2012

Biochemical Conversions of Lignocellulosic Biomass for Sustainable Fuel-Ethanol Production in the Upper Midwest

Michael James Brodeur-Campbell
Michigan Technological University

BIOCHEMICAL CONVERSIONS OF LIGNOCELLULOSIC BIOMASS FOR
SUSTAINABLE FUEL-ETHANOL PRODUCTION IN THE UPPER MIDWEST

By

Michael J. Brodeur-Campbell

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

(Chemical Engineering)

MICHIGAN TECHNOLOGICAL UNIVERSITY

2012

This dissertation, "Biochemical Conversions of Lignocellulosic Biomass for Sustainable Fuel-Ethanol Production in the Upper Midwest," is hereby approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING.

Department of Chemical Engineering

Signatures:

Dissertation Advisor   _____

David Shonnard

Department Chair   _____

S. Komar Kawatra

Date   _____

To my family,

     The one I was born to

     And the one we made.

Table of Contents

List of Figures

List of Tables

Preface

I would like to gratefully acknowledge the help of Jordan Klinger for his work on Chapter 3. My contribution to Chapter 3 was to perform the literature review, plan and help conduct the experiments, analyze the data, and write up the results.

Acknowledgements

Abstract

Biofuels are an increasingly important component of worldwide energy supply. This research aims to understand the pathways and impacts of biofuels production, and to improve these processes to make them more efficient. In Chapter 2, a life cycle assessment (LCA) is presented for cellulosic ethanol production from five potential feedstocks of regional importance to the upper Midwest – hybrid poplar, hybrid willow, switchgrass, diverse prairie grasses, and logging residues – according to the requirements of Renewable Fuel Standard (RFS). Direct land use change emissions are included for the conversion of abandoned agricultural land to feedstock production, and computer models of the conversion process are used in order to determine the effect of varying biomass composition on overall life cycle impacts. All scenarios analyzed here result in greater than 60% reduction in greenhouse gas emissions relative to petroleum gasoline. Land use change effects were found to contribute significantly to the overall emissions for the first 20 years after plantation establishment. Chapter 3 is an investigation of the effects of biomass mixtures on overall sugar recovery from the combined processes of dilute acid pretreatment and enzymatic hydrolysis. Biomass mixtures studied were aspen, a hardwood species well suited to biochemical processing; balsam, a high-lignin softwood species, and switchgrass, an herbaceous energy crop with high ash content. A matrix of three different dilute acid pretreatment severities and three different enzyme loading levels was used to characterize interactions between pretreatment and enzymatic hydrolysis. Maximum glucose yield for any species was 70% of

theoretical for switchgrass, and maximum xylose yield was 99.7% of theoretical

for aspen. Supplemental β-glucosidase increased glucose yield from enzymatic

hydrolysis by an average of 15%, and total sugar recoveries for mixtures could

be predicted to within 4% by linear interpolation of the pure species results.

Chapter 4 is an evaluation of the potential for producing *Trichoderma reesei*

cellulose hydrolases in the *Kluyveromyces lactis* yeast expression system. The

exoglucanases Cel6A and Cel7A, and the endoglucanase Cel7B were inserted

separately into the *K. lactis* and the enzymes were analyzed for activity on

various substrates. Recombinant Cel7B was found to be active on

carboxymethyl cellulose and Avicel powdered cellulose substrates. Recombinant

Cel6A was also found to be active on Avicel. Recombinant Cel7A was produced,

but no enzymatic activity was detected on any substrate. Chapter 5 presents a

new method for enzyme improvement studies using enzyme co-expression and

yeast growth rate measurements as a potential high-throughput expression and

screening system in *K. lactis* yeast. Two different *K. lactis* strains were evaluated

for their usefulness in growth screening studies, one wild-type strain and one

strain which has had the main galactose metabolic pathway disabled. Sequential

transformation and co-expression of the exoglucanase Cel6A and

endoglucanase Cel7B was performed, and improved hydrolysis rates on Avicel

were detectable in the cell culture supernatant. Future work should focus on

hydrolysis of natural substrates, developing the growth screening method, and

utilizing the *K. lactis* expression system for directed evolution of enzymes.

Chapter 1: Cellulosic Ethanol: A More Sustainable Transportation Fuel for the United States

Observed increases in average global surface temperatures in recent decades can be largely attributed to increases in greenhouse gas concentrations in the atmosphere (IPCC 2007a). The dominant cause of increased greenhouse gas concentrations is the burning of fossil fuels for energy production. One of the major uses of fossil fuels is in petroleum-based transportation fuels, which accounted for 27% of U.S. greenhouse gas emissions in 2010 (EPA 2010a). Near-term replacements to fossil-derived transportation fuels will need to be technologically feasible, economically viable, and largely compatible with existing infrastructure in order to be successful in the marketplace.

Ethanol is one such biofuel, and cellulosic ethanol has the potential to provide a more sustainable transportation fuel if it can be successfully commercialized. In the recent Billion-Ton Update (U.S. DOE 2011) it was estimated that agricultural residues and other woody feedstocks could economically provide nearly one billion dry metric tons of lignocellulosic biomass in the United States for less than $60/ton by the year 2022. If converted to ethanol this could displace up to 30% of gasoline used in transportation. One of the key barriers to converting this biomass to ethanol is the recalcitrance of lignocellulosic biomass to hydrolysis (U.S. DOE 2006).

This dissertation presents experimental and modeling research targeted towards

commercialization of sustainable cellulosic ethanol production in the upper

Midwest region of the United States. Chapter 2 presents a life cycle assessment

(LCA) of cellulosic ethanol production from five potential feedstocks of regional

importance to the upper Midwest. This LCA is performed according to the

requirements of the Renewable Fuel Standard, the recent legislation passed by

the U.S. Congress setting forth targets and emissions standards for biofuels

being blended into the U.S. transportation fuel supply. Chapter 3 investigates the

effect on overall sugar recovery from processing mixed feedstocks for a

commercial biorefinery. While previous studies have largely focused on process

optimization for a single feedstock, a viable commercial biorefinery will likely

have to simultaneously process many different feedstocks from multiple sources.

Chapter 4 studies the potential for producing cellulose hydrolyzing enzymes from

the wood-rot fungus *Trichoderma reesei* in the yeast *Kluyveromyces lactis*.

Cellulase enzyme production is among the top contributors to both operating

costs and environmental burdens in a cellulosic ethanol facility. Enzyme

production is also one of the best opportunities for process improvements, as

enzymatic hydrolysis is still a developing technology and still has a great deal of

potential for improvement. Enzyme expression in yeast offers opportunities for

improving process economics and efficiency, however not all enzymes can be

properly produces in a yeast expression host. Chapter 5 proposes a new

approach to cellulase enzyme studies, using the *K. lactis* yeast expression

system. Because of the large number and low specific activity of individual

enzymes, enzyme improvement studies are hampered by the lack of a truly high-throughput enzyme expression and screening system. *K. lactis* possesses the ability to express multiple different enzymes simultaneously. Combined with its ability to metabolize many of the different kinds of sugars present in lignocellulosic biomass, enzyme co-expression in *K. lactis* holds promise for enzyme improvement studies which relate directly toward the goal of consolidated bioprocessing (CBP), creating a single micro-organism which can hydrolyze lignocellulosic biomass and ferment it directly to ethanol or other high-value chemicals (Lynd et al. 2002; Zhang et al. 2006a).

Chapter 2: Life Cycle Assessment of Cellulosic Ethanol Production for Five Upper Midwestern Regional Scenarios

2.1 INTRODUCTION

In 2007 the Intergovernmental Panel on Climate Change (IPCC) concluded that most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations (IPCC 2007a). Transportation accounted for 27% of greenhouse gas emissions in the United States in 2010 (EPA 2010a), with over 90% of transportation fuels being derived from fossil petroleum (IPCC 2007b). Biofuels have the potential to reduce greenhouse gas emissions from transportation, since renewable biomass feedstocks remove carbon dioxide from the atmosphere during biomass growth. However, the dominant biofuel used for transportation in the United States currently is ethanol derived from corn starch. Corn starch ethanol production is still very fossil-fuel intensive and only results in approximately 20% reduction of greenhouse gas emissions relative to gasoline (EPA 2010b), although other recent studies show some corn ethanol pathways able to achieve greenhouse gas emissions reductions of up to 67% relative to gasoline (Wang et al. 2007; Liska et al. 2009). Additionally there are other negative social and environmental consequences of converting this potential food source into fuel, such as increased food prices (Mitchell 2008), increased nutrient

leaching and eutrophication (Sheehan 2009), and reduced biodiversity (Fargione et al. 2009).

Several studies indicate that cellulosic ethanol, a biofuel derived mainly from inedible plant materials, has the potential to provide a renewable transportation fuel with a much better environmental profile than corn ethanol, if it can be successfully commercialized (Tilman et al. 2006; Kline et al. 2009; EPA 2010b). This study evaluates the environmental effects of cellulosic ethanol produced from several potentially important feedstocks for the upper Midwest; hybrid poplar, hybrid willow, switchgrass, diverse prairie grass, and logging residues according to the requirements of the Renewable Fuel Standard. The Renewable Fuel Standard (RFS) of the Energy Independence and Security Act of 2007 (EISA) was enacted to encourage more sustainable biofuel production in the United States (110th Congress 2007), including cellulosic ethanol and other advanced biofuels. It also establishes several categories for renewable fuels, defines what may be considered renewable biomass, and sets production targets for biofuels through 2022. It mandates that lifecycle greenhouse gas emissions of biofuels be determined according to specific guidelines, including evaluating direct and indirect land use change emissions from biomass production. Lifecycle greenhouse gas emissions must meet a 60% reduction threshold in order to be considered a cellulosic biofuel under the RFS. In addition to greenhouse gas emissions it also stipulates other environmental and social impacts of biofuels production be considered, including: air quality, water quality

and water consumption, wildlife habitat, ecosystem health, food prices, fuel
prices, energy security, rural economic development, and others.

To be eligible for consideration under the RFS, biofuels must be made from
allowable feedstocks. These include planted crops and crop residues from
agriculture land cleared prior to December 19, 2007 and actively managed or
fallow on that date for switchgrass and prairie grass, and planted trees and tree
residues from tree plantations established prior to December 19, 2007 and
actively managed or fallow on that date for tree plantations (110th Congress
2007).

In response to the RFS mandate, the Environmental Protection Agency (EPA)
has performed comprehensive lifecycle assessments for conventional corn
ethanol, soybean biodiesel, sugarcane ethanol, and switchgrass ethanol, and
has determined that they will meet the requirements set forth in the RFS (EPA
2010b), known as the RFS2. The report concludes that switchgrass will likely
achieve a GHG reduction of 110% relative to petroleum gasoline. In addition,
this report has also identified several biofuels pathways which are similar enough
to those already analyzed that they are likely to also meet the requirements of
the RFS, although they have not evaluated any of the other feedstocks
investigated here.

