH ratio in a recombinant *Escherichia coli*. J Nanoelectron Optoe 6:1-5

Chou CJ et al. (2007) Impact of substrate glycoside linkage and elemental sulfur on bioenergetics of and hydrogen production by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl Environ Microbiol 73:6842-6853

Claassen PAM et al. (1999) Utilisation of biomass for the supply of energy carriers. Appl Microbiol Biotechnol 52:741-755

Cleries R, Galvez J, Espino M, Ribes J, Nunes V, de Heredia ML (2012) BootstRatio: A web-based statistical analysis of fold-change in qPCR and RT-qPCR data using resampling methods. Comput Biol Med 42:438-445

Congressional Budget Office (2010) Using Biofuel Tax Credits to Achieve Energy and Environmental Policy Goals. http://www.cbo.gov/publication/21444 (accessed May 2013)

Cueto PH, Mendez BS (1990) Direct selection of *Clostridium acetobutylicum* fermentation mutants by a proton suicide method. Appl Environ Microbiol 56:578-580

da Silva GP, Mack M, Contiero J (2009) Glycerol: a promising and abundant carbon source for industrial microbiology. Biotechnol Adv 27:30-39

da Silva SS, Chandel AK, Wickramasinghe SR, Dominguez JM (2012) Fermentative production of value-added products from lignocellulosic biomass. J Biomed Biotechnol 2012:826162

Dabrock B, Bahl H, Gottschalk G (1992) Parameters affecting solvent production by *Clostridium pasteurianum*. Appl Environ Microbiol 58:1233-1239

Daniels L, Belay N, Rajagopal BS (1986) Assimilatory reduction of sulfate and sulfite by methanogenic bacteria. Appl Environ Microbiol 51:703-709

Daniels L, Wessels D (1984) A method for the spectrophotometric assay of anaerobic enzymes. Anal Biochem 141:232-237

Darrett RH, Drisham CM (1995) Biochemistry. Saunders College Publishing, New York

Das D, Veziroğlu TV (2001) Hydrogen production by biological processes: a survey of literature. Int J Hydrogen Energy 26:13-28

Datta R, Zeikus JG (1985) Modulation of acetone-butanol-ethanol fermentation by carbon monoxide and organic acids. Appl Environ Microbiol 49:522-529

De Lacey AL et al. (2002) IR spectroelectrochemical study of the binding of carbon monoxide to the active site of *Desulfovibrio fructosovorans* Ni-Fe hydrogenase. J Biol Inorg Chem 7:318-326

de Vrije T et al. (2007) Glycolytic pathway and hydrogen yield studies of the extreme thermophile *Caldicellulosiruptor saccharolyticus*. Applied microbiology and biotechnology 74:1358-1367

Deibel RH, Niven CF, Jr. (1964) Pyruvate Fermentation by *Streptococcus Faecalis*. J Bacteriol 88:4-10

Demain AL, Newcomb M, Wu JH (2005) Cellulase, clostridia, and ethanol. Microbiol Mol Biol Rev 69:124-154

Demirbas A (2009) Biohydrogen: For future engine fuel demands. Springer, Trabzon

Deng Y, Olson DG, Zhou J, Herring CD, Joe Shaw A, Lynd LR (2013) Redirecting carbon flux through exogenous pyruvate kinase to achieve high ethanol yields in *Clostridium thermocellum*. Metab Eng 15:151-158

Dennis C (1972) Breakdown of cellulose by yeast species. J Gen Microbiol 71:409-411

Desvaux M (2005) *Clostridium cellulolyticum*: model organism of mesophillic cellulolytic clostridia. FEMS Microbiology Reviews 29:741-764

Desvaux M (2006) Unravelling carbon metabolism in anaerobic cellulolytic bacteria. Biotechnol Prog 22:1229-1238

Desvaux M, Guedon E, Petitdemange H (2000) Cellulose catabolism by *Clostridium cellulolyticum* growing in batch culture on defined medium. Appl Environ Microbiol 66:2461-2470

Desvaux M, Guedon E, Petitdemange H (2001a) Carbon flux distribution and kinetics of cellulose fermentation in steady-state continuous cultures of *Clostridium cellulolyticum* on a chemically defined medium. J Bacteriol 183:119-130

Desvaux M, Guedon E, Petitdemange H (2001b) Kinetics and metabolism of cellulose degradation at high substrate concentrations in steady-state continuous cultures of Clostridium cellulolyticum on a chemically defined medium. Appl Environ Microbiol 67:3837-3845

Desvaux M, Guedon E, Petitdemange H (2001c) Metabolic flux in cellulose batch and cellulose-fed continuous cultures of *Clostridium cellulolyticum* in response to acidic environment. Microbiology 147:1461-1471

Desvaux M, Petitdemange H (2001) Flux analysis of the metabolism of *Clostridium cellulolyticum* grown in cellulose-fed continuous culture on a chemically defined medium under ammonium-limited conditions. Appl Environ Microbiol 67:3846-3851

Desvaux Ml, Guedon E, Petitdemange H (2001d) Kinetics and metabolism of cellulose degradation at high substrate concentrations in steady-state continuous cultures of

Clostridium cellulolyticum on a chemically defined medium. Applied and Environmental Microbiology 67:3837-3845

Dien BS, Cotta MA, Jeffries TW (2003) Bacteria engineered for fuel ethanol production: current status. Appl Microbiol Biotechnol 63:258-266

Diez-Gonzalez F, Bond DR, Jennings E, Russell JB (1999) Alternative schemes of butyrate production in *Butyrivibrio fibrisolvens* and their relationship to acetate utilization, lactate production, and phylogeny. Arch Microbiol 171:324-330

Ding YR, Ronimus RS, Morgan HW (2001) *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. J Bacteriol 183:791-794

Dror TW, Morag E, Rolider A, Bayer EA, Lamed R, Shoham Y (2003a) Regulation of the cellulosomal CelS (cel48A) gene of *Clostridium thermocellum* is growth rate dependent. J Bacteriol 185:3042-3048

Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2003b) Regulation of expression of scaffoldin-related genes in *Clostridium thermocellum*. J Bacteriol 185:5109-5116

Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2005) Regulation of major cellulosomal endoglucanases of *Clostridium thermocellum* differs from that of a prominent cellulosomal xylanase. J Bacteriol 187:2261-2266

Dürre P, Kuhn A, Gottschalk G (1986) Treatment with allyl alcohol selects specifically for mutants of *Clostridium acetobutylicum* defective in butanol synthesis. FEMS Microbiol Lett 36:77-81

Dwivedi RC et al. (2008) Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. Anal Chem 80:7036-7042

Elsen S, Dischert W, Colbeau A, Bauer CE (2000) Expression of uptake hydrogenase and molybdenum nitrogenase in Rhodobacter capsulatus is coregulated by the RegB-RegA two-component regulatory system. J Bacteriol 182:2831-2837

Energy Information Administration (2006a) Annual Energy Outlook 2006, DOE/EIA-0383, February 2006. ftp://ftp.eia.doe.gov/forecasting/0484(2006).pdf (accessed May 2013)

Energy Information Administration (2006b) International Energy Outlook 2006, DOE/EIA-0484, June 2006. ftp://ftp.eia.doe.gov/forecasting/0383(2006).pdf (accessed May 2013)

Energy Information Administration (2009) International Energy Outlook 2009, DOE/EIA-0484, May 2009. http://arsiv.setav.org/ups/dosya/25025.pdf (accessed May 2013)

Eriksen NT, Nielsen TM, Iversen N (2008) Hydrogen production in anaerobic and microaerobic *Thermotoga neapolitana*. Biotechnol Lett 30:103-109

Evans DJ, Pickett CJ (2003) Chemistry and the hydrogenases. Chem Soc Rev 32:268-275

Fauque G et al. (1988) The three classes of hydrogenases from sulfate-reducing bacteria of the genus *Desulfovibrio*. FEMS Microbiol Rev 4:299-344

Felsenstein J (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39:783-791

Florens L, Washburn MP (2006) Proteomic analysis by multidimensional protein identification technology. Methods Mol Biol 328:159-175

Freier D, Mothershed CP, Wiegel J (1988) Characterization of *Clostridium thermocellum* JW20. Appl Environ Microbiol 54:204-211

Friedrich B, Buhrke T, Burgdorf T, Lenz O (2005) A hydrogen-sensing multiprotein complex controls aerobic hydrogen metabolism in *Ralstonia eutropha*. Biochem Soc Trans 33:97-101

Garvie EI (1980) Bacterial lactate dehydrogenases. Microbiol Rev 44:106-139

Gilar M, Olivova P, Daly AE, Gebler JC (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. J Sep Sci 28:1694-1703

Girbal L, Soucaille P (1994) Regulation of *Clostridium acetobutylicum* metabolism as revealed by mixed-substrate steady-state continuous cultures: role of NADH/NAD ratio and ATP pool. J Bacteriol 176:6433-6438

Girbal L, Vasconcelos I, Saint-Amans S, Soucaille P (1995) How neutral red modified carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH. FEMS Microbiol Rev 16:151-162

Glusker JP (1971) Aconitase. In: Boyer PD (ed) The Enzymes. Academic Press, New York, pp 413-439

Gold ND, Martin VJ (2007) Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. J Bacteriol 189:6787-6795

Goodwin S, Zeikus JG (1987) Physiological adaptations of anaerobic bacteria to low pH: metabolic control of proton motive force in *Sarcina ventriculi*. J Bacteriol 169:2150-2157

Gottwald M, Gottschalk G (1985) The internal pH of *Clostridium acetobutylicum* and its effect on the shift from acid to solvent formation. Arch Microbiol 143:42-46

Gowen CM, Fong SS (2010) Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of *Clostridium thermocellum*. Biotechnol J 5:759-767

Grupe H, Gottschalk G (1992) Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift Induction. Appl Environ Microbiol 58:3896-3902

Guedon E, Desvaux M, Payot S, Petitdemange H (1999a) Growth inhibition of *Clostridium cellulolyticum* by an inefficiently regulated carbon flow. Microbiology 145 (Pt 8):1831-1838