As cellulosic ethanol has been a topic of much discussion and research recently, there are also many other valuable techno-economic and life cycle assessments available. The National Renewable Energy Laboratory (NREL) has developed an Aspen Plus process model for cellulosic ethanol production, described in the design report by Humbird et al. (Humbird et al. 2011). It establishes a most likely technological and economic scenario for processing of corn stover, but does not evaluate any potential environmental consequences.

Many other life cycle assessments have also been performed on poplar and switchgrass, and generally report greenhouse gas reductions in the range of 57-115%. Hsu et al. (Hsu et al. 2010) built on an earlier techno-economic model developed at NREL, and described in the design report of Aden et al. (Aden et al. 2002). They evaluated environmental burdens for biochemical conversion of switchgrass, corn stover, and wheat straw to ethanol, as well as an advanced corn grain process, and the thermochemical conversion of forest residues. They found GHG reductions for switchgrass bioethanol production of approximately 57% compared to petroleum gasoline using product displacement allocation (see Methods section for more on allocation methods), although they neglect land use change effects.

Adler et al. (Adler et al. 2007) found roughly comparable results for both poplar and switchgrass at approximately 115% GHG emissions reduction for each, using the DAYCENT model to determine land use emissions and co-product

displacement for electricity generated from lignin. Cherubini and Jungmeier (Cherubini and Jungmeier 2010) determined a GHG reduction of 80% for a specific switchgrass bioethanol production pathway, using energy allocation and accounting for land use change in the same manner as this study.   At the low end, Spartari et al. (Spartari et al. 2005) found GHG reductions of approximately 57% for switchgrass bioethanol production, using co-product displacement allocation and assuming no net carbon flux to or from the land. No complete studies have been found in the literature from plantation establishment through ethanol production and end use on willow or diverse prairie scenarios.

The RFS also mandates that studies be made of the impact of biofuels production "on the environment, including on air quality, climate change, conversion of wetlands, ecosystems, wildlife habitat, water quality, and water supply" (110th Congress 2007). For the study reported in this dissertation we have chosen to evaluate eutrophication potential as an indicator of water quality as this is the category most significantly impacted by biomass production, and particulate matter emissions less than 2.5 μm (PM2.5) for air quality as this category has significant implication for human health.

While eutrophication potential and particulate matter emissions will certainly have an effect on the environment, they do not give a complete understanding of the overall effect of biofuels production. One of the most useful indicators of ecosystem health would be a direct measure of biodiversity in the affected area.

Biodiversity is an important consideration for ecosystem health, since diverse

ecosystems are more robust (Tilman 1996), and can produce more biomass with

fewer inputs than monocultures (Tilman et al. 2006).

A quantitative understanding of the importance of species diversity goes back at

least to the work of Edward Simpson (Simpson 1949). This Simpson index,

when applied to ecological studies, is a measure of the probability that two

individuals taken at random will be members of the same species. Species

richness, another indicator of biodiversity, is simply a count of the number of

different species in a given habitat, and has a number of different ways of

estimation (Grubb 1997; Chazdon et al. 1998). By definition, plant biodiversity is

lower in a dedicated monoculture, such as a poplar or switchgrass plantation,

than it would be in a polyculture, such as the diverse prairie scenario. This has

follow-on effects for the wildlife, with some species being unable to find suitable

habitat in monoculture biomass plantations (Webster et al. 2012). Work is in

progress to establish biodiversity indicators suitable for use in the framework of

LCA, such as the work at Oak Ridge National Laboratory (McBride et al. 2011),

and here at Michigan Technological University (Webster et al. 2010). This will be

of enormous benefit to future life cycle assessments of biofuels production once

the data are well-established. Currently, collecting original data of biodiversity in

feedstock plantations and modeling indirect land use change effects are outside

the scope of this study.

The requirement of determining direct and indirect land use change effects has been one of the more controversial provisions of the RFS. In the debate on direct land use change effects, Fargione et al. (Fargione et al. 2008) point out that clearing land for biofuels plantations can emit a great deal of carbon dioxide initially, which creates a "carbon debt" that takes a long time to pay back. Kline et al. (Kline et al. 2009) make the counter-argument that when biomass feedstocks are grown in appropriate locations under sustainable conditions, a number of environmental benefits accrue, including reduced greenhouse gas emissions. This illustrates the importance of explicitly stating all assumptions about land use change in life cycle assessments (LCAs) of biofuels. In this LCA, plantations were assumed to have been established on abandoned agricultural land.   As abandoned agricultural land generally has lower carbon content both above and below ground than native ecosystems. It should therefore create very little "carbon debt," and may even result in a net carbon sequestration. This aspect will be evaluated in greater depth in the Discussion section

The RFS is one of several recent policies which explicitly adopt a life cycle perspective. Many of these policies deal with waste management or product labeling policy (Curran 1997). The requirement to include indirect land use change has caused a debate LCA community on the very nature of what makes a good LCA. One effect of policies such as the RFS on the field of Life Cycle Assessment has been to broaden the scope of LCA, which has developed from a simple energy analysis to a comprehensive environmental burden analyses

(Guinée et al. 2010). The continued influence of policies such as these on the field of LCA is to cause them to include not only direct environmental effects, but also to consider indirect effects. This has caused LCA practitioners to develop what has become known as a Consequential LCA (Brander et al. 2009). These Consequential LCAs include a broader scope and consider other environmental and social concerns, and the field is moving towards a full Life Cycle Sustainability Assessment (Zamagni 2012). However, these developments are still in progress and are not accepted by all practitioners as valid. Notwithstanding these new developments, determining indirect land use change effects is beyond the scope of the limited life cycle assessment presented in this study.

As shown in this literature review of cellulosic ethanol LCAs, there are a variety of study assumptions that make it difficult to compare ethanol produced from different lignocellulosic feedstocks. In this study we address this deficiency by conducting LCAs on five different feedstocks using consistent modeling tools and assumptions within the context of the upper Midwest region of the United States. This study seeks to contribute to the understanding of cellulosic ethanol production, using the language of the RFS as a guide to ensure that an environmentally and politically relevant analysis is performed. The five feedstocks evaluated here are hybrid poplar, hybrid willow, switchgrass, diverse prairie grasses, and logging residues. A discussion of the acceptability of producing these feedstocks on abandoned agriculture land with respect to the

definition of renewable biomass in the RFS is also included. This study has four main objectives:

1. Enhance previous lifecycle assessments for hybrid poplar and switchgrass by including direct land use change, water, and air quality indicators
2. Expand the number of feedstocks evaluated to include hybrid willow, diverse prairie grasses, and logging residues
3. Develop a better understanding of the effect varying biomass composition and moisture content will have on inputs and outputs for the conversion process, and on life cycle environmental impact of cellulosic ethanol
4. Determine the effect of direct land use change from conversion of abandoned agricultural land to energy crop production on lifecycle greenhouse gas emissions for cellulosic ethanol

The primary focus of this study has been on greenhouse gas emissions including land use change impacts, and a determination of whether the fuel production pathways modeled here will meet the requirements of the RFS for cellulosic biofuels. Preliminary water and air quality impacts are also included to provide a more environmentally robust analysis. A qualitative discussion of biodiversity and indirect land use change in LCAs is also included. . Although the RFS does not consider forest resources such as hybrid poplar established on abandoned agricultural land after 2007 as allowable feedstocks for biofuels, this study will investigate the greenhouse gas implications of this land use change decision.

2.2 METHODS

The LCA software tool SimaPro 7.2 was used to calculate the lifecycle environmental impacts for each scenario, using the EcoInvent database as the primary source of inventory data. Feedstock production impacts were determined by accessing the relevant literature for each biomass type. Direct land use change CO_2 emissions were calculated according to the method of the Intergovernmental Panel on Climate Change, using a Tier 1 approach (IPCC 2006). Nitrous oxide (N_2O) emissions were calculated on a nitrogen basis using and emission factor of 1.325% of the amount of nitrogen applied in the fertilizer. Particulate matter less than 2.5 μm was used as the primary air quality indicator as this is the fraction which has the greatest implication for human health. Eutrophication potential was chosen as the primary water quality indicator as this is the category which is most significantly affected by biomass production. Aspen Plus 7.2 process simulation software was used to model the NREL conversion process for the differing biomass compositions in order to determine the effect on conversion inputs and outputs. Detailed descriptions are provided in the appropriate sections below.

2.2.1 Goal, Scope, Functional Unit, and Allocation Methods
The goal of this study was to evaluate the environmental impacts of ethanol production from five lignocellulosic feedstocks relevant to the upper Midwest

region. This study aims to satisfy the requirements of the Renewable Fuel Standard (RFS) as given in the Energy Independence and Security Act of 2005 (EISA) by evaluating impacts for global climate change including direct land use effects.

The scope of this study includes emissions from all key inputs for feedstock production, feedstock transportation, conversion to ethanol, product distribution, and end use in an automobile, as seen in Figure 2.1. The majority of lifecycle inputs are associated with the biomass production and feedstock conversion steps, and are described in more detail as indicated in the following sections. The functional unit is one megajoule (lower heating value, LHV) of ethanol in the delivered fuel, to facilitate comparisons with other life cycle assessments of fossil and renewable transportation fuels.

| Biomass Production | Transportation | Feedstock Conversion | Distribution | End Use |

Figure 2.1: Life Cycle Diagram of Cellulosic Ethanol Production

Because the conversion process generates two energy product streams, both ethanol and electricity, results for two LCA methodologies are presented. In the displacement method all impacts are allocated to the ethanol, but excess electricity is assumed to displace grid electricity at the regional mix, providing a

credit to the ethanol for the co-product electricity. The regional electricity mix of primary energy sources was determined from the EPA Emissions & Generation Resource Integrated Database (eGRID) 2005 data for the states of Michigan, Ohio, Indiana, Illinois, and Wisconsin. In the energy allocation method, environmental burdens are allocated to both the ethanol and electricity based on the ratio of the primary energy content for the two streams, ethanol and electricity.

2.2.2 Feedstock Production

A feedstock production rate of 83,333 kg/hr (dry weight) was set for all feedstocks to provide sufficient material for a commercial scale biorefinery, based on the recent report of Humbird et al. (Humbird et al. 2011). Biomass production in oven dry tonnes per hectare per year varies from 0.56 to 13.5 (see Table 2.1 below), averaged over the productive lifespan of the plantation.

Table 2.1
Summary of feedstock production and transportation

	Poplar	Willow	Switchgrass	Prairie	Harvesting Residue
Productivity (tonne/(ha.yr))	13.5	12	9	5	0.56
Harvest Cycle (yr)	7	4	1	1	N/A
Area cultivated (ha)	52,000	58,000	78,000	140,000	1,250,000
Feed transport distance (km)	41	43	50	67	199

2.2.2.1 Production Inputs

Inputs for plantation energy crops (poplar, willow, and switchgrass) include fuel use for nursery, site preparation, planting, maintenance, and harvesting;

fertilizers; and herbicides and pesticides. Hybrid poplar and hybrid willow were assumed to be produced in short rotation woody feedstock plantations according to Gasol et al. 2009 (Gasol et al. 2009) and Heller et al. 2003 (Heller et al. 2003), respectively. Switchgrass production inputs were taken from Sampson 2007 (Sampson 2007), and modified for divers prairie grasses based on Tilman 2006 (Tilman et al. 2006). Logging residue production requirements were determined from surveys in previous work performed at Michigan Technological University (Reis 2009). Eutrophication impacts on watersheds through fertilizer runoff leaching was assumed to be 30% of the amount applied, according to IPCC guidelines (IPCC 2000). A detailed listing of all feedstock production inputs is included in Appendix A.

2.2.2.2 Direct Land Use Change Carbon Calculations

Direct land use change CO_2 emissions from feedstock production were calculated according to the method described by the Intergovernmental Panel on Climate Change (IPCC) in their "2006 Guidelines for National Greenhouse Gas Inventories," (IPCC 2006). Biomass plantations were assumed to be established on abandoned agricultural land, which was modeled as "moderately degraded grassland" according to the IPCC categorization. Conversion was modeled according to the IPCC categories of "moderately degraded grassland" to "forest" for poplar and willow scenarios, and to "improved grasslands" for switchgrass and mixed species prairie scenarios. No land use change was assumed for the logging residues because the land remains in the same land use category.

32

A Tier 1 approach was used, which assumes that the carbon flux for the land was in equilibrium prior to conversion, with no net carbon emissions or sequestrations. After conversion the land moves towards a new carbon equilibrium state over a period of 20 years. After 20 years in the new land use category the land is assumed to again be in equilibrium with no further net carbon accumulation or release.

Carbon is modeled in four categories on the land; above ground biomass (AGB), below ground biomass (BGB), soil organic carbon (SOC), and dead organic matter (DOM). Calculations for each category are presented in detail in Appendix A, with summary results presented here in Table 2.2.

Table 2.2
Land Use Change Carbon Sequestration (tonne C/(ha.yr))

	Poplar	Willow	Switchgrass	Prairie	Harvesting Residue
AGB	-0.056	-0.056	-0.056	-0.056	0
BGB	0.294	0.037	0.455	0.238	0
SOC	0.178	0.178	1.142	0.697	0
DOM	0	0	0	0	0
Total	0.416	0.159	1.541	0.879	0

The CO_2 emission (-0,056 tonne C/(ha.yr)) for above ground biomass represents the standing biomass in the field that was cleared but not utilized in order to establish the plantation. This biomass is mineralized relatively rapidly and becomes a carbon emission to the atmosphere. Below ground biomass for the four plantation scenarios was calculated in the same manner as for the above

ground biomass, using the ratio of above to below ground biomass indicated by the IPCC document. Soil organic carbon was calculated relative to the reference soil organic carbon to a depth of 30 cm for cold temperate moist ecosystems with sandy soils and modified for the different management scenarios, as detailed in Appendix A. Dead organic matter net carbon fluxes are assumed to be zero for Tier 1 calculation methods.

Direct land use change emissions were allocated entirely to the main product, ethanol, in the displacement method. In the energy allocation method, dLUC emissions were allocated to ethanol according to the energy allocation factor.

2.2.3 Feedstock Transportation

The feedstock transportation stage represents the transportation required to convey raw biomass to a central processing facility and includes fuel use and infrastructure maintenance from the EcoInvent database. Feedstock transportation distances were calculated according to the equation given in Wright et al. 2008 (Wright et al. 2008).

Eq. 2.1
$$\bar{r}_{circle} = \frac{2}{3}\tau\sqrt{\frac{F}{\pi * Y * f}}$$

Where:

$\bar{r}_{circle}$ is the average distance to the processing plant from a circular surrounding area

τ is a tortuosity factor, 1.5 is used as characteristic of a rectangular grid

F is the feed requirement, 83,333 kg/hr

Y is the biomass yield (see Table 2.1 for values)

f is the fraction of land devoted to feedstock production, 0.1 is used for this study

This formula gives an average one-way transportation distance varying from 41 km for poplar to 199 km for logging residues (see Table 2.1 above). Transportation was assumed to be by road using diesel powered truck. Transportation burdens include the water weight in the moist feedstock, and were calculated for a two-way trip under the assumption that trucks had to return empty to the field to collect more biomass.

2.2.4 Feedstock Conversion

The feedstock conversion stage models the processing of biomass feedstocks into ethanol, co-products, and wastes. This stage includes all conversion process inputs and waste disposal. Feedstock composition for the conversion process was varied for each scenario based on typical values for each species reported in the literature. Little information is available about protein concentration of woody feedstocks harvested for energy production. Protein content of woody feedstocks was assumed to be lower than for the herbaceous feedstocks and set to a value of 1%. Feedstock composition values were normalized to achieve 100% mass closure to allow for use in the Aspen Plus

process modeling environment. These data are shown in Table 2.3 below. Raw

data on feedstock compositions are given in Appendix A.

Table 2.3
Normalized feedstock composition by percent on a dry weight basis.

	Poplar	Willow	Switchgrass	Prairie	Harvesting Residue
Extract	7.12	4.75	10.82	14.15	4.74
Glucan	43.64	39.38	32.54	30.90	46.20
Xylan	14.02	22.32	21.48	18.82	15.43
Galactan	0.94	1.52	0.99	1.17	0.97
Arabinan	0.87	0.60	2.66	3.11	1.00
Mannan	2.03	1.40	0.35	0.38	3.91
Lignin	26.07	22.28	19.63	18.36	23.39
Ash	1.59	1.98	5.71	7.71	0.29
Acetate	2.66	4.77	2.91	2.37	3.07
Protein	1.06	0.99	2.91	3.04	0.98
SUM	100	100	100	100	100

Feedstock compositions are given on a dry weigh basis. However, for the

process simulation, woody feedstocks (poplar, willow, and logging residues) were

assumed to contain 50% moisture by weight, and herbaceous feedstocks

(switchgrass and prairie) were assumed to contain 15% moisture by weight.

The feedstock conversion process was modeled using Aspen Plus V7.2 to

determine how input and output quantities would vary for each feedstock

depending on the biomass composition and moisture content. The Aspen Plus

model file was obtained from the National Renewable Energy Laboratory (NREL)

at <http://www.nrel.gov/extranet/biorefinery/aspen_models/>. This model was

designed for the processing of corn stover to cellulosic ethanol, and contains the

most recent data available for the conversion process. A full description of the

process is contained in the 2011 design report of Humbird et al. (Humbird et al.

2011). Major conversion process inputs and products are summarized here in

Table 2.4.

Table 2.4
Conversion inputs, and ethanol and electricity production for different feedstocks.
All units are kg/hr unless otherwise noted.

	Poplar	Willow	Switchgrass	Prairie	Harvesting Residue
Biomass feedstock (dry)	83,333	83,333	83,333	83,333	83,333
Sulfuric acid	1,980	1,980	1,980	1,980	1,980
Ammonia	1,387	1,852	1,404	1,279	1,489
Corn steep liquor	1,358	1,354	1,307	1,297	1,383
Diammonium phosphate	142.7	144.9	142.9	142.4	143.8
Sorbitol	44.52	44.86	44.56	44.50	44.69
Glucose	3,542	3,197	2,631	2,508	3,751
Nutrients	83.87	75.69	62.28	59.38	88.81
Corn oil	16.63	15.01	12.35	11.78	17.61
Sulfur dioxide	21.05	19.00	15.63	14.90	22.29
WWT chemicals	5,333	7,431	5,559	5,025	5,764
Lime	763.3	742.2	834.6	876.1	760.9
Make-up water	111,166	91,816	158,580	158,982	100,621
Ethanol Production	22,404	23,596	21,363	19,877	23,972
Excess Electricity (kW)	14,539	8,877	15,456	16,241	10,326

Conversion process inputs and products were generated by the Aspen Plus

model based on feedstock composition, including; sugars, lignin, ash, acetate,

and moisture. Sulfuric acid use for pretreatment is one area of note. Although

experience indicates that high ash species such as switchgrass require greater

pretreatment severity for optimal sugar release, acid use in the NREL Aspen Plus

model is invariant; equal amounts of acid are used in each pretreatment

scenario. Pretreatment severity may also be increased by increasing

pretreatment time at the same acid concentration, which is the mechanism

utilized in this study. The high ammonia use for willow is due to the high acetate

content of that feedstock, which is converted to acetic acid during pretreatment

and then needs to be neutralized prior to fermentation. The consequences of

conversion inputs and products on the lifecycle performance for each scenario

are evaluated in the Discussion section.

2.2.5 Product Distribution and End Use

Product (ethanol) distribution was assumed to be 240 km by road using diesel

powered trucks. Transportation burdens were calculated for the one-way trip

under the assumption that tanker trucks would be fully utilized on any return trip

to carry other fuel products. End-use combustion in a motor vehicle is assumed

to be carbon neutral as all carbon in the ethanol fuel is biogenic in origin and

fixed from the atmosphere during biomass growth, creating a closed cycle.

2.2.6 Inventory and Impact Assessment

The inventory of emissions and energy demand for each of the inputs listed in

tables above were generated using the most technology, geographic, and time

relevant ecoprofiles in the EcoInvent™ database in SimaPro. Electricity

consumption was modeled on an upper Midwest regional average grid, as

described in Appendix A, and the EcoInvent database is considered to be a

modern and comprehensive life cycle inventory database. Global warming

potentials for all gaseous emissions was from the IPCC GWP 100a method in SimaPro; GWP = 1 for CO_2, 25 for CH_4, and 298 for N_2O. Nitrous oxide nitrogen emissions (N_2O-N) from fertilizer applications during feedstock production were assumed to be 1.325% of the amount of nitrogen applied. Particulate matter less than 2.5 μm and eutrophication potential were determined using the TRACI method in SimaPro (Bare 2012).