Guedon E, Desvaux M, Petitdemange H (2000a) Kinetic analysis of *Clostridium cellulolyticum* carbohydrate metabolism: importance of glucose 1-phosphate and glucose 6-phosphate branch points for distribution of carbon fluxes inside and outside cells as revealed by steady-state continuous culture. J Bacteriol 182:2010-2017

Guedon E, Desvaux M, Petitdemange H (2002) Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. Appl Environ Microbiol 68:53-58

Guedon E, Payot S, Desvaux M, Petitdemange H (1999b) Carbon and electron flow in *Clostridium cellulolyticum* grown in chemostat culture on synthetic medium. J Bacteriol 181:3262-3269

Guedon E, Payot S, Desvaux M, Petitdemange H (2000b) Relationships between cellobiose catabolism, enzyme levels, and metabolic intermediates in *Clostridium cellulolyticum* grown in a synthetic medium. Biotechnol Bioeng 67:327-335

Gunsalus IC, Wood AJ (1942) The Dehydrogenation of Alcohols by Streptococci of Group B. J Bacteriol 44:523-528

Guss AM, Lynd LR (2012) *Clostridium thermocellum* strains for enhanced ethanol production and method of their use. In: Organization WIP (ed). WO/2012/109578, U.S.A

Gyan S, Shiohira Y, Sato I, Takeuchi M, Sato T (2006) Regulatory loop between redox sensing of the NADH/NAD(+) ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in Bacillus subtilis. J Bacteriol 188:7062-7071

Haft DH et al. (2001) TIGRFAMs: a protein family resource for the functional identification of proteins. Nucleic Acids Res 29:41-43

Hall AH (1999) Bioedit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98

Hallahan DL, Fernandez VM, Hatchikian EC, Hall DO (1986) Differential inhibition of catalytic sites in *Desulfovibrio gigas* hydrogenase. Biochimie 68:49-54

Hallenbeck PC (2005) Fundamentals of the fermentative production of hydrogen. Water Sci Technol 52:21-29

Hallenbeck PC, Benemann, J.R. (2002) Biological hydrogen production; fundamentals and limiting processes. International Journal of Hydrogen Energy 27:1123-1505

Hamilton-Brehm SD, Mosher JJ, Vishnivetskaya T, Podar M, Carroll S, Allman S, Phelps TJ, Keller M, Elkins JG. *Caldicellulosiruptor obsidiansis* sp. nov., an anaerobic, extremely thermophilic, cellulolytic bacterium isolated from Obsidian Pool, Yellowstone National Park. Appl Environ Microbiol 76:1014-1020

Han B, Higgs RE (2008) Proteomics: from hypothesis to quantitative assay on a single platform. Guidelines for developing MRM assays using ion trap mass spectrometers. Brief Funct Genomic Proteomic 7:340-354

Hattori S (2008) Syntrophic acetate-oxidizing microbes in methanogenic environments. Microbes Environ 23:118-127

Hayashi H, Takagi KI, Fukumura M, Kimura T, Karita S, Sakka K, Ohmiya K. (1997) Sequence of xynC and properties of XynC, a major component of the *Clostridium thermocellum* cellulosome. J Bacteriol 179:4246-4253

He Q, Lokken PM, Chen S, Zhou J (2009) Characterization of the impact of acetate and lactate on ethanolic fermentation by *Thermoanaerobacter ethanolicus*. Bioresour Technol 100:5955-5965

Heinonen JK, Drake HL (1988) Comparative assessment of inorganic pyrophosphate and pyrophosphatase levels of

Escherichia coli, Clostridium pasteurianum, and *Clostridium thermoaceticum*. FEMS Microbiol Lett 52:205-208

Herrero AA, Gomez RF, Roberts MF (1985a) 31P NMR studies of *Clostridium thermocellum*. Mechanism of end product inhibition by ethanol. J Biol Chem 260:7442-7451

Herrero AA, Gomez RF, Roberts MF (1985b) Growth inhibition of *Clostridium thermocellum* by carboxylic acids: a mechanism based on uncoupleing by weak acids. Appl Microbiol Biotechnol 22:53-62

Holwerda EK, Hirst KD, Lynd LR (2012) A defined growth medium with very low background carbon for culturing *Clostridium thermocellum*. J Ind Microbiol Biotechnol 39:943-947

Homann T, Tag C, Biebl H, Deckwer W-T, Schink B (1990) Fermentation of glycerol to 1,3-propanediol by *Klebsiella* and *Citrobacter* strains. Appl Microbiol Biotechnol 33:121-126

Horn M, Krupp F (2009) Earth: The Sequel: The Race to Reinvent Energy and Stop Global Warming. W. W. Norton & Company, New York

Huang L, Forsberg CW, Gibbins LN (1986) Influence of external pH and fermentation products on *Clostridium acetobutylicum* intracellular pH and cellular distribution of fermentation products. Appl Environ Microbiol 51:1230-1234

Iannotti EL, Kafkewitz D, Wolin MJ, Bryant MP (1973) Glucose fermentation products in *Ruminococcus albus* grown in continuous culture with *Vibrio succinogenes*: changes caused by interspecies transfer of H_2. J Bacteriol 114:1231-1240

Ingram LO (1990) Ethanol tolerance in bacteria. Crit Rev Biotechnol 9:305-319

Ingram LO, Conway T, Clark DP, Sewell GW, Preston JF (1987) Genetic engineering of ethanol production in *Escherichia coli*. Appl Environ Microbiol 53:2420-2425

Intergovernmental Panel on Climate Change (2007) Climate Change 2007: Synthesis Report. http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf (accessed May 2013)

International Energy Agency (2006) Energy Technology Perspectives 2006 – Scenarios & Strategies to 2050, 2006. http://www.iea.org/Textbase/npsum/enertech2006SUM.pdf (accessed May 2013)

Islam R, Cicek N, Sparling R, Levin D (2006) Effect of substrate loading on hydrogen production during anaerobic fermentation by *Clostridium thermocellum* 27405. Appl Microbiol Biotechnol 72:576-583

Islam R, Cicek N, Sparling R, Levin D (2009) Influence of initial cellulose concentration on the carbon flow distribution during batch fermentation by *Clostridium thermocellum* ATCC 27405. Appl Microbiol Biotechnol 82:141-148

Ito T, Nakashimada Y, Senba K, Matsui T, Nishio N (2005) Hydrogen and ethanol production from glycerol-containing wastes discharged after biodiesel manufacturing process. J Biosci Bioeng 100:260-265

Jackson BE, Bhupathiraju VK, Tanner RS, Woese CR, McInerney MJ (1999) *Syntrophus aciditrophicus* sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms. Arch Microbiol 171:107-114

Ji XJ, Huang H, Li S, Du J, Lian M (2008) Enhanced 2,3-butanediol production by altering the mixed acid fermentation pathway in *Klebsiella oxytoca*. Biotechnol Lett 30:731-734

Johnson EA, Madia A, Demain AL (1981) Chemically defined minimal medium for growth of the anaerobic cellulolytic thermophile *Clostridium thermocellum*. Appl Environ Microbiol 41:1060-1062

Jolkver E, Emer D, Ballan S, Kramer R, Eikmanns BJ, Marin K (2009) Identification and characterization of a bacterial transport system for the uptake of pyruvate, propionate, and acetate in *Corynebacterium glutamicum*. J Bacteriol 191:940-948

Julliand V, de Vaux A, Millet L, Fonty G (1999) Identification of *Ruminococcus flavefaciens* as the predominant cellulolytic bacterial species of the equine cecum. Appl Environ Microbiol 65:3738-3741

Junelles AM, Janati-Idrissi R, Petitdemange E, Gay R (1988) Iron effect on acetone-butanol fermentation. Curr Microbiol 17:299-303

Junelles AM, Janati-Idrissi R, Petitdemange H, Gay R (1987) Effect of pyruvate on glucose metabolism in *Clostridium acetobutylicum*. Biochimie 69:1183-1190

Kadar Z et al. (2004) Yields from glucose, xylose, and paper sludge hydrolysate during hydrogen production by the extreme thermophile *Caldicellulosiruptor saccharolyticus*. Applied Biochemistry and Biotechnology 113-16:497-508

Kahel-Raifer H et al. (2010) The unique set of putative membrane-associated anti-sigma factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. FEMS Microbiol Lett 308:84-93

Kanai T et al. (2005) Continuous hydrogen production by the hyperthermophilic archaeon, *Thermococcus kodakaraensis* KOD1. J Biotechnol 116:271-282

Kanehisa M et al. (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36:D480-484

Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA (2007) Quantitative multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry ans stable isotope dilution. Mol Cell Proteomics 6:2212-2229

Kataeva IA et al. (2009) Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. Journal of Bacteriology 191:3760-3761

Kell DB, Peck MW, Rodger G, Morris JG (1981) On the permeability to weak acids and bases of the cytoplasmic membrane of *Clostridium pasteurianum*. Biochem Biophys Res Commun 99:81-88

Kengen SW, de Bok FA, van Loo ND, Dijkema C, Stams AJ, de Vos WM (1994) Evidence for the operation of a novel Embden-Meyerhof pathway that involves

ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. J Biol Chem 269:17537-17541

Kim BH, Bellows P, Datta R, Zeikus JG (1984) Control of carbon and electron flow in *Clostridium acetobutylicum* fermentations: Utilization of carbon monoxide to inhibit hydrogen production and to enhance butanol yields. Appl Environ Microbiol 48:764-770

Kleihues L, Lenz O, Bernhard M, Buhrke T, Friedrich B (2000) The H_2 sensor of *Ralstonia eutropha* is a member of the subclass of regulatory [NiFe] hydrogenases. J Bacteriol 182:2716-2724

Kleykamp D (2008) Oil and the world economy. Tamkang Journal of International Affairs 12:51-117

Knappe J, Sawers G (1990) A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. FEMS Microbiol Rev 6:383-398

Kok B (1973) Photosynthesis. In: Gibbs M, Hollaender MA, Kok B, Krampitz LO (eds) Proceedings of the workshop on bio-solar hydrogen conversion, Bethesda

Kosaric N, Duvnjak Z, Farkas A, Sahm H, Bringer-Meyer S (2005) Ethanol. In: Weinheim (ed) Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH

Kraemer JT, Bagley DM (2006) Supersaturation of dissolved H(2) and CO (2) during fermentative hydrogen production with N(2) sparging. Biotechnol Lett 28:1485-1491

Krokhin OV, Spicer V (2009) Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides. Anal Chem 81:9522-9530

Kuhad RC, Gupta R, Singh A (2011) Microbial cellulases and their industrial applications. Enzyme Res 2011:280696

Kumar N, Ghosh A, Das D (2001) Redirection of biochemical pathways for the enhancement of H_2 production by *Enterobacter cloacae*. Biotechnol Lett 23:537-541

Kumar P, Barrett DM, Delwiche MJ, Streve P (2009) Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. Ind Eng Chem Res 48:3713-3729

Kurkin S, Meuer J, Koch J, Hedderich R, Albracht SP (2002) The membrane-bound [NiFe]-hydrogenase (Ech) from Methanosarcina barkeri: unusual properties of the iron-sulphur clusters. Eur J Biochem 269:6101-6111

Kurose N et al. (1989) Characterization of new strains of *Clostridium thermocellum* and the *cel*A gene from a strain. Agric Biol Chem 53:3179-3185

Lacis LS, Lawford HG (1988) Ethanol production from xylose by *Thermoanaerobacter ethanolicus* in batch and continuous culture. Archives of Microbiology 150:48-55

Lacis LS, Lawford HG (1991) *Thermoanaerobacter ethanolicus* growth and product yield from elevated levels of xylose or glucose in continuous cultures. Applied and Environmental Microbiology 57:579-585

Lakhal R et al. (2010) Effect of oxygen and redox potential on glucose fermentation in *Thermotoga maritima* under controlled physicochemical conditions. Int J Microbiol 2010:896510

Lamed R, Zeikus JG (1980a) Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanaerobium brockii*. J Bacteriol 144:569-578

Lamed R, Zeikus JG (1980b) Glucose fermentation pathway of *Thermoanaerobium brockii*. J Bacteriol 141:1251-1257

Lamed R, Zeikus JG (1981) Thermostable, ammonium-activated malic enzyme of *Clostridium thermocellum*. Biochim Biophys Acta 660:251-255

Lamed RJ, Lobos JH, Su TM (1988) Effects of stirring and hydrogen on fermentation products of *Clostridium thermocellum*. Appl Environ Microbiol 54:1216-1221

Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 4:222

Lee CK, Ordal ZJ (1967) Regulatory effect of pyruvate on the glucose metabolism of *Clostridium thermosaccharolyticum*. J Bacteriol 94:530-536

Lee HS, Vermaas WF, Rittmann BE (2010) Biological hydrogen production: prospects and challenges. Trends Biotechnol 28:262-271

Lee YJ, Miyahara T, Noike T (2001) Effect of iron concentration on hydrogen fermentation. Bioresour Technol 80:227-231

Levin D, Carere C, Ramachandran U, Rydzak T, Saunders J (2011a) Fermentative Biofuels: Prospects of Practical Application (Chapter 12). In: Veziroğlu TN (ed) Carbon Neutral Fuels and Energy Carriers Taylor & Francis Group, pp 603-635

Levin DB, Carere C, Ramachandran U, Rydzak T, Saunders J (2011b) Fermentative Biofuels: Prospects of Practical Application In: Veziroğlu TN (ed) In Carbon Neutral Fuels and Energy Carriers. Taylor & Francis Group

Levin DB, Islam R, Cicek N, Sparling R (2006) Hydrogen production by *Clostridium thermocellum* 27405 from cellulosic biomass substrates. International Journal of Hydrogen Energy 31:1496-1503

Levin DB, Pitt L, Love M (2004) Biohydrogen production: prospects and limitations to practical application. International Journal of Hydrogen Energy 29:173-185

Li HF, Knutson BL, Nokes SE, Lynn BC, Flythe MD (2012a) Metabolic control of *Clostridium thermocellum* via inhibition of hydrogenase activity and the glucose transport rate. Appl Microbiol Biotechnol 93:1777-1784

Li Y et al. (2012b) Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. Biotechnol Biofuels 5:2

Liang T-M, Cheng S-S, Wu K-L (2002) Behavioral study on hydrogen fermentation reactor installed with silicone rubber membrane. Int J Hydrogen Energy 27:1157-1165

Lin WR, Lee CC, Hsu JJ, Hamel JF, Demain AL (1998a) Properties of acetate kinase activity in *Clostridium thermocellum* cell extracts. Appl Biochem Biotechnol 69:137-145

Lin WR et al. (1998b) Purification and characterization of acetate kinase from *Clostridium thermocellum*. Tetrahedron 54:15915-15925

Lissolo T, Pulvin S, Thomas D (1984) Reactivation of the hydrogenase from Desulfovibrio gigas by hydrogen. Influence of redox potential. J Biol Chem 259:11725-11729

Liu H, Grot S, Logan BE (2005) Electrochemically assisted microbial production of hydrogen from acetate. Environ Sci Technol 39:4317-4320

Liu H, Sadygov RG, Yates JR, 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193-4201

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

Liu X, Zhu Y, Yang ST (2006) Construction and characterization of ack deleted mutant of *Clostridium tyrobutyricum* for enhanced butyric acid and hydrogen production. Biotechnol Prog 22:1265-1275

Ljunggren M, Willquist K, Zacchi G, van Niel EW (2011) A kinetic model for quantitative evaluation of the effect of hydrogen and osmolarity on hydrogen production by *Caldicellulosiruptor saccharolyticus*. Biotechnol Biofuels 4:31

Lorowitz W, Clark D (1982) *Escherichia coli* mutants with a temperature-sensitive alcohol dehydrogenase. J Bacteriol 152:935-938

Loubiere P, Salou P, Leroy MJ, Lindley ND, Pareilleux A (1992) Electrogenic malate uptake and improved growth energetics of the malolactic bacterium *Leuconostoc oenos* grown on glucose-malate mixtures. J Bacteriol 174:5302-5308

Lovitt RW, Shen GJ, Zeikus JG (1988) Ethanol production by thermophilic bacteria: biochemical basis for ethanol and hydrogen tolerance in *Clostridium thermohydrosulfuricum*. Journal of Bacteriology 170:2809-2815

Lu Y, Zhang C, Zhao H, Xing XH (2012) Improvement of hydrogen productivity by introduction of NADH regeneration pathway in *Clostridium paraputrificum*. Appl Biochem Biotechnol 167:732-742

Luers F, Seyfried M, Daniel R, Gottschalk G (1997) Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression of the gene encoding 1,3-propanediol dehydrogenase. FEMS Microbiol Lett 154:337-345

Luo HW, Zhang H, Suzuki T, Hattori S, Kamagata Y (2002) Differential expression of methanogenesis genes of *Methanothermobacter thermoautotrophicus* (formerly *Methanobacterium thermoautotrophicum*) in pure culture and in cocultures with fatty acid-oxidizing syntrophs. Appl Environ Microbiol 68:1173-1179

Lynd LR, Cruz CH (2010) Make way for ethanol. Science 330:1176

Lynd LR, Grethlein HE (1987) Hydrolysis of dilute acid pretreated mixed hardwood and purified microcrystalline cellulose by cell-free broth from *Clostridium thermocellum*. Biotechnol Bioeng 29:92-100

Lynd LR, Grethlein HE, Wolkin RH (1989) Fermentation of cellulosic substrates in batch and continuous culture by *Clostridium thermocellum*. Appl Environ Microbiol 55:3131-3139

Lynd LR, van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of cellulosic biomass: an update. Current Opinion in Biotechnology 16:577-583

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506-577, table of contents

Maeda T, Sanchez-Torres V, Wood TK (2007) Enhanced hydrogen production from glucose by metabolically engineered Escherichia coli. Appl Microbiol Biotechnol 77:879-890

Magnusson L, Cicek N, Sparling R, Levin D (2009) Continuous hydrogen production during fermentation of alpha-cellulose by the thermophillic bacterium *Clostridium thermocellum*. Biotechnol Bioeng 102:759-766

Maness P-C, Thammannagowda S (2008) Fermentation and electrohydrogenic approaches to hydrogen production. http://www.hydrogen.energy.gov/pdfs/progress08/ii_f_6_maness.pdf

Markowitz VM et al. (2006) The integrated microbial genomes (IMG) system. Nucleic Acids Res 34:D344-348

McLaughlin MJ, Sullivan RP, Zhao H (2009) Directed enzyme evolution and high-throughput screening. In: Tao J, Lin G-Q, Liese A (eds) Biocatalysts for the

Pharmaceutical Industry: Discovery, Development, and Manufacturing John Wiley & Sons, Ltd, Chichester

McQueen P et al. (2012) Information-dependent LC-MS/MS acquisition with exclusion lists potentially generated on-the-fly: Case study using a whole cell digest of *Clostridium thermocellum*. Proteomics 12:1-10

Meinecke B, Bertram J, Gottschalk G (1989) Purification and characterization of the pyruvate-ferredoxin oxidoreductase from *Clostridium acetobutylicum*. Arch Microbiol 152:244-250

Membrillo-Hernandez J, Echave P, Cabiscol E, Tamarit J, Ros J, Lin EC (2000) Evolution of the adhE gene product of *Escherichia coli* from a functional reductase to a dehydrogenase. Genetic and biochemical studies of the mutant proteins. J Biol Chem 275:33869-33875

Mertens E (1993) ATP versus pyrophosphate: glycolysis revisited in parasitic protists. Parasitol Today 9:122-126

Mertens E, De Jonckheere J, Van Schaftingen E (1993) Pyrophosphate-dependent phosphofructokinase from the amoeba *Naegleria fowleri*, an AMP-sensitive enzyme. Biochem J 292 (Pt 3):797-803

Meyer CL, McLaughlin JK, Papoutsakis T (1985) The effect of CO on growth and product formation in batch cultures of *Clostridium acetobutylicum*. Biotechnol Lett 7:37-42