2.3 RESULTS

2.3.1 GHG Emissions for Co-product Displacement Allocation

Greenhouse gas emissions for the co-product displacement life cycle assessment method are presented graphically below in Figures 2.2 and 2.3, for scenarios including land use change impacts in the first 20 years after plantation establishment (Figure 2.2), and for subsequent time periods without land use change impacts (Figure 2.3).
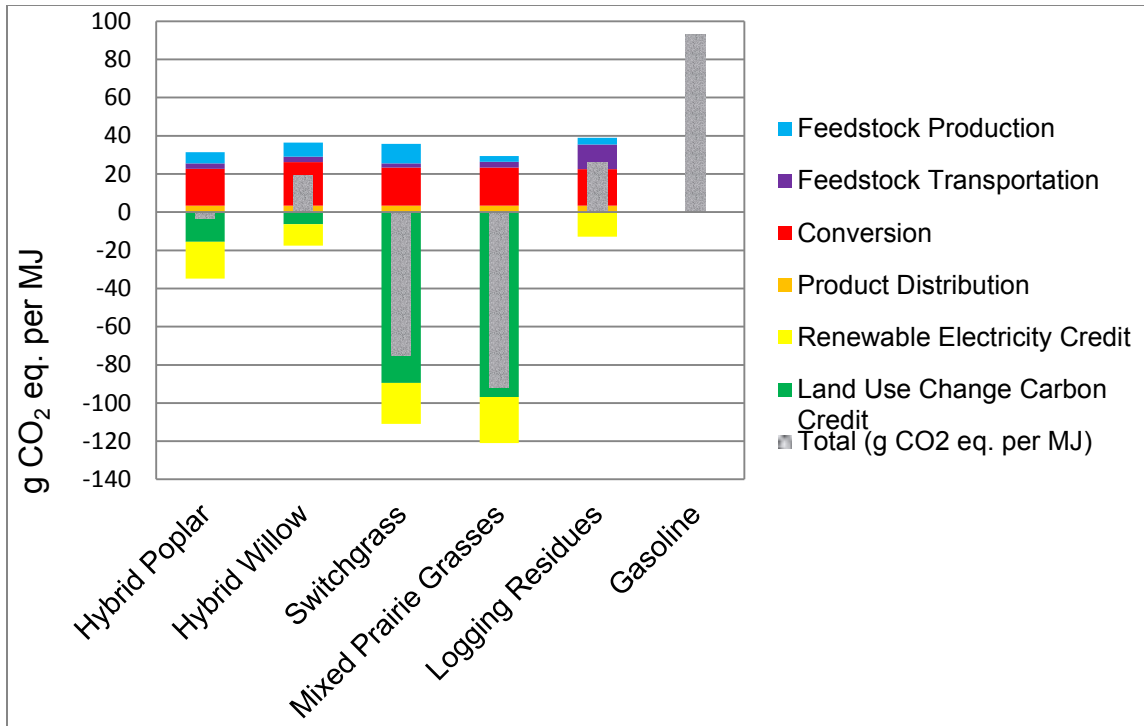
Figure 2.2: Greenhouse gas emissions for the first 20 years after plantation establishment with direct land use change calculations, displacement method.
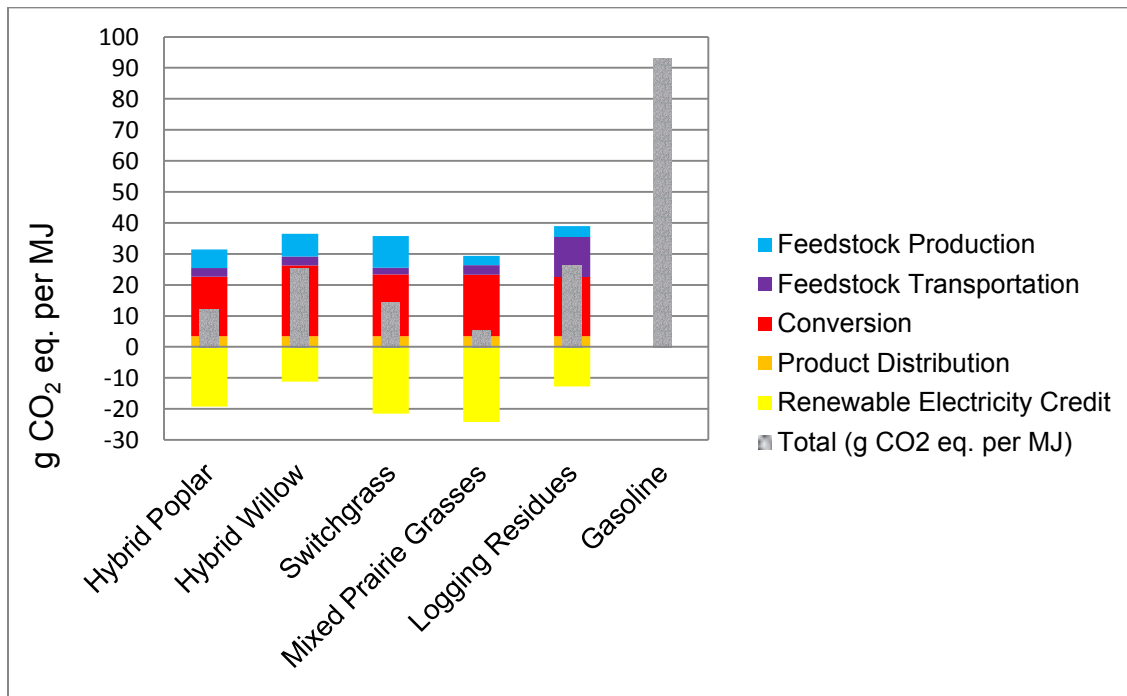
Figure 2.3: Greenhouse gas emissions for time periods greater than 20 years after plantation establishment without direct land use change, displacement method.

The feedstock conversion process is the largest contributor to greenhouse gas emissions for all cases, followed by feedstock production for most scenarios. However, feedstock transportation is a significant portion of the greenhouse gas emissions for logging residues due to the low productivity per acre, and subsequently large transportation distances for that feedstock. As can be seen, land use change impacts can have a considerable effect on the overall carbon balance, especially for herbaceous feedstocks which accumulate a significant amount of their total biomass in underground root systems and soil organic carbon.

Emissions and savings in CO_2 eq. emissions relative to petroleum gasoline are given numerically for all stages and sub-stages within feedstock production and conversion in Table 2.5 below.

Total greenhouse gas reductions vary from 71.8% for logging residues to 199% for prairie grasses where large carbon sequestration into the soil takes place over 20 years following plantation establishment. Percent reductions in greenhouse gas emissions are calculated relative to the emission determined for gasoline by the EPA in the RFS Final Rule (EPA 2010b) (98 kg CO_2 eq./MM BTU or 92.9 g CO_2 eq./MJ).

Table 2.5

GHG results, displacement, g CO_2 eq. per MJ and percent reduction relative to gasoline

Lifecycle Stage	Poplar	Willow	Switchgrass	Prairie	Logging Residues	Gasoline
Feedstock Production	5.97	7.37	10.2	3.02	3.54	
Fertilizers	*5.22*	*5.11*	*7.85*	*0.00*	*0.00*	
Herbicides and Pesticides	*0.09*	*0.05*	*0.20*	*0.00*	*0.00*	
Machinery and Fuels	*0.66*	*2.02*	*2.13*	*2.96*	*3.54*	
Other	*0.00*	*0.18*	*0.03*	*0.06*	*0.00*	
Feedstock Transportation	2.82	2.84	2.13	3.07	12.9	
Conversion	19.2	22.8	19.9	19.8	19.1	
Pretreatment	*4.70*	*6.04*	*5.12*	*5.03*	*4.69*	
Fermentation	*1.38*	*1.33*	*1.45*	*1.55*	*1.30*	
Enzyme Production	*3.35*	*2.87*	*2.61*	*2.67*	*3.32*	
Waste Water Treatment	*8.81*	*11.7*	*9.64*	*9.36*	*8.90*	
Boiler Chemicals	*0.95*	*0.88*	*1.09*	*1.23*	*0.89*	
Remaining Processes	*0.00*	*0.00*	*0.03*	*0.00*	*0.00*	
Product Distribution	3.45	3.45	3.45	3.45	3.45	
Renewable Electricity	-19.3	-11.2	-21.5	-24.2	-12.8	
Land Use Change	-15.5	-6.32	-89.5	-96.8	0.00	
Total w/ LUC Percent reduction	-3.36 104%	19.0 79.5%	-75.2 181%	-91.6 199%	26.2 71.8%	92.9 0%
Total w/o LUC Percent reduction	12.1 86.9%	25.3 72.8%	14.2 84.7%	5.18 94.4%	26.2 71.8%	92.9 0%

Within the feedstock production stage fertilizers (especially nitrogen fertilizers) are the single largest contributor to greenhouse gas emissions in scenarios where fertilizers are used. Machinery and their fuel use is the next largest category, or the largest category in scenarios which use no fertilizer.

Within the conversion process, waste water treatment is the largest single contributor to greenhouse gas emissions for all scenarios, followed by

pretreatment, then enzyme production. Waste water treatment chemicals are

modeled as sodium hydroxide, the primary purpose of which is to control pH in

the waste water treatment pond where denitrification takes place.

2.3.2 GHG Emissions for Energy Allocation

For the energy allocation scenario, both environmental burdens and the land use

change carbon credit are allocated based on the primary energy of the two

product streams. Electricity production in the conversion facility is 30.4% efficient

at converting primary biomass energy into electricity. This leads to the energy

allocation factors for ethanol given in Table 2.6 below.

Table 2.6
Ethanol energy allocation factors

	Poplar	Willow	Switchgrass	Prairie	Harvesting Residues
Ethanol primary energy	603,788	635,912	575,733	535,685	646,045
Electricity primary energy	172,172	105,122	183,032	192,328	122,282
Ethanol allocation factor	77.8%	85.8%	75.9%	73.6%	84.1%