Meyer CL, Roos JW, Papoutsakis T (1986) Carbon monoxide gas leads to alcohol production and butyrate uptake without acetone formation in continuous of *Clostridium acetobutylicum*. Appl Microbiol Biotechnol 24:159-167

Meyer J (2007) [FeFe] hydrogenases and their evolution: a genomic perspective. Cell Mol Life Sci 64:1063-1084

Michaelis L, Hill ES (1933) The viologen indicators. J Gen Physiol 16:859-873

Mishra S, Beguin P, Aubert JP (1991) Transcription of *Clostridium thermocellum* endoglucanase genes celF and celD. J Bacteriol 173:80-85

Mizuno O, Dinsdale R, Hawkes FR, Hawkes DL, Noike T (2000) Enhancement of hydrogen production from glucose by nitrogen gas sparging. Bioresour Technol 73:59-65

Moat AG, Foster JW, Spector MP (2002) Microbial Physiology (4th ed.) Wiley-Liss Inc., New York. p. 414

Morais PB, Rosa CA, Linardi VR, Carazza F, Nonato EA (1996) Production of fuel alcohol by *Saccharomyces* strains from tropical habitats. Biotechnol Lett 18:1351-1356

Morimoto K, Kimura T, Sakka K, Ohmiya K (2005) Overexpression of a hydrogenase gene in *Clostridium paraputrificum* to enhance hydrogen gas production. FEMS Microbiol Lett 246:229-234

Mountfort DO, Kaspar HF (1986) Palladium-mediated hydrogenation of unsaturated hydrocarbons with hydrogen gas released during anaerobic cellulose degradation. Appl Environ Microbiol 52:744-750

Mukund S, Adams MW (1995) Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Biol Chem 270:8389-8392

Munro SA, Zinder SH, Walker LP (2009) The fermentation stoichiometry of *Thermotoga neapolitana* and influence of temperature, oxygen, and pH on hydrogen production. Biotechnol Prog 25:1035-1042

Nair RV, Bennett GN, Papoutsakis ET (1994) Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. J Bacteriol 176:871-885

Nakajima A, Ueda R (1999) Improvement of microalgal photosynthetic productivity by reducing the content of light harvesting pigment. J Appl Phycol 11:195-201

Nakajima A, Ueda R (2000) The effect of reducing light-harvesting pigment on marine microalgal productivity. J Appl Phycol 12:285-290

Nakase T, Suzuki M, Takashima M, Hamamoto M, Hatano T, Fukui S (1994) A taxonomic study on cellulolytic yeast and yeast-like microorganisms isolated in Japan. I: Ascomycetous yeast genera *Candida* and *Williopsis*, and a yeast-like genus *Prototheca*. J Gen Appl Microbiol 40:519-531

Nandi R, Sengupta S (1998) Microbial production of hydrogen: an overview. Crit Rev Microbiol 24:61-84

Nataf Y et al. (2009) *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. Proc Natl Acad Sci U S A 107:18646-18651

Nath K, Das D (2004) Improvement of fermentative hydrogen production: various approaches. Appl Microbiol Biotechnol 65:520-529

Nath K, Kumar A, Das D (2006) Effect of some environmental parameters on fermentative hydrogen production by *Enterobacter cloacae* DM11. Can J Microbiol 52:525-532

National Research Council (1999) Review of the research strategy for biomass-derived transportation fuels. National Academy Press, Washington, DC

Newcomb M, Chen CY, Wu JH (2007) Induction of the celC operon of *Clostridium thermocellum* by laminaribiose. Proc Natl Acad Sci U S A 104:3747-3752

Newcomb M, Millen J, Chen CY, Wu JH (2011) Co-transcription of the celC gene cluster in *Clostridium thermocellum*. Appl Microbiol Biotechnol 90:625-634

Ng TK, Ben-Bassat A, Zeikus JG (1981) Ethanol production by thermophilic bacteria: Fermentation of cellulosic substrates by cocultures of *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. Appl Environ Microbiol 41:1337-1343

Ng TK, Zeikus JG (1982) Differential metabolism of cellobiose and glucose by *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. J Bacteriol 150:1391-1399

Nguyen TA, Han SJ, Kim JP, Kim MS, Sim SJ (2010) Hydrogen production of the hyperthermophilic eubacterium, *Thermotoga neapolitana* under N_2 sparging condition. Bioresour Technol 101 Suppl 1:S38-41

Nguyen TAD, Pyo Kim J, Sun Kim M, Kwan Oh Y, Sim SJ (2008) Optimization of hydrogen production by hyperthermophilic eubacteria, *Thermotoga maritima* and

Thermotoga neapolitana in batch fermentation. International Journal of Hydrogen Energy 33:1483-1488

Nguyen TN, Borges KM, Romano AH, Noll KM (2001) Differential gene expression in *Thermotoga neapolitana* in response to growth substrate. FEMS Microbiol Lett 195:79-83

Oak Ridge National Laboratory (2007) *Clostridium thermocellum* ATCC 27405 analysis files. http://genome.ornl.gov/microbial/cthe/ (accessed May 2008)

Ohta K, Beall DS, Mejia JP, Shanmugam KT, Ingram LO (1991a) Genetic improvement of *Escherichia coli* for ethanol production: chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. Appl Environ Microbiol 57:893-900

Ohta K, Beall DS, Mejia JP, Shanmugam KT, Ingram LO (1991b) Metabolic engineering of *Klebsiella oxytoca* M5A1 for ethanol production from xylose and glucose. Appl Environ Microbiol 57:2810-2815

Old WM et al. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4:1487-1502

Ouhib-Jacobs O, Lindley ND, Schmitt P, Clavel T (2009) Fructose and glucose mediates enterotoxin production and anaerobic metabolism of *Bacillus cereus* ATCC14579(T). J Appl Microbiol 107:821-829

Özkan M, Ylmaz, E., Lynd, L.R., Özcengiz, G. (2004) Cloning and expression of the *Clostridium thermocellum* L-lactate dehydrogenase in *Escherichia coli* and enzyme characterization. Canadian Journal of Microbiology 50:845-851

Park W, Hyun SH, Oh SE, Logan BE, Kim IS (2005) Removal of headspace CO_2 increases biological hydrogen production. Environ Sci Technol 39:4416-4420

Patel GB, Khan AW, Agnew BJ, Colvin JR (1982) Metabolism of *Acetivibrio cellulolyticus* during optimized growth on glucose, cellobiose and cellulose. Eur J Appl Microbiol Biotechnol 16:212-218

Patni NJ, Alexander JK (1971a) Catabolism of fructose and mannitol in *Clostridium thermocellum*: presence of phosphoenolpyruvate: fructose phosphotransferase, fructose 1-phosphate kinase, phosphoenolpyruvate: mannitol phosphotransferase, and mannitol 1-phosphate dehydrogenase in cell extracts. J Bacteriol 105:226-231

Patni NJ, Alexander JK (1971b) Utilization of glucose by *Clostridium thermocellum*: presence of glucokinase and other glycolytic enzymes in cell extracts. J Bacteriol 105:220-225

Pauss A, Andre G, Perrier M, Guiot SR (1990) Liquid-to-gas mass transfer in anaerobic processes: Inevitable transfer limitations of methane and hydrogen in the biomethanation process. Appl Environ Microbiol 56:1636-1644

Payot S, Guedon E, Gelhaye E, Petitdemange H (1999) Induction of lactate production associated with a decrease in NADH cell content enables growth resumption of *Clostridium cellulolyticum* in batch cultures on cellobiose. Res Microbiol 150:465-473

Pei J et al. (2010) *Thermoanaerobacter* spp. control ethanol pathway via transcriptional regulation and versatility of key enzymes. Metab Eng 12:420-428

Pei J et al. (2011) The mechanism for regulating ethanol fermentation by redox levels in *Thermoanaerobacter ethanolicus*. Metab Eng 13:186-193

Penfold DW, Sargent F, Macaskie E (2006) Inactivation of the *Escherichia coli* K-12 twin-arginine translocation system promotes increased hydrogen production. FEMS Microbiol Lett 262:135-137

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567

Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC (1998) X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. Science 282:1853-1858

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45

Pfenninger-Li XD, Albracht SP, van Belzen R, Dimroth P (1996) NADH:ubiquinone oxidoreductase of *Vibrio alginolyticus*: purification, properties, and reconstitution of the Na+ pump. Biochemistry 35:6233-6242

Pfitzer S, Decker K (1972) Quantitative determination of adenine nucleotides in growing cell cultures. Zentralbl Bakteriol Orig A 220:428-431

Pierik AJ, Roseboom W, Happe RP, Bagley KA, Albracht SP (1999) Carbon monoxide and cyanide as intrinsic ligands to iron in the active site of [NiFe]-hydrogenases. NiFe(CN)2CO, Biology's way to activate H2. J Biol Chem 274:3331-3337

Posewitz MC, King PW, Smolinski SL, Zhang L, Seibert M, Ghirardi ML (2004) Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. J Biol Chem 279:25711-25720

Preiss J (1984) Bacterial glycogen synthesis and its regulation. Annu Rev Microbiol 38:419-458

Preiss J, Romeo T (1989) Physiology, biochemistry and genetics of bacterial glycogen synthesis. Adv Microb Physiol 30:183-238

Rachman MA, Furutani Y, Nakashimada Y, Kakizono T, Nishio N (1997) Enhanced hydrogen production in altered mixed acid fermentation of glucose by *Enterobacter aerogenes*. J Ferment Bioeng 83:358-363

Raman B, McKeown CK, Rodriguez M, Jr., Brown SD, Mielenz JR (2011) Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. BMC Microbiol 11:134

Raman B et al. (2009) Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. PLoS One 4:e5271

Rando RR (1972) In situ generation of irreversible enzyme inhibitors. Nat New Biol 237:53

Rao G, Mutharasan R (1987) Altered electron flow in continuous cultures of *Clostridium acetobutylicum* induced by viologen dyes. Appl Environ Microbiol 53:1232-1235

Rao G, Ward PJ, Mutharasan R (1987) Manipulation of end-product distribution in strict anaerobes. Ann N Y Acad Sci 506:76-83

Rappsilber J, Ryder U, Lamond AI, Mann M (2002) Large-scale proteomic analysis of the human spliceosome. Genome Res 12:1231-1245

Ren Z, Ward TE, Logan BE, Regan JM (2007) Characterization of the cellulolytic and hydrogen-producing activities of six mesophilic Clostridium species. J Appl Microbiol 103:2258-2266

Renewable Fuels Association (2012) Accelerating Industry Innovation - 2012 Ethanol Industry Outlook". Renewable Fuels Association. http://ethanolrfa.3cdn.net/d4ad995ffb7ae8fbfe_1vm62ypzd.pdf (accessed May 2013)

Rhee SK, Pack MY (1980) Effect of environmental pH on fermentation balance of *Lactobacillus bulgaricus*. J Bacteriol 144:217-221

Rhodes CJ (2008) The oil question: nature and prognosis. Sci Prog 91:317-375

Riederer A et al. (2011) Global gene expression patterns in *Clostridium thermocellum* as determined by microarray analysis of chemostat cultures on cellulose or cellobiose. Appl Environ Microbiol 77:1243-1253

Robinson JR, Sagers RD (1972) Phosphotransacetylase from *Clostridium acidiurici*. J Bacteriol 112:465-473

Rodrigue J-P, Comtois C (2013) The Geography of Transport systems. In: Routledge JP (ed), 3rd edn, New York

Rothstein DM (1986) *Clostridium thermosaccharolyticum* strain deficient in acetate production. J Bacteriol 165:319-320

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-386

Ruehl C (2011) BP Energy Outlook 2030. http://www.seec.surrey.ac.uk/Events/EventDocuments/ChristofRuehlSlides.pdf (accessed May 2013)

Rupprecht J, Hankamer B, Mussgnug JH, Ananyev G, Dismukes C, Kruse O (2006) Prespectives and advances of biological H_2 production in microorganisms. Appl Microbiol Biotechnol 72:422-449

Rydzak T, Levin DB, Cicek N, Sparling R (2009) Growth phase-dependant enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405. J Biotechnol 140:169-175

Rydzak T, Levin DB, Cicek N, Sparling R (2011) End-product induced metabolic shifts in *Clostridium thermocellum* ATCC 27405. Appl Microbiol Biotechnol 92:199-209

Rydzak T et al. (2012) Proteomic analysis of *Clostridium thermocellum* core metabolism: Relative protein expression profiles and growth phase-dependent changes in protein expression. BMC Microbiol 12:214

Saint-Amans S, Girbal L, Andrade J, Ahrens K, Soucaille P (2001) Regulation of carbon and electron flow in *Clostridium butyricum* VPI 3266 grown on glucose-glycerol mixtures. J Bacteriol 183:1748-1754

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425

Sander R (1999) Compilation of Henry's law constants for inorganic and organic species of potential importance in environmental chemistry (Version 3). Max-Planck Institute of Chemistry, Mainz, Germany

Sauer U, Eikmanns BJ (2005) The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. FEMS Microbiology Reviews 29:765-794

Sawers G (1994) The hydrogenases and formate dehydrogenases of Escherichia coli. Antonie Van Leeuwenhoek 66:57-88

Sawers G, Bock A (1988) Anaerobic regulation of pyruvate formate-lyase from *Escherichia coli* K-12. J Bacteriol 170:5330-5336

Schäfer T, Schönheit P (1991) Pyruvate metabolism of the hyperthermophilic archaebacterium *Pyrococcus furiosus*. Arch Microbiol:366-377

Schicho RN, Ma K, Adams MW, Kelly RM (1993) Bioenergetics of sulfur reduction in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 175:1823-1830

Schmehl M et al. (1993) Identification of a new class of nitrogen fixation genes in Rhodobacter capsulatus: a putative membrane complex involved in electron transport to nitrogenase. Mol Gen Genet 241:602-615

Schmidt S, Biegel E, Muller V (2009) The ins and outs of Na(+) bioenergetics in Acetobacterium woodii. Biochim Biophys Acta 1787:691-696

Schröder C, Selig M, Schönheit P (1994) Glucose fermentation to acetate, CO_2 and H_2 in the anaerobic hyperthermophilic eubacterium *Thermotoga maritima*: involvement of the Embden-Meyerhof pathway. Archives of Microbiology 161:460-470

Schut GJ, Adams MW (2009) The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. J Bacteriol 191:4451-4457

Schwarz WH (2001) The cellulosome and cellulose degradation by anaerobic bacteria. Appl Microbiol Biotechnol 56:634-649

Schwarz WH (2003) A list of cellulolytic bacteria page. http://www.wzw.tum.de/mbiotec/cellmo.htm (accessed May 2013)

Scopes RK, Griffiths-Smith K (1986) Fermentation capabilities of *Zymomonas mobilis* glycolytic enzymes. Biotechnol Lett 8:653-656

Searchinger T et al. (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. Science 319:1238-1240

Shaw AJ, Hogsett DA, Lynd LR (2009) Identification of the [FeFe]-hydrogenase responsible for hydrogen generation in *Thermoanaerobacterium saccharolyticum* and demonstration of increased ethanol yield via hydrogenase knockout. J Bacteriol 191:6457-6464

Shaw AJ et al. (2008) Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. Proc Natl Acad Sci U S A 105:13769-13774

Sims R, Taylor M, Saddler JN (2008) From 1st to 2nd Generation Biofuel Technologies: An overview of current industry and RD&D activities. In: Bioenergy IEAI (ed)

Soboh B, Linder D, Hedderich R (2004) A multisubunit membrane-bound [NiFe] hydrogenase and an NADH-dependent Fe-only hydrogenase in the fermenting bacterium *Thermoanaerobacter tengcongensis*. Microbiology 150:2451-2463

Sparling R, Carere C, Rydzak T, Schellenberg J, Levin D (2012) Thermodynamic and Biochemical Aspect of Hydrogen Production by Dark Fermentation (Chapter 10) In: Azbar N, Levin D (eds) State of the Art and Progress in Production of Biohydrogen. Bentham Science Publishing, pp 160-188

Sparling R, Islam R, Cicek N, Carere C, Chow H, Levin DB (2006) Formate synthesis by *Clostridium thermocellum* during anaerobic fermentation. Can J Microbiol 52:681-688

Spinnler HE, Lavigne B, Blanchere H (1986) Pectinolytic activity of *Clostridium thermocellum:* Its use for anaerobic fermentation of sugar beet pulp. Appl Microbiol Biotechnol 23:434-437

Sridhar J, Eiteman MA, Wiegel JW (2000) Elucidation of enzymes in fermentation pathways used by *Clostridium thermosuccinogenes* growing on inulin. Appl Environ Microbiol 66:246-251

Srinivasan S (2009) The food v. fuel debate: A nuanced view of incentive structures. Renewable Energy 34:950-954

Stevens BHJ, Payne J (1977) Cellulase and xylanase production by yeasts of the genus *Trichosporon*. J Gen Microbiol 100:381-393

Stevenson DM, Weimer PJ (2005) Expression of 17 Genes in *Clostridium thermocellum* ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. Applied and Environmental Microbiology 71:4672-4678

Strobel HJ (1995) Growth of the thermophilic bacterium *Clostridium thermocellum* in continuous culture. Current Microbiology 31:210-214

Strobel HJ, Caldwell FC, Dawson KA (1995) Carbohydrate transport by the anaerobic thermophile *Clostridium thermocellum* LQRI. Appl Environ Microbiol 61:4012-4015

Susskind BM, Warren LG, Reeves RE (1982) A pathway for the interconversion of hexose and pentose in the parasitic amoeba *Entamoeba histolytica*. Biochem J 204:191-196

Svetlichnyi VA, Svetlichnaya TP, Chernykh NA, Zavarzin GA (1990) *Anaerocellum thermophilum* gen. nov. sp. nov: an extremely thermophilic cellulolytic eubacterium isolated from hot-springs in the Valley of Geysers. Microbiology 59:598-604

Takahata Y, Nishijima M, Hoaki T, Maruyama T (2001) *Thermotoga petrophila* sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the Kubiki oil reservoir in Niigata, Japan. Int J Syst Evol Microbiol 51:1901-1909

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731-2739

Tang YJ et al. (2009) Analysis of metabolic pathways and fluxes in a newly discovered thermophilic and ethanol-tolerant Geobacillus strain. Biotechnol Bioeng 102:1377-1386

Tatusov RL et al. (2003) The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4:41

Thauer RK, Jungermann K, Decker K (1977) Energy conservation in chemotrophic anaerobic bacteria. Bacteriol Rev 41:100-180

The World Bank (2012) Fossil Fuel Energy Consumption (% of total). http://data.worldbank.org/indicator/EG.USE.COMM.FO.ZS (accessed May 2013)

Tokuda H, Nakamura T, Unemoto T (1981) Potassium ion is required for the generation of pH-dependent membrane potential and delta pH by the marine bacterium *Vibrio alginolyticus*. Biochemistry 20:4198-4203

Tokuda H, Unemoto T (1981) A respiration-dependent primary sodium extrusion system functioning at alkaline pH in the marine bacterium Vibrio alginolyticus. Biochem Biophys Res Commun 102:265-271

Trinh CT, Li J, Blanch HW, Clark DS (2011) Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol. Appl Environ Microbiol 77:4894-4904

Tripathi SA et al. (2010) Development of pyrF-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a pta mutant. Appl Environ Microbiol 76:6591-6599

Tripathi SA et al. (2012) Development of pyrF-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a pta mutant. Appl Environ Microbiol 76:6591-6599

Tyurin MV, Desai SG, Lynd LR (2004) Electrotransformation of *Clostridium thermocellum*. Appl Environ Microbiol 70:883-890

Tyurin MV, Sullivan CR, Lynd LR (2005) Role of spontaneous current oscillations during high-efficiency electrotransformation of thermophilic anaerobes. Appl Environ Microbiol 71:8069-8076