Results for the energy allocation assessment are presented graphically in

Figures 2.4 and 2.5 below, both for scenarios including land use change impacts

in the first 20 years after plantation establishment (Figure 2.4), and for

subsequent time periods without land use change impacts (Figure 2.5).
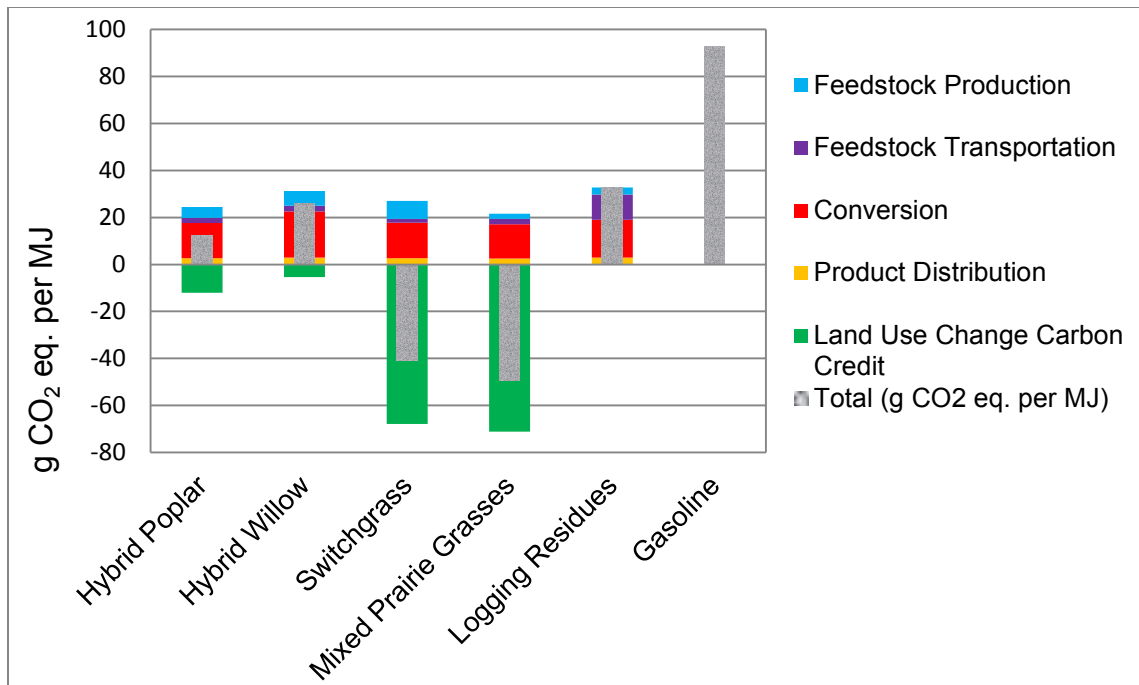
Figure 2.4: Greenhouse gas emissions with direct land use change, energy allocation method.
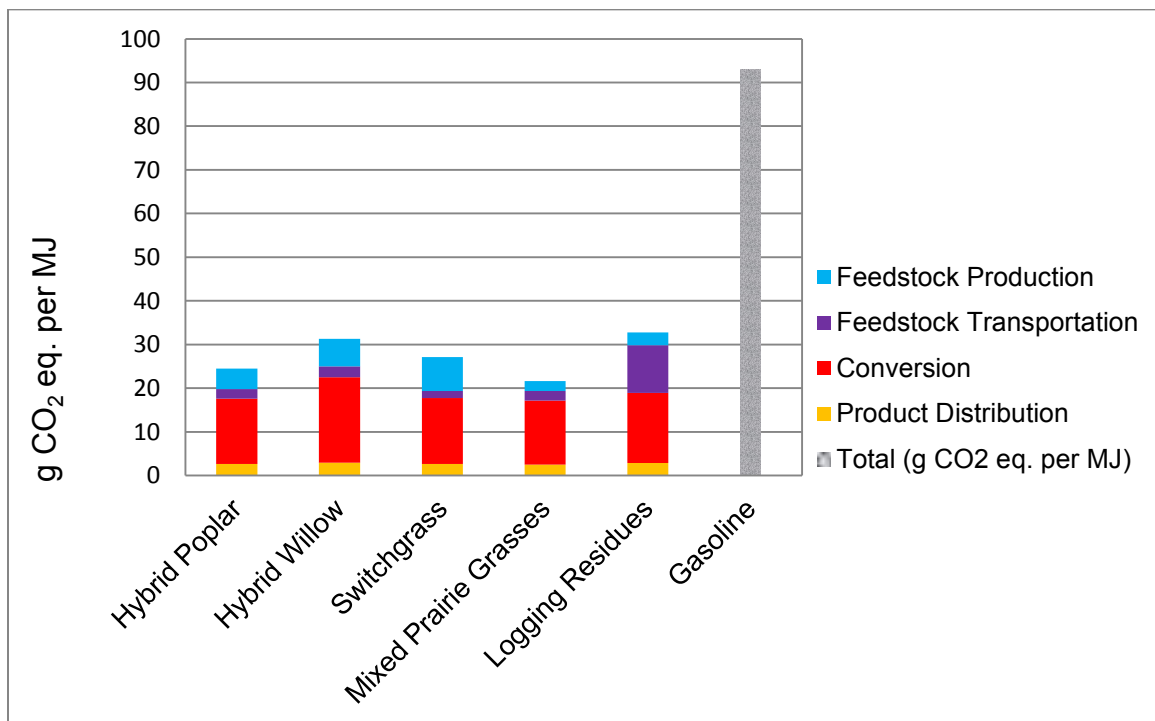
Figure 2.5: Greenhouse gas emissions without direct land use change, energy allocation method.

The energy allocation method shows slightly less favorable results for greenhouse gas emissions compared to the displacement method, although the patterns among the feedstocks and life cycle stages are still the same. Conversion is the largest contributor to emissions, followed by feedstock production, with biomass transportation being significant only for logging residues. Herbaceous feedstocks show strong advantages for the first twenty years when significant biomass is accumulating below ground due to land use change effects, after which both woody and herbaceous feedstocks show roughly comparable results. Results are shown numerically in Table 2.7.

Table 2.7
GHG results, allocation g CO_2 eq. per MJ and percent reduction relative to gasoline

	Poplar	Willow	Switchgrass	Prairie	Logging residues	Gasoline
Feedstock Production	4.65	6.32	7.74	2.22	2.98	
Fertilizers	4.06	4.39	5.95	0.00	0.00	
Herbicides and Pesticides	0.07	0.05	0.15	0.00	0.00	
Machinery and Fuels	0.52	1.73	1.61	2.18	2.98	
Other	0.00	0.15	0.02	0.04	0.00	
Feedstock Transportation	2.19	2.44	1.62	2.26	10.8	
Conversion	14.9	19.6	15.1	14.6	16.1	
Pretreatment	3.66	5.18	3.88	3.70	3.94	
Fermentation	1.07	1.14	1.10	1.14	1.09	
Enzyme Production	2.61	2.46	1.98	1.96	2.79	
Waste Water Treatment	6.86	10.0	7.31	6.89	7.48	
Boiler Chemicals	0.74	0.75	0.83	0.91	0.74	
Remaining Processes	0.00	0.00	0.02	0.00	0.00	
Product Distribution	2.68	2.96	2.62	2.54	2.90	
Land Use Change	-12.1	-5.42	-67.9	-71.2	0.00	
Total w/ LUC	12.4	25.9	-40.8	-49.6	32.8	92.9
Percent reduction	86.7%	72.1%	144%	153%	64.7%	0%
Total w/o LUC	24.5	31.3	27.1	21.6	32.8	92.9
Percent reduction	73.7%	66.3%	70.8%	76.7%	64.7%	0%

Total greenhouse gas reductions for the energy allocation method vary from

64.7% for logging residues to 144% for prairie grasses where large carbon

sequestration into the soil takes place in the 20 years following plantation

establishment.

2.3.3 Air and Water Emissions

Preliminary air and water quality data for the product displacement scenario are

presented below in figures 2.6 and 2.7.
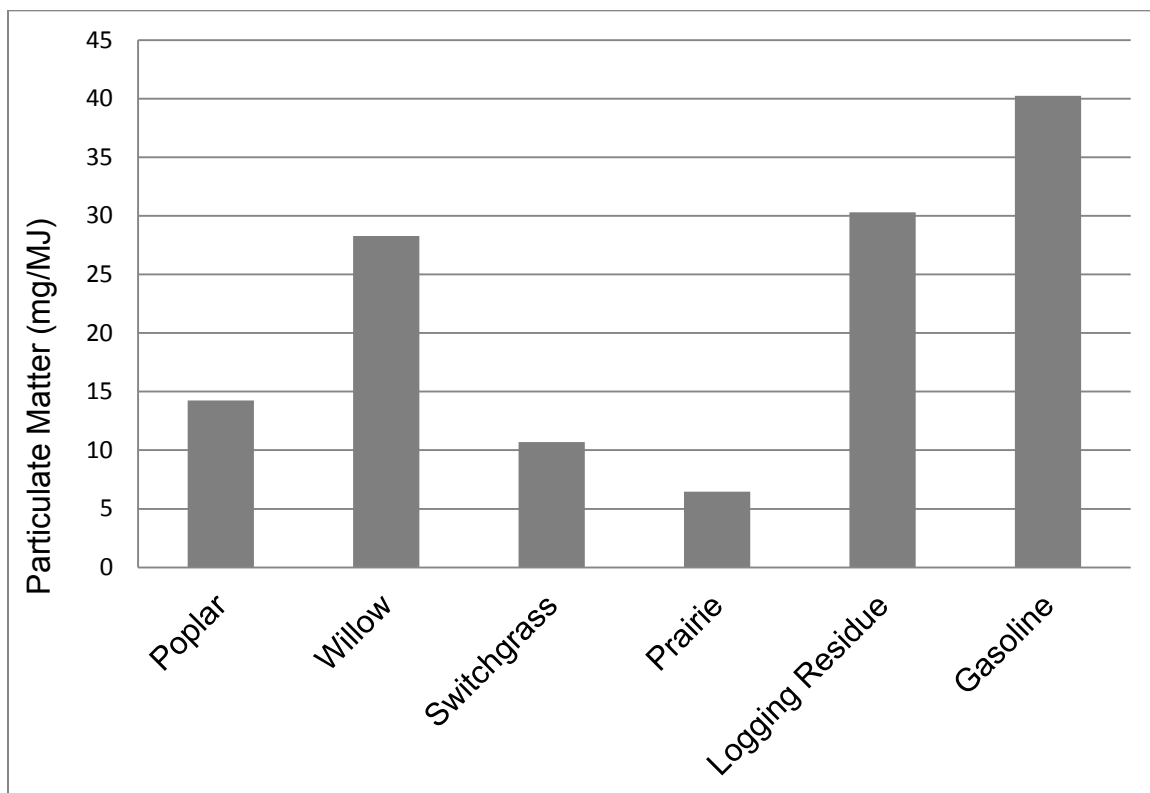
Figure 2.6: Respiratory effects of ethanol production

Particulate matter emissions are generally higher for the woody feedstocks than

for the herbaceous feedstocks, and all feedstocks show lower emissions than

gasoline. Particulate matter emissions for logging residues are dominated by feedstock transportation, while for willow they are associated with other fuel use for feedstock production and handling. These data are only for emissions from fuel and other infrastructure burdens and do not include particulates directly produced by feedstock production steps, e.g. harvesting or chipping.
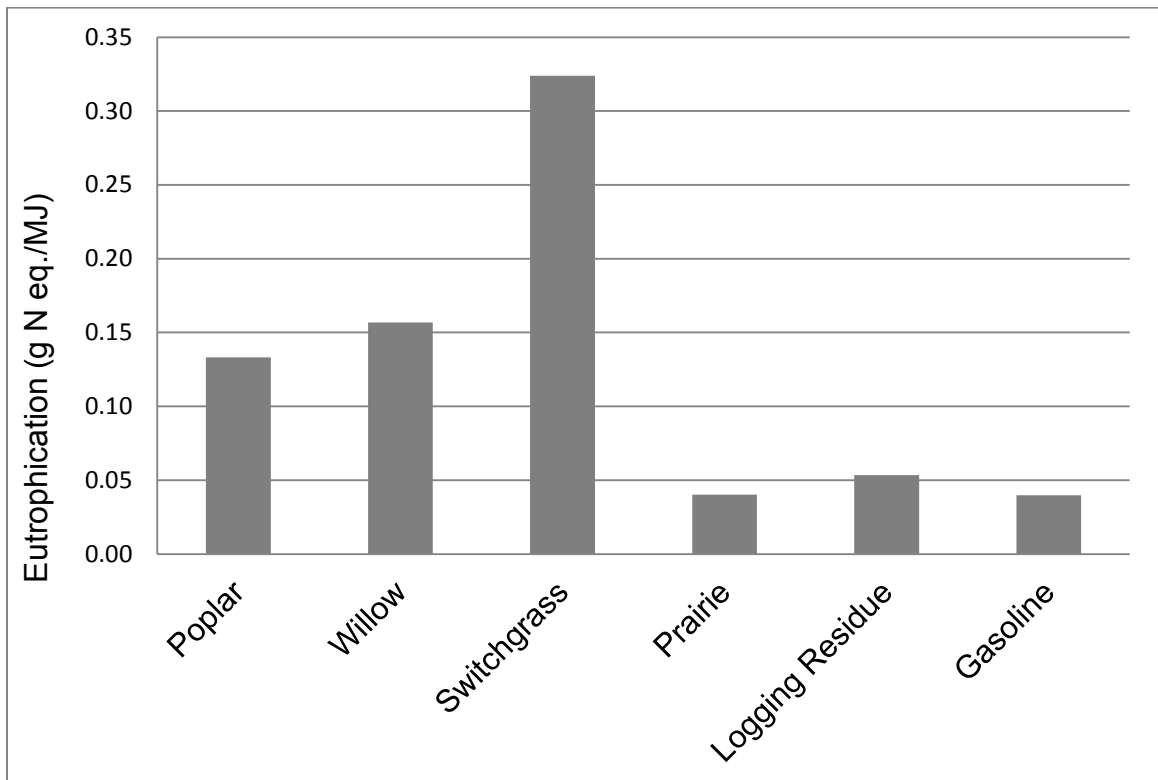
Figure 2.7: Eutrophication potential of ethanol production