U.S. Department of Energy (2007) Alternative Fuels data Center. Energy Independence and Security Act of 2007. http://www.afdc.energy.gov/laws/eisa (accessed May 2013)

van de Werken HJ et al. (2008) Hydrogenomics of the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. Appl Environ Microbiol 74:6720-6729

Vardar-Schara G, Maeda T, Wood TK (2008) Metabolically engineered bacteria for producing hydrogen via fermentation. Microb Biotechnol 1:107-125

Vasconcelos I, Girbal L, Soucaille P (1994) Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. J Bacteriol 176:1443-1450

Veit A, Akhtar MK, Mizutani T, Jones PR (2008) Constructing and testing the thermodynamic limits of synthetic NAD(P)H:H2 pathways. Microb Biotechnol 1:382-394

Vey JL, Yang J, Li M, Broderick WE, Broderick JB, Drennan CL (2008) Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme. Proc Natl Acad Sci U S A 105:16137-16141

Veziroğlu TN (1997) Hydrogen movement and the next action: fossil fuels industry and sustainability economics. Int J Hydrogen Energy 22:551-556

Vignais PM (2008) Hydrogenases and H(+)-reduction in primary energy conservation. Results Probl Cell Differ 45:223-252

Vignais PM, Billoud, B., Meyer, J. (2001) Classification and phylogeny of hydrogenases. FEMS Microbiology Reviews 25:455-501

Wagner N, Tran QH, Richter H, Selzer PM, Unden G (2005) Pyruvate fermentation by *Oenococcus oeni* and *Leuconostoc mesenteroides* and role of pyruvate dehydrogenase in anaerobic fermentation. Appl Environ Microbiol 71:4966-4971

Wang J, Wan W (2008) Influence of Ni^{2+} concentration on biohydrogen production. Bioresour Technol 99:8864-8868

Wang M, Wu M, Huo H (2007) Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. Environ Res Lett 2:024001

Wang S, Huang H, Moll J, Thauer RK (2010) NADP+ reduction with reduced ferredoxin and NADP+ reduction with NADH are coupled via an electron-bifurcating enzyme complex in *Clostridium kluyveri*. J Bacteriol 192:5115-5123

Warnick TA, Methe BA, Leschine SB (2002) *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. Int J Syst Evol Microbiol 52:1155-1160

Weimer PJ, Zeikus JG (1977) Fermentation of cellulose and cellobiose by *Clostridium thermocellum* in the absence of Methanobacterium thermoautotrophicum. Appl Environ Microbiol 33:289-297

Wells JE, Russell JB, Shi Y, Weimer PJ (1995) Cellodextrin efflux by the cellulolytic ruminal bacterium *Fibrobacter succinogenes* and its potential role in the growth of nonadherent bacteria. Appl Environ Microbiol 61:1757-1762

Wheals AE, Basso LC, Alves DM, Amorim HV (1999) Fuel ethanol after 25 years. Trends Biotechnol 17:482-487

Wiegel J, Ljungdahl LG (1981) *Thermoanaerobacter ethanolicus* gen. nov., spec. nov., a new, extreme thermophilic, anaerobic bacterium. Archives of Microbiology 128:343-348

Willquist K, van Niel EW (2010) Lactate formation in *Caldicellulosiruptor saccharolyticus* is regulated by the energy carriers pyrophosphate and ATP. Metab Eng 12:282-290

Willquist K, Zeidan AA, van Niel EW (2010) Physiological characteristics of the extreme thermophile *Caldicellulosiruptor saccharolyticus*: an efficient hydrogen cell factory. Microb Cell Fact 9:89

Winkelman JW, Clark DP (1984) Proton suicide: general method for direct selection of sugar transport- and fermentation-defective mutants. J Bacteriol 160:687-690

Xing D, Ren N, Li Q, Lin M, Wang A, Zhao L (2006) *Ethanoligenens harbinense* gen. nov., sp. nov., isolated from molasses wastewater. Int J Syst Evol Microbiol 56:755-760

Xue Y, Xu Y, Liu Y, Ma Y, Zhou P (2001) *Thermoanaerobacter tengcongensis* sp. nov., a novel anaerobic, saccharolytic, thermophilic bacterium isolated from a hot spring in Tengcong, China. Int J Syst Evol Microbiol 51:1335-1341

Yamamoto Y, Sato Y, Takahashi-Abbe S, Takahashi N, Kizaki H (2000) Characterization of the *Streptococcus mutans* pyruvate formate-lyase (PFL)-

activating enzyme gene by complementary reconstitution of the In vitro PFL-reactivating system. Infect Immun 68:4773-4777

Yang SJ et al. (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "*Anaerocellum thermophilum*" DSM 6725. Applied and Environmental Microbiology 75:4762-4769

Yin Y, Kirsch JF (2007) Identification of functional paralog shift mutations: conversion of *Escherichia coli* malate dehydrogenase to a lactate dehydrogenase. Proc Natl Acad Sci U S A 104:17353-17357

Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H (2005) Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing Escherichia coli strains. Appl Environ Microbiol 71:6762-6768

Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H (2007) Efficient induction of formate hydrogen lyase of aerobically grown Escherichia coli in a three-step biohydrogen production process. Appl Microbiol Biotechnol 74:754-760

Zaunmuller T, Eichert M, Richter H, Unden G (2006) Variations in the energy metabolism of biotechnologically relevant heterofermentative lactic acid bacteria during growth on sugars and organic acids. Appl Microbiol Biotechnol 72:421-429

Zhang YH, Lynd LR (2005a) Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci U S A 102:7321-7325

Zhang YH, Lynd LR (2005b) Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. J Bacteriol 187:99-106

Zheng XJ, Yu HQ (2005) Inhibitory effects of butyrate on biological hydrogen production with mixed anaerobic cultures. J Environ Manage 74:65-70

Zhu J, Shimizu K (2005) Effect of a single-gene knockout on the metabolic regulation in *Escherichia coli* for D-lactate production under microaerobic condition. Metab Eng 7:104-115

Zhu W, Smith JW, Huang CM (2010) Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol 2010:840518

Zieske LR (2006) A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J Exp Bot 57:1501-1508

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson B, Vogel H (eds) Evolving Genes and Proteins. New York, New York, pp 97–166

Zverlov VV, Schwarz WH (2008) Bacterial cellulose hydrolysis in anaerobic environmental subsystems: *Clostridium thermocellum* and *Clostridium stercorarium*, thermophilic plant-fiber degraders. Ann N Y Acad Sci 1125:298-307

Zybailov B, Coleman MK, Florens L, Washburn MP (2005) Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. Anal Chem 77:6218-6224

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. J Proteome Res 5:2339-2347

Appendix

Supplementary Table 1: Cofactor specificity (ATP or PP_i) of phosphofructokinases based on sequence alignments. Alignments of key residues determining ATP or PP_i specificity, as determined by Bapteste *et al*. (Bapteste et al. 2003) and Bielen *et al*. (Bielen et al. 2010), were performed using BioEdit v.7.0.9.0. The *P. furiosus* and *Th. kodakarensis* genes are very distinct (different COG and different KO) and are annotated as Archaeal phosphofructokinases.

6-Phosphofructokinase Genes				
Organism	**Locus Tag**	**AA #104**	**AA #124**	**ATP or PPi Dependence**
Bacillus cereus ATCC 14579	BC4600	G	G	ATP
Caldanaerobacter subterraneus subsp. tengcongensis MB4	TTE1816	G	G	ATP
Ethanoligenens harbinense YUAN-3T, DSM 18485	Ethha_1347	G	G	ATP
Geobacillus thermoglucosidasius C56-YS93	Geoth_0897	G	G	ATP
P. furiosus DSM 3638	PF0312	S	E	ATP
	PF1784	L	F	ATP
Thermococcus kodakaraensis KOD1	TK0376	L	F	ATP
	TK1110	S	E	ATP
Thermotoga maritima MSB8	TM0209	G	A	ATP
	TM0289	D	K	PPi
Thermotoga neapolitana DSM 4359	CTN_0395	D	K	PPi
	CTN_0476	G	A	ATP
Thermotoga petrophila RKU-1	Tpet_0623	D	K	PPi
	Tpet_0715	G	A	ATP
Caldicellulosiruptor bescii DSM 6725	Athe_1265	G	G	ATP
	Athe_1824	D	K	PPi
Caldicellulosiruptor saccharolyticus DSM 8903	Csac_1830	G	G	ATP
	Csac_2366	D	K	PPi
Clostridium cellulolyticum H10	Ccel_2223	D	K	PPi
	Ccel_2612	G	G	ATP
Clostridium phytofermentans ISDg	Cphy_0336	G	G	ATP
	Cphy_3345	D	K	PPi
Clostridium thermocellum ATCC 27405	Cthe_0347	D	K	PPi
	Cthe_1261	G	G	ATP
Clostridium thermocellum DSM 4150	CtherDRAFT_1670	G	G	ATP
Thermoanaerobacter pseudethanolicus 39E	Teth39_0494	D	K	PPi
	Teth39_0683	G	G	ATP

Supplementary Table 2: Relative abundance indexes and changes in protein expression levels of proteins involved cellulose and hemicellulose breakdown and transport. Shotgun and 4-plex 2D-HPLC-MS/MS data identifying protein relative abundance indexes, changes in protein expression, and vector differences indicating statistical relevance of changes in expression are provided. Given the excessive size of the table it could not be reproduced in this thesis, but can be accessed as an excel file through the following link: http://www.biomedcentral.com/content/supplementary/1471-2180-12-214-s3.xlsm

Supplementary Table 3: Relative abundance indexes and changes in protein expression levels of proteins involved in conversion of phosphoenolpyruvate to end-products. Shotgun and 4-plex 2D-HPLC-MS/MS data identifying protein relative abundance indexes, changes in protein expression, and vector differences indicating statistical relevance of changes in expression are provided. Given the excessive size of the table it could not be reproduced in this thesis, but can be accessed as an excel file through the following link: http://www.biomedcentral.com/content/supplementary/1471-2180-12-214-s4.xlsm