Fertilizer use drives eutrophication potential, for the switchgrass case which has the highest fertilizer inputs the eutrophication potential is highest. Poplar and willow had moderate fertilizer inputs and have moderate eutrophication potential. Prairie and logging residues have no fertilizer inputs, and they have the lowest eutrophication potential, similar to gasoline, although they still have some

eutrophication potential due to burdens of fuel use and infrastructure maintenance.

2.4 DISCUSSION

Total greenhouse gas emissions for all scenarios evaluated here are below the 60% reduction threshold required to meet the requirements of a cellulosic biofuel as defined by the RFS, using either displacement or energy allocation methods. However, savings were greater for displacement compared to energy allocation. Direct land use change effects during the initial 20 years immediately following plantation establishment result in a significant sequestration of carbon on the land for all feedstocks other than logging residues, although even after the land reaches equilibrium the greenhouse gas reductions are still greater than 60% for all pathways evaluated here. Because abandoned agricultural land generally has lower carbon content than the energy crops studied here, the "carbon debt" for utilizing these lands is in fact negative, creating a "carbon credit". Although tree plantations established on abandoned agricultural land would not be considered renewable biomass under the language of the RFS, they clearly show environmental benefits in this analysis. This may be an area where science can inform policy makers to improve the policy language to better accomplish policy goals.

This life cycle assessment has benefited from the availability of the NREL process model simulations in Aspen Plus to investigate the consequences of feedstock composition and moisture content on conversion inputs and overall life cycle impacts. As expected, ethanol production was seen to be most strongly affected by the carbohydrate content of the feedstock, with higher sugar content resulting in more ethanol produced. More surprisingly, electricity export was found to be dependent on both lignin and moisture content of the feedstock (see Table 2.4). As lignin is the primary component of the boiler fuel, high lignin content results in higher electricity production. However, higher moisture content of the feedstock results in more electricity being consumed during the conversion process, primarily in the feedstock handling (mechanical conveyance) and pretreatment areas, resulting in less excess electricity available for export. Moisture content of the feedstock was also found to effect make-up water requirements in the conversion process in the expected manner, with wetter feedstocks requiring less fresh water make-up in the process. This finely detailed understanding of the consequences of changes in process conditions would not be possible in the absence of such a detailed process model.

The results developed in this study fall in the middle range of the studies surveyed, approximately 70-80% for both switchgrass and poplar scenarios excluding land use change effects, although switchgrass shows a greater potential for land use change carbon sequestrations immediately following plantation establishment according to the IPCC Tier 1 method. Most of the other

studies reviewed found that one-half to two-thirds of life cycle GHG emissions result from the feedstock production stage. Our results show generally less than half of the life cycle GHG emissions coming from the feedstock production stage, with one-half to two-thirds of GHG emissions coming from the conversion stage. One major difference is that the previous model developed by NREL in 2002 assumes lime is used to condition the hydrolysate from pretreatment, which adjusts the pH and also precipitates sulfate and gypsum. However, removal of gypsum also results in sugar loss. The 2011 NREL cellulosic ethanol simulation model assumes ammonia is used to condition the hydrolysate. This results in greater sugar recovery, but also results in greater base consumption during waste water treatment to control pH during denitrification.

Air and water quality indicators presented here are preliminary and do not include all possible sources of emissions. They do however indicate relatively promising pathways among the scenarios evaluated here. Eutrophication potential is dominated by fertilizer use, highest for switchgrass and lowest for prairie. Particulate emissions seem to be higher for woody feedstocks than for herbaceous ones, although this does not hold true for all cases. Particulate matter emissions generally reflect the amount of fuel used in preparing the feedstocks.

2.5 FUTURE WORK

Modeling studies such as these are valuable in indicating the most likely effects of cellulosic ethanol production under various scenarios, and pointing towards promising directions for research. However, models are only as accurate as the assumptions upon which they are built. More experimental data are needed from field studies to validate the results presented here, especially for commercial-scale production of cellulosic ethanol production. The IPCC Tier 1 methods utilized in this study are a beginning, but represent the least detailed level of landscape modeling for carbon fluxes. More detailed models, such as the Canadian Forest Service-Carbon Budget Model are currently being employed in our laboratory to provide a better understanding of carbon fluxes on the landscape level. Additionally, experimental validation from field studies would also be valuable to verify these model results. Continued monitoring of land use changes around the globe is also important. While the actual contribution of policies such as the RFS to indirect land use change may be tenuous and difficult to establish with certainty, eventually the sum of data may be helpful in developing full consequential life cycle assessments of biofuels production.

Improved data on water and air quality effects of biomass production and harvesting are also important. Much regional variation exists for these impact categories, and detailed regional understanding is important for accurate lifecycle understanding. Biodiversity indicators which can be adapted for use in the

framework of life cycle analysis will also be valuable for evaluations of biomass

production, especially for energy production scenarios.

Finally, this study indicates areas for potential improvement in the Renewable

Fuel Standard itself. While woody biomass planted on abandoned agricultural

land is not considered renewable biomass according to the language of the RFS,

this study clearly indicates potential environmental benefits from such a biomass

production scenario. Since the intended goal of the RFS is to encourage more

sustainable transportation fuel production, this study indicates possibilities for

improving the language of the RFS to better accomplish those policy goals.

2.6 CONCLUSIONS

Cellulosic ethanol holds promise as a near-term renewable transportation fuel

with much better environmental performance than petroleum gasoline or corn

ethanol. Policies such as the RFS can encourage more sustainable energy

infrastructure by providing incentives for biofuel production, and establishing

lifecycle environmental impact performance criteria for production pathways.

However, the language of the RFS needs to be considered in light of the best

scientific evidence of the time. Results such as those presented here indicate

that much environmental benefit can be achieved by allowing woody feedstock

plantations to be established on abandoned agricultural land, which would be

prohibited as renewable biomass under the current regulatory framework. The

RFS represents one of a number of policies which embrace a life cycle perspective. These policies are driving the evolution of life cycle assessment from a simple attributional method towards a more inclusive consequential method.

All biofuel production scenarios examined here meet the greenhouse gas reduction requirements of the RFS for cellulosic biofuels. Feedstock production and conversion steps were the most significant contributors to overall greenhouse gas emissions. Based on the IPCC Tire 1 method, direct land use change effects show significant influence on the greenhouse gas emissions profile, especially during the period of time immediately following plantation establishment while the land is moving towards a new equilibrium state. Herbaceous feedstocks show significant promise for carbon sequestration when cultivated on abandoned agriculture land due to the significant amount of below ground biomass and soil organic carbon that accumulates. Eutrophication potential (water quality indicator) points to a need to control fertilizer runoff to improve environmental performance, however experimental data are needed to verify this result. Air quality impacts seem to be dominated by fuel used in feedstock production. Ultimately, more data will be needed on biodiversity effects from establishing and maintaining biomass plantations to determine overall environmental performance for these cellulosic ethanol production pathways.

Calculated estimates of land use change impacts, such as the IPCC Tier 1

method, are valuable for indicating the potential results of biofuel production,

however more experimental data are needed to corroborate these results.

Process simulations such as these implemented in Aspen Plus can also provide

valuable insight into how process variables such as feedstock composition and

biomass moisture content can effect overall lifecycle results. Although

simulations are a necessary step towards building a robust commercial

biorefinery, such models ultimately need validation using pilot- and commercial-

scale demonstration.

Chapter 3: Feedstock Mixture Effects on Sugar Monomer Recovery Following Dilute Acid Pretreatment and Enzymatic Hydrolysis[1]

3.1 INTRODUCTION

Much work has been done to understand and optimize the pretreatment and enzymatic hydrolysis conditions for various feedstocks including poplar and aspen (Grohmann et al. 1985; Chung et al. 2005; Jensen et al. 2010), spruce and pine (Martínez et al. 1997; Tengborg et al. 1998), switchgrass (Chung et al. 2005), and agricultural residues such as corn stover, wheat straw, and bagasse (Grohmann et al. 1985; Lloyd and Wyman 2005; Kabel et al. 2007; Guo et al. 2008). However, most studies have focused on optimizing hydrolysis conditions for a single species at a time, comparing them side by side for best performance. Alternately, the work done by the Consortium for Applied Fundamentals and Innovation (CAFI) group (Wyman et al. 2005) has evaluated a number of different pretreatment technologies, but applied them to a single well-defined feedstock – corn stover – to provide comparative performance data.

A commercially viable biorefinery will likely have to process mixed feedstocks from a variety of sources, for both economic and environmental reasons. These are likely to include materials such as softwoods, hardwoods, agricultural residues, municipal and industrial waste streams, and herbaceous or woody

---

[1] The material contained in this chapter was previously published in *Bioresource Technology* 116:320-326. See Appendix C for documentation of permission to republish this material.

energy crops. Although Lynd (Lynd and Grethlein 1987) did study enzymatic hydrolysis of a mixed hardwood substrate (90% maple, 10% birch), the effect of processing mixtures on overall sugar recovery is largely uninvestigated. Previous work in our laboratory has investigated the effect of processing northern Midwestern regional mixed forest feedstock mixtures on dilute acid hydrolysis (Jensen et al. 2008). It was shown that processing mixed forest feedstocks has no synergistic or antagonistic effects on dilute acid pretreatment, but little is known about the effects of processing mixed feedstocks on enzymatic hydrolysis or overall sugar recovery.

This study investigates the effects of forest biomass mixtures on overall sugar monomer recovery from the combined dilute acid hydrolysis and enzymatic hydrolysis. Aspen was chosen as a model hardwood species due to its high suitability for processing by this method. Balsam was chosen as a high-lignin, softwood adjunct to investigate the effect of elevated lignin on enzymatic hydrolysis. Both of these feedstocks are common woody species found in the Upper Midwest region of the U.S. Switchgrass was chosen as a model herbaceous energy crop with high ash content, because it can also be cultivated on lands in the Upper Midwest region including abandoned agricultural lands in forest regions or set aside lands in agricultural regions. Experiments were performed on each pure species, and on 50:50 blends of aspen:balsam and aspen:switchgrass.

We wished to test the hypothesis that overall sugar yields from mixed feedstocks could be predicted by a simple linear interpolation model based on the behavior of the pure species. It has been shown that dilute acid hydrolysis behavior is affected by ash content, neutralizing the free protons in solution and lowering the effective pretreatment severity. This effect can be overcome by increasing the acid concentration, or allowing a longer time for the pretreatment reaction. No effect of ash on enzymatic hydrolysis is expected as the enzyme system is buffered to an optimal pH for the reaction to take place. Lignin is known to reduce enzymatic hydrolysis efficiency of lignocellulosic biomass by several mechanisms; physical separation of enzymes from cellulosic substrate, binding of enzymes to lignin, precipitation of enzymes with soluble lignin, and inactivation of enzyme phenolic residues (McMillan 1994; Sewalt et al. 1997). Deviation from linear behavior for the softwood-hardwood mixture would indicate that significant enzyme inactivation is taking place from unique lignin-enzyme interactions, while strictly linear behavior would suggest that balsam lignin is sufficiently similar to aspen lignin and that the system is already saturated in enzyme-lignin interactions.

3.2 MATERIALS AND METHODS

A simplified process flow diagram for biochemical conversion of lignocellulosic biomass to sugars is found below in Figure 3.1. Ground, sieved material is treated with dilute sulfuric acid (0-2% wt) at elevated temperature (150-200°C) to

hydrolyze hemicellulose. Solid residue containing cellulose and lignin, after washing, are subjected to enzymatic hydrolysis. Dilute acid pretreatment and enzymatic hydrolysis conditions are described in detail in the Research Methods section.
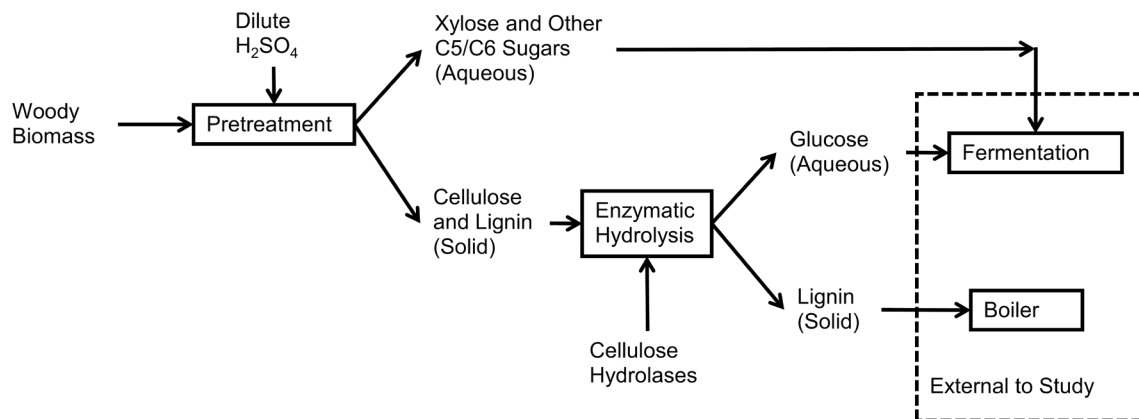
Figure 3.1: Simplified process flow diagram of biomass conversion

Previous work in our laboratory (Jensen et al. 2010) has determined the composition of the feedstocks and the optimal pretreatment time for these pure feedstocks under various conditions of acid concentration (0.25-0.75 wt% H_2SO_4) and temperature (150-175°C). The optimal pretreatment time is defined as the time at which the maximum amount of monomer and soluble oligomer sugars are recoverable in the liquid phase, with low concentrations of degradation products such as furfural and 5-hydroxymethyl furfural (HMF). Expected sugar recoveries for pure species at 0.5 wt% H_2SO_4 and 160°C are given in Table 3.1 below and served as benchmarks for this study. Feedstock composition (expressed as monomer equivalents) is given in Table 3.2. Dilute acid hydrolysis times were 24 minutes for aspen, 20 minutes for balsam, and 72 minutes for switchgrass.

Enzymatic hydrolysis results were measured after three days at 50C, pH 4.8, 250

RPM reciprocal shaking.

Table 3.1

Expected xylose and glucose recoveries (both monomer and oligomer expressed
as monomer equivalents) for pure species at 0.5 wt% H_2SO_4 and 160°C
(Adapted from (Jensen et al. 2010).

	Dilute Acid Pretreatment		Enzymatic Hydrolysis		Total Sugars	
	Xylose	*Glucose*	*Xylose*	*Glucose*	*Xylose*	*Glucose*
Aspen	87%	3%	5%	65%	92%	68%
Balsam	69%	6%	0%	3%	69%	10%
Switchgrass	70%	12%	10%	65%	81%	77%

Table 3.2

Feedstock composition.

	Feedstock Composition					
Species	*Glucan*	*Xylan*	*Galactan*	*Arabinan*	*Mannan*	*Lignin*
Aspen[1]	52.43%	14.60%	3.52%	2.41%	5.32%	26.69%
Balsam[1]	47.09%	6.23%	5.45%	5.41%	11.49%	36.04%
Switchgrass[2]	31.47%	19.73%	1.31%	2.59%	0.18%	21.36%

[1] As determined by Yat (Yat et al. 2008).
[2] Provided by the National Renewable Energy Laboratory, Golden CO (NREL).

A primary concern in this current study is to determine the effect of processing

mixed lignocellulosic feedstocks on sugar recovery from enzymatic hydrolysis.

However, because dilute acid pretreatment severity can have an effect on

subsequent processing steps, a matrix of different pretreatment conditions and

enzyme loadings must be studied simultaneously. For this study a single

pretreatment acid concentration and temperature (0.5 wt% H_2SO_4, 160°C) was

chosen, and the pretreatment severity was varied by adjusting the pretreatment

time. Three pretreatment times were studied at ½x, 1x, and 2x the optimal

pretreatment time for the mixture or pure species. At each pretreatment time three samples were taken for analysis and subsequent enzymatic hydrolysis. Three different enzyme loadings were examined for each pretreatment time condition, giving a matrix of nine unique pretreatment and enzymatic hydrolysis conditions for each pure species or mixed species feedstocks. Enzyme loading levels were 0.1, 0.25, or 0.5 ml enzyme formulation/g-carbohydrate in the feedstock, from lowest to highest manufacturer recommendation (see section 2.3). Six sets of experiments were performed, two for pure aspen, one for balsam, one for switchgrass, one for a 50:50 mix of aspen and balsam, and one for a 50:50 mix of aspen and switchgrass.

3.2.1 Biomass Preparation

Prior to hydrolysis each biomass species was debarked (for aspen and balsam), dried, chipped, hammer milled, and screened to a uniform sample size. The sample preparation was performed in accordance with NREL's Laboratory Analytical Procedure NREL/TP-510-42620, with modifications to adjust for appropriate sieve sizes. Drying was performed at 100°C for 24 hours. A knife mill (Thomas Wiley® NR. 3557524 359264) was used to grind the dried biomass to an appropriate range of sizes. A screening apparatus (W.S. TYLER ROTAP model RX-29, serial 9774) was used in the particle size differentiation. The particle size used in these experiments was between 20 and 28 mesh (0.853mm-0.599mm, Tyler Mesh).

3.2.2 Dilute Acid Pretreatment

One-half gram (0.5 g) of dry sample was weighed and placed into one of a set of

nine 316 stainless steel Swagelok cylindrical reactor tubes (total volume = 6.7

ml). With one end cap secured, 4.5 ml of dilute sulfuric acid (0.5 wt%) was

added to achieve 10 wt% biomass and 0.045 g H_2SO_4/g biomass in the reactor,

and the second end cap was tightened to 40 N-m of torque to assure a tight seal.

After allowing the dilute acid solution to diffuse into the wood particles

(approximately ½ hour, inverting every 10 minutes), the reactors tubes were

submerged into a silicon oil (Dow Corning 550 fluid) constant temperature bath

preheated to 175°C. The bath and reactors reach thermal equilibrium at 160°C

within five minutes. Tubes were removed from the oil bath in sets of three

(replicate samples) at the previously determined times, ½x, 1x, and 2x the

optimal pretreatment time. Reactor tubes removed from the oil bath were

immediately submerged in an ice-water bath to stop the reaction.

The reactors were opened and 3 ml of liquid was collected in 14 ml centrifuge

tubes. Liquid samples were centrifuged at 13,000 RPM for five minutes in a

Marathon 21K centrifuge (Fisher Scientific). Following centrifugation the liquid

supernatant was filtered through a 0.22 μm polycarbonate membrane filter

(VWR). One milliliter was neutralized with 10M NaOH to pH 5-6 and analyzed by

HPLC as described below to determine sugar monomer concentrations in the

liquid. Sugar concentrations were measured by refractive index and calculated

relative to standards of known concentration. In order to determine total oligomer

sugar concentrations one milliliter of the filtered liquid was further acidified to 4 wt% sulfuric acid and autoclaved for 60 minutes at 121°C to complete hydrolysis of the oligomers. The autoclave procedure hydrolyzes any dissolved oligomers into monomers with some sugar degradation. To account for sugar degradation a sugar recovery standard (SRS) of known sugar concentration is processed in the same manner as the samples to determine the fraction of sugars degraded to dehydration products and tars. Following the autoclave step, all standards and samples were neutralized to a pH between 5 and 6, centrifuged to remove any particulate matter, and the supernatant analyzed by HPLC in the same manner as the monomer samples

Solid material from each reactor was combined with solids precipitated from centrifugation, filtered on a glass fiber filter (VWR, grade 691), and washed with excess distilled water. Solids were then treated by enzymatic hydrolysis as described in section 3.2.3.

3.2.3 Enzymatic Hydrolysis

Solids, including the glass fiber filter, were placed into 50mL Erlenmeyer flasks with 47mL of distilled water and 2.5 ml of 1M sodium citrate buffer pH 4.5 to achieve enzymatic hydrolysis conditions of 1% solids loading (based on initial dry biomass), 50mM citrate, and pH 4.8. Antibiotics were added, 200µL of 10 mg/mL tetracycline (Sigma) and 150µL of 10 mg/mL cycloheximide (MP Biomedical), to prevent microbial growth. Negative controls containing glass fiber filters with no

biomass were prepared in an identical manner. Each flask was tightly sealed

with a rubber stopper and introduced to an incubated bench-top orbital shaker

(Lab-Line model 3527) and allowed to come to thermal equilibrium for 1 hour at

50°C and 250RPM shaking. After reaching thermal equilibrium, the Accelerase

1500 (Genencor) was added to the flasks at 0.1, 0.25, or 0.5 milliliters per gram

carbohydrate (lowest to highest manufacturer recommended loading). By

varying the levels of enzyme at each pretreatment condition, the combined

effects of enzyme loading and pretreatment severity was studied to optimize for

total sugar recovery.