Supplementary Table 4: List of selected peptides, transitions states, corresponding peak areas, and retention times used to determine changes in protein expression in response to exogenous pyruvate using MRM. Given the excessive size of the table it

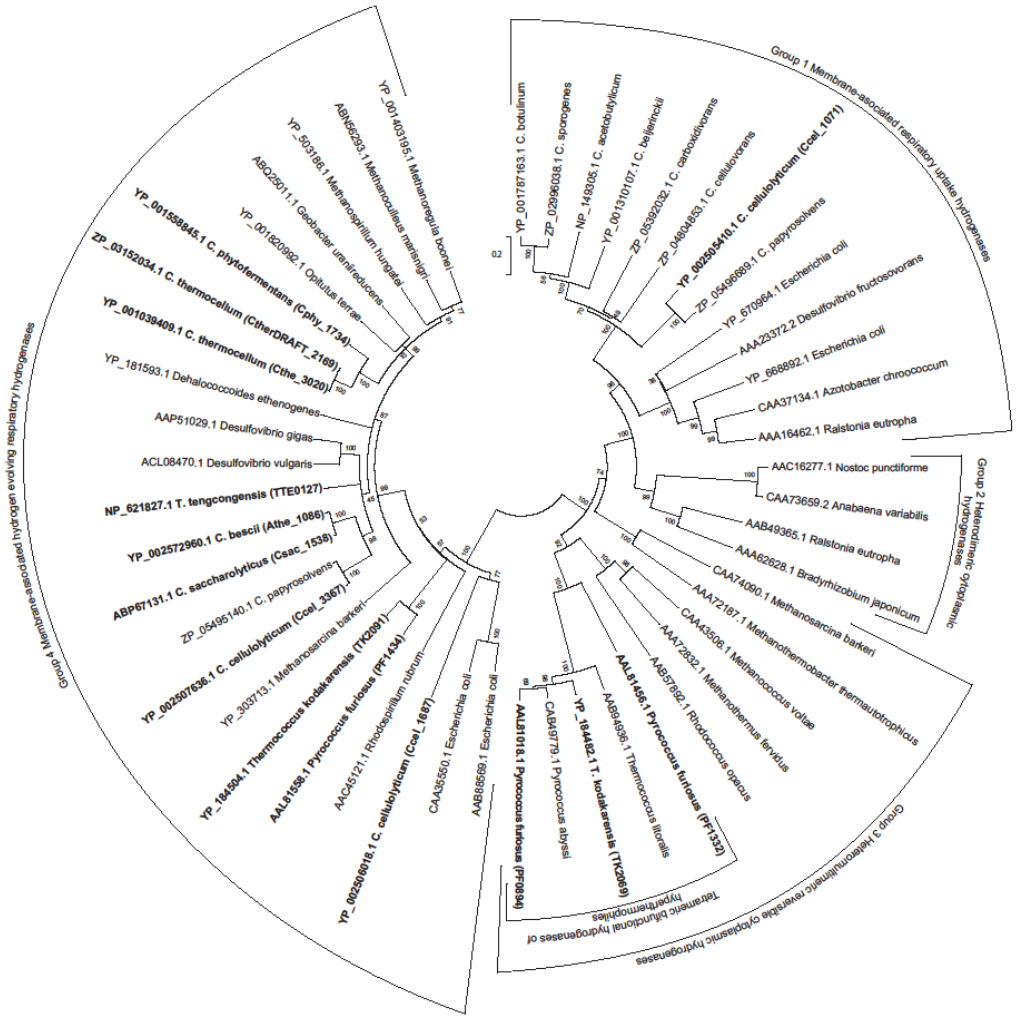
could not be reproduced in this thesis, but can be accessed as an excel file through the following link:

https://www.dropbox.com/s/vzm05ghitwpaq74/MRM%20Supplementary%20File%201%20-%20Pyr%20data.xlsx

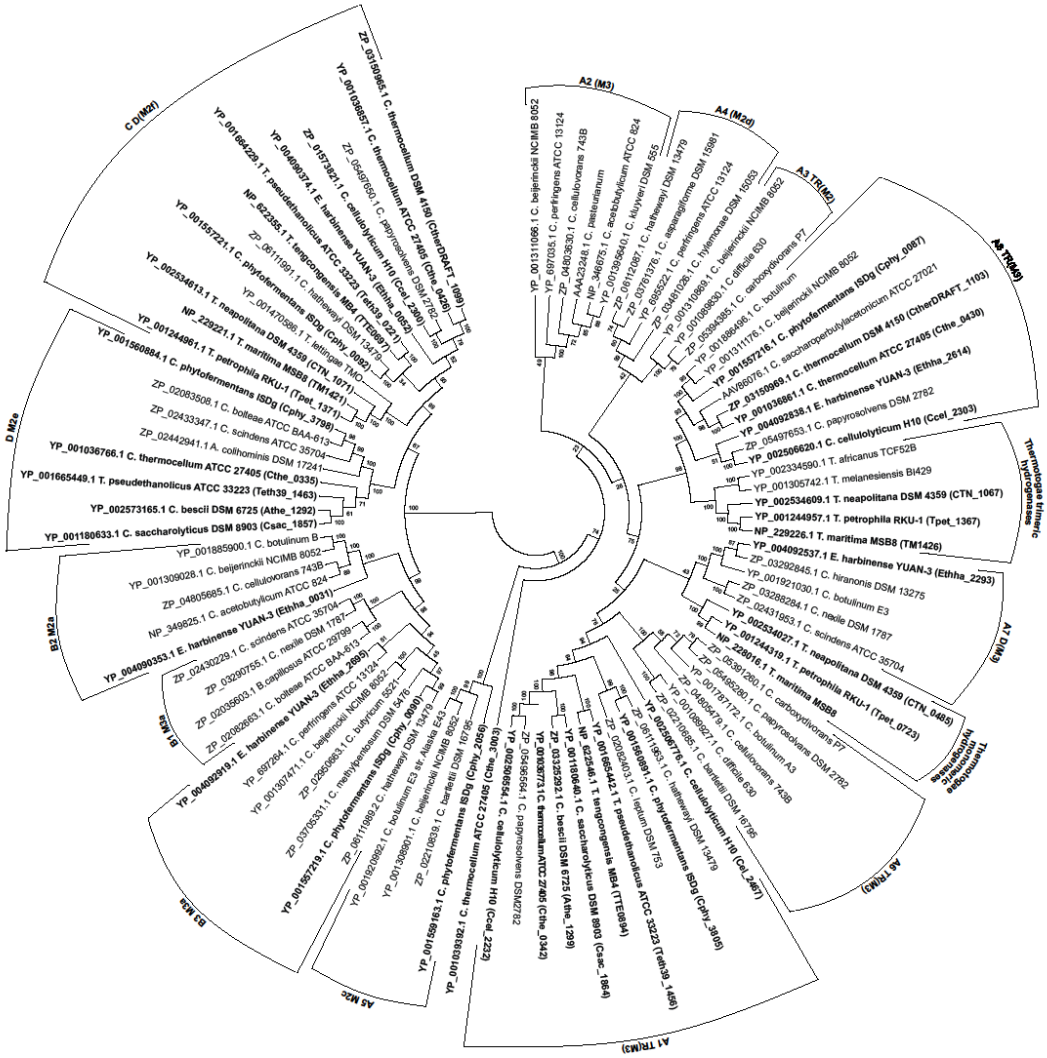
Supplementary Table 5: List of selected peptides, transitions states, corresponding peak areas, and retention times used to determine changes in protein expression in response to exogenous pyruvate using MRM. Given the excessive size of the table it could not be reproduced in this thesis, but can be accessed as an excel file through the following link:

https://www.dropbox.com/s/iscehyg5u88mn3y/MRM%20Supplementary%20File%202%20-%20HPP%20data.xlsx

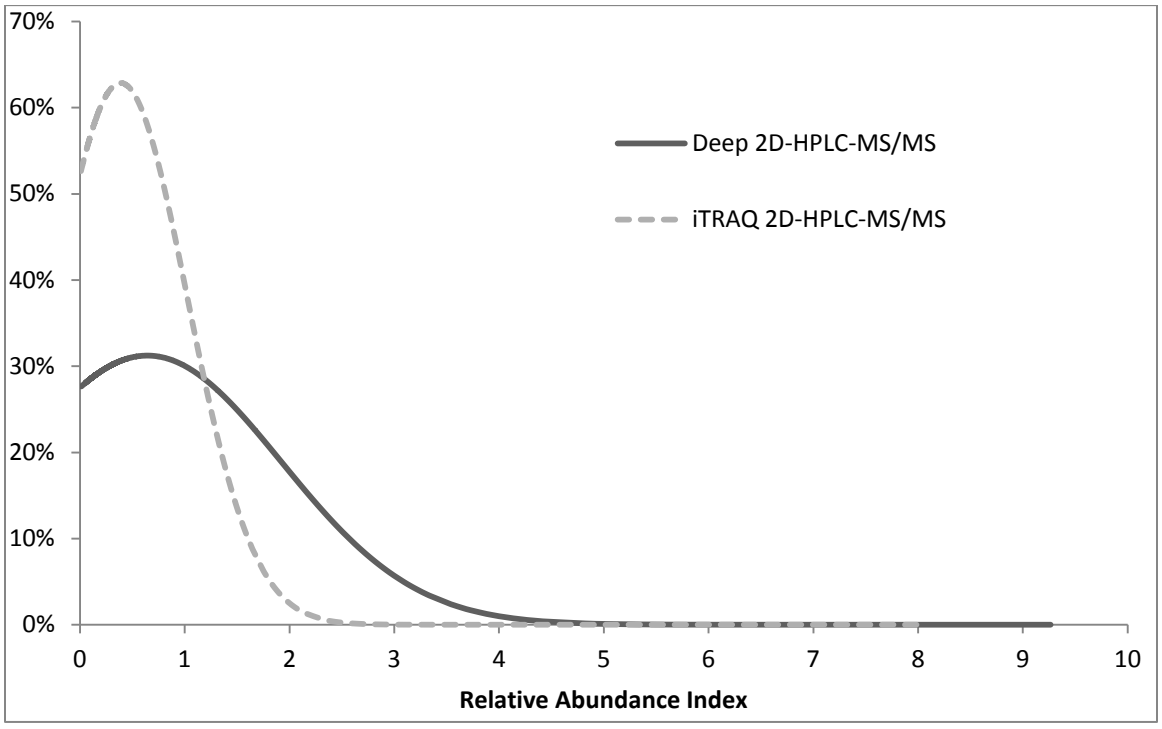
Supplementary Figure 1: Phylogenetic clustering of [NiFe] hydrogenases large (catalytic) subunits. Catalytic (large) subunits of [NiFe] H_2ases were identified based upon the modular signatures as described by Calusinska *et al*. [16]. Species considered in this manuscript are highlighted and corresponding H_2ase gene loci are provided. This image can also be accessed as a PDF file through the following link: **http://www.biomedcentral.com/content/supplementary/1471-2180-12-295-s2.pdf**