A 3 ml liquid sample was removed from each flask at 0, 24, 48, and 72 hours.

Samples were filtered through a 0.22 μm polycarbonate membrane filter and

analyzed by HPLC for sugar concentration in the same manner as the

pretreatment samples.

3.2.4 HPLC Detection and Sugar Analysis

An Agilent 1200 Series HPLC with refractive index detector was used for sugar

determination.   A Bio-Rad de-ashing column was connected inline prior to the

analytical column to remove ions via ion exchange. A Bio-Rad Aminex HPC-87P

column was used to separate sugars for analysis. Sugar standards containing

known concentrations of glucose, xylose, galactose, arabinose, and mannose

were analyzed along with samples to determine retention times for different

sugars, and to create a standard curve to correlate sugar concentration with

detector response peak area. Arabinose and mannose were not effectively separated by this method, and their data were taken as a combined arabinose plus mannose peak.

Four different sugar concentrations were used to generate a standard curve for monomer sugar determination within samples. In order to accurately determine overall sugar recoveries, all dilutions with acid and base were factored into the calculation of sugar concentrations. Analysis of the SRS allows for determination of the concentration of monomer sugars that are degraded during the autoclave procedure. Taking into account the level of monomer degradation in addition to dilution factors, the total amount of monomer and oligomer sugars released during dilute acid pretreatment was calculated.

3.2.5 Sugar Recovery Model for Biomass Mixtures

A linear interpolation model (Eq. 3.1 below) was used to predict sugar recoveries for the combined steps of dilute acid pretreatment and enzymatic hydrolysis of mixed feedstock streams based on the known yields obtainable for pure species.

Eq. 3.1: $$P_{AB} = x_A P_A + x_B P_B$$

Where:

P_{AB} is the percent of sugar recovery for the mixture of species A and B

P_A is the percent of sugar recovery from pure species A

x_A is the mass fraction of species A

P_B is the percent of sugar recovery from pure species B

x_B is the mass fraction of species B

3.3 RESULTS AND DISCUSSION

Sugar concentrations were measured for each reaction condition, and sugar

recoveries for the optimal conditions are summarized here in Table 3.3. A

detailed discussion of each step follows in the appropriately marked sections.

Table 3.3

Experimental sugar yields (both monomer and oligomer expressed as monomer equivalents) for dilute acid hydrolysis (DAH), enzymatic hydrolysis (EH), and maximum total sugar yields (combined dilute acid pretreatment and enzymatic hydrolysis) at optimal pretreatment times and high enzyme loading for pure species and mixtures.

	DAH - monomer				DAH - oligomer				EH - monomer	
	Glucose	*Xylose*	*Galactose*	*Ar+Man*	*Glucose*	*Xylose*	*Galactose*	*Ar+Man*	*Glucose*	*Xylose*
Balsam	6.82%	67.91%	37.10%	53.04%	N/D	4.07%	1.37%	3.35%	2.28%	N/D
Aspen/Balsam	5.00%	77.64%	32.88%	44.16%	0.50%	4.04%	N/D	N/D	27.34%	8.16%
Aspen (1)	2.79%	75.05%	20.30%	19.19%	0.02%	7.49%	1.80%	0.99%	46.68%	4.29%
Aspen (2)	3.63%	87.36%	28.21%	24.76%	N/D	5.89%	N/D	N/D	61.47%	6.46%
Aspen/Switchgrass	5.61%	76.01%	45.12%	38.52%	N/D	2.71%	1.54%	2.67%	52.51%	3.81%
Switchgrass	8.92%	62.93%	128.65%	94.82%	2.73%	5.58%	22.62%	0.98%	59.05%	5.33%

66

3.3.1 Dilute Acid Pretreatment

For each experiment yield of xylose monomer and soluble oligomer were seen to vary with both time of reaction and feedstock type (Figure 3.2). Figure 3.2 shows the results for xylose, the main sugar released during pretreatment, for five of the six experiments performed; one each for aspen, balsam, and switchgrass, a 50:50 blend of aspen and switchgrass, and a 50:50 blend of aspen and balsam. Xylose yields are shown for all dilute acid pretreatment conditions. Values and error bars show the mean and standard deviation of three separate reactor trials. Graphs of all sugar monomer and oligomer yields from dilute acid pretreatment are provided in Section B.1 of Appendix B.
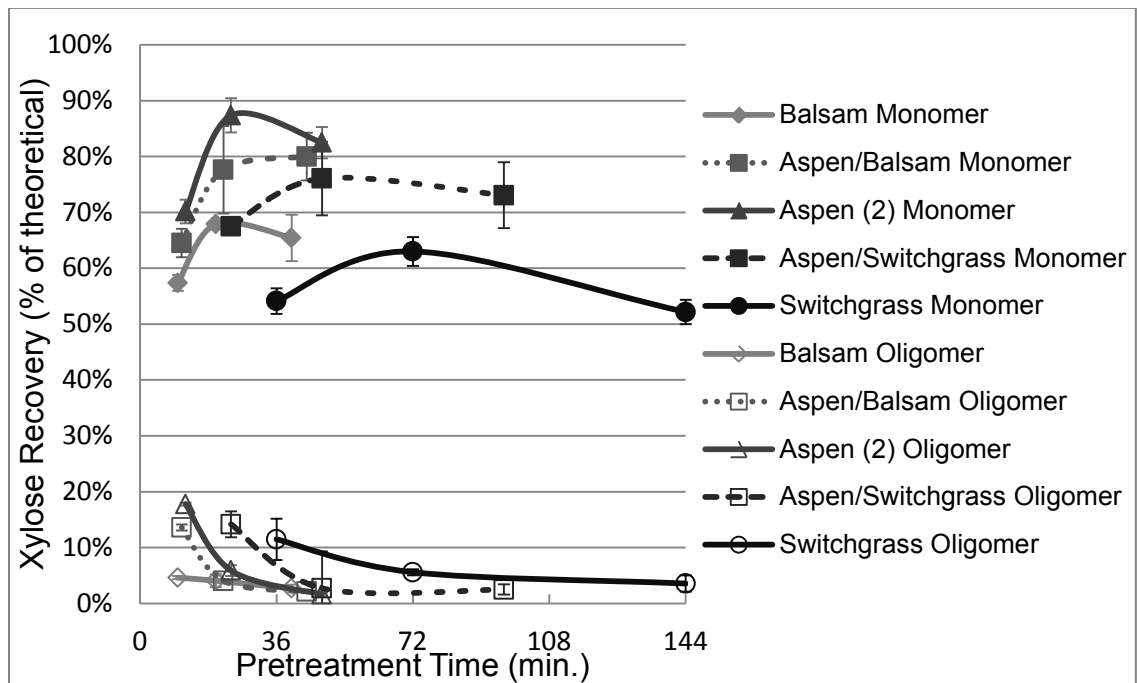
Figure 3.2: Total xylose yields (monomer plus dissolved oligomer expressed as monomer equivalents) following dilute acid hydrolysis for aspen, balsam, switchgrass, 50:50 aspen:balsam blend, and 50:50 aspen:switchgrass blend.

Peak xylose monomer plus oligomer yields for the dilute acid pretreatment varied from 69-93% of the theoretical maximum as calculated from species compositions. Pure switchgrass gave the lowest xylose recovery and pure aspen the highest. Each pure species and mixture showed a peak sugar recovery near the expected optimal time based on prior work (Jensen, Morinelly et al. 2010). Earlier time points showed incomplete hydrolysis with greater quantities of soluble oligomeric sugars in the hydrolyzate. Later time points exhibited lower yields of sugars due to production of decomposition products such as furfural and HMF in the hydrolyzate.

The first experiment with aspen showed incomplete hydrolysis compared to previous studies performed in our lab (compare aspen from Table 3.1 with aspen (1) from Table 3.3), with low xylose recoveries and large standard deviations. This experiment was repeated and the second attempt showed much higher xylose recovery and narrower standard deviations. Balsam and switchgrass xylose recoveries behaved in the expected manner from previous studies, matching the expected results given in Table 3.1. Switchgrass hydrolysis times were much longer than for aspen and balsam due to the acid-neutralizing effects of the relatively high ash content of switchgrass compared to aspen and balsam. Blends of aspen:balsam and aspen:switchgrass gave xylose recovery values in between the results obtained for the pure species (a result also shown in Jensen, Morinelly et al. 2010).

Some glucose, approximately 3-12% depending on species, is also liberated during dilute acid hydrolysis, presumably from the hemicellulose and a small amount of cellulose hydrolysis. However, most of the glucose is recovered from cellulose hydrolysis as described in Section 3.2.

Although the focus of this study was on the major sugars xylose and glucose, recoveries of the minor sugars may be relevant for a commercial biorefinery. The minor sugars galactose, arabinose, and mannose account for 1-5% of the total sugars present in the feedstock, except for balsam which contains over 11% mannose by weight. Recoveries for these sugars varied from approximately 20% for galactose in the first aspen experiment, to over 50% for arabinose plus mannose in the balsam experiment. The yield of galactose from the switchgrass sample shows greater than 100% recovery, indicating that the galactose peak may not have been completely separate from the xylose peak which elutes very close to it in time, and that maintenance of the HPLC may have been overdue. However, galactose represents only a small fraction, approximately 1%, of the total sugars in switchgrass, and therefore small differences in sugar measurement will result in large differences on a percent basis. No galactose, arabinose, or mannose was detected during enzymatic hydrolysis.

Furfural was quantified in the hydrolyzate from dilute acid pretreatment, and was seen to increase with pretreatment time resulting in lower xylose and higher furfural concentrations for the more severe pretreatment conditions. Furfural

concentrations varied from 0.04-2.88 g/l depending on species and pretreatment severity. Switchgrass, with the highest xylan content and the longest pretreatment time for any experiment, was seen to be most susceptible to furfural production with 2.88 g/l of furfural detected at the most severe pretreatment condition. Balsam, with the lowest xylan content and shortest pretreatment time, produced only 0.44 g/l of furfural at the most severe pretreatment condition.

Furfural concentrations at the optimal pretreatment condition were 0.49 g/l for the aspen (1) experiment, 0.60 g/l for the aspen (2) experiment, 0.18 g/l for balsam, 1.76 g/l for switchgrass, 0.36 g/l for the aspen:balsam blend, and 0.67 for the aspen:switchgrass blend. These levels compare favorably with the reported 2.4 g/l threshold of toxicity