Supplementary Figure 2: Phylogenetic clustering of [FeFe] hydrogenases large (catalytic) subunits. Catalytic (large) subunits of [FeFe] H_2ases were identified based upon the modular signatures as described by Calusinska *et al*. (Calusinska et al. 2010). Species considered in this manuscript are highlighted and corresponding H_2ase gene loci are provided. This image can also be accessed as a PDF file through the following link: **http://www.biomedcentral.com/content/supplementary/1471-2180-12-295-s3.pdf**

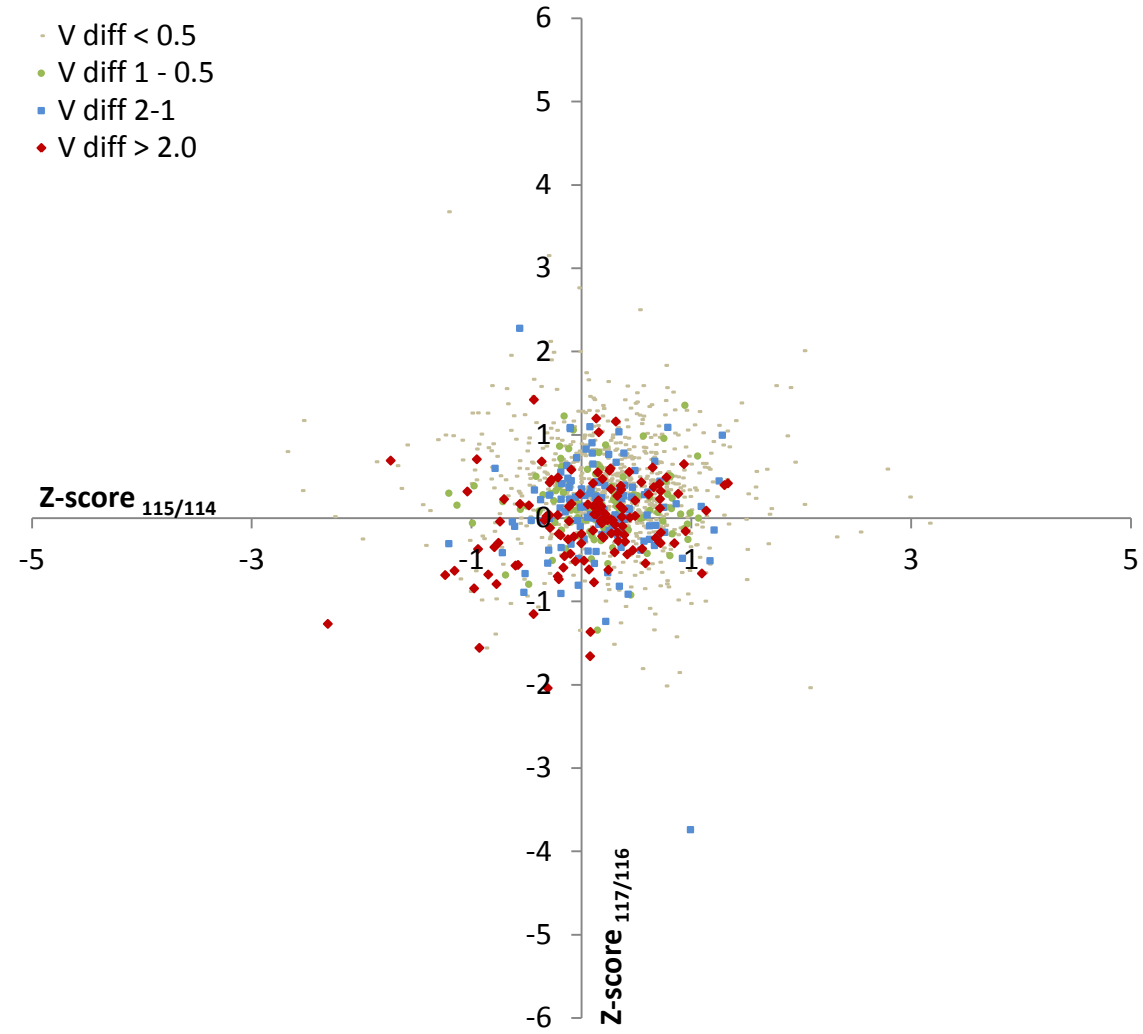


Supplementary Figure 3: Relative abundance index (RAI) distribution using shotgun (deep) and 4-plex 2D-HPLC-MS/MS. RAI distribution values follow a similar trend using both acquisition methods, however RAI per given protein was lower using 4-plex 2D-HPLC-MS/MS.

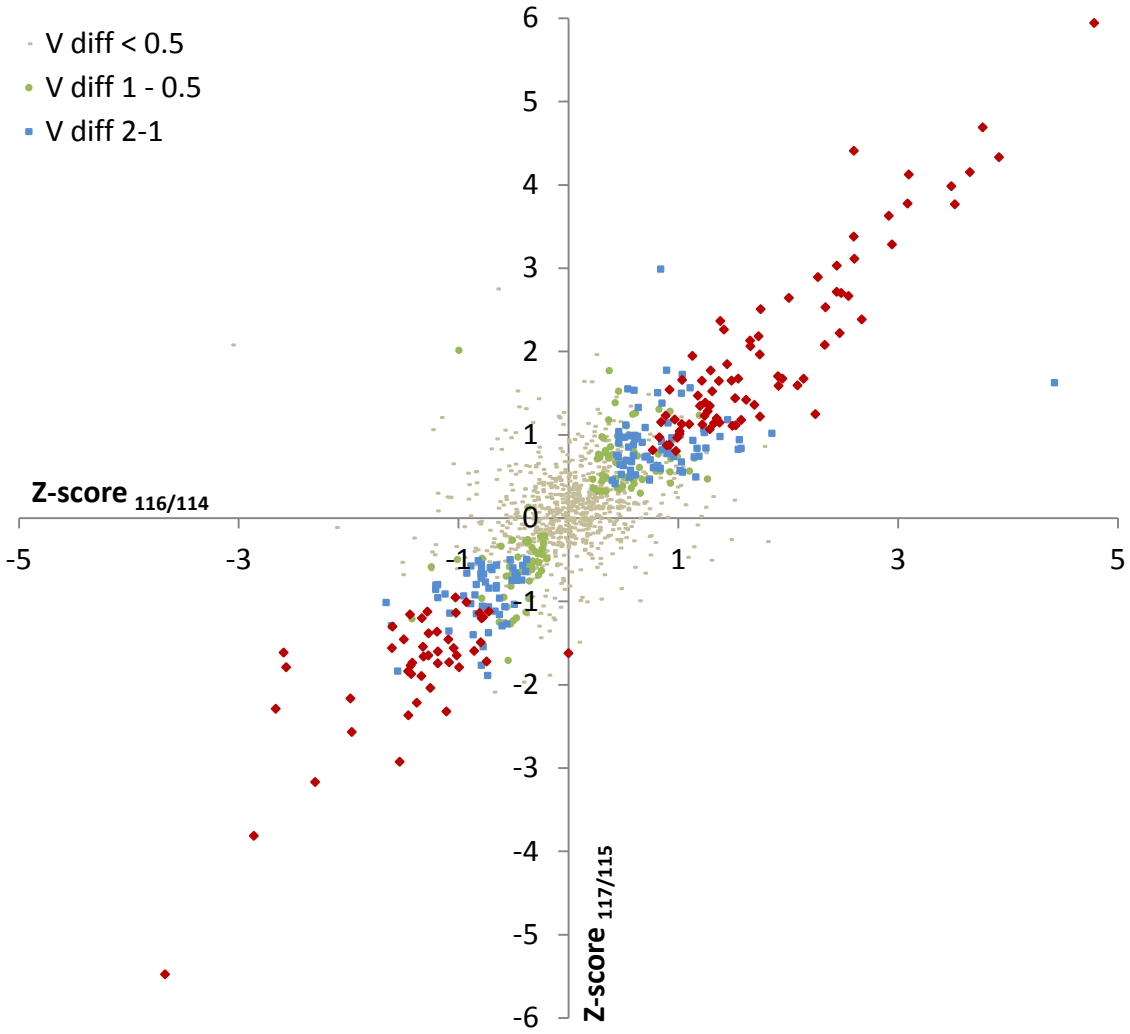


Supplementary Figure 4: Correlation of protein iTRAQ ratios for biological replicates. Protein z-score value ratios (A) among stationary and exponential phase biological replicates (reporter ion ratio 114/115 *vs* 116/117) and (B) between stationary *vs* exponential phase cell lysates (reporter ion ratio 116/114 *vs* 117/115) illustrating correlation between biological replicates. Positive correlation is represented by points in quadrants 1 and 3.

A)


V diff < 0.5
V diff 1 - 0.5
V diff 2-1
V diff > 2.0
Z-score 115/114
Z-score 117/116
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-5
-3
-1
1
3
5

B)


V diff < 0.5
V diff 1 - 0.5
V diff 2-1
Z-score 116/114
Z-score 117/115
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-5
-3
-1
1
3
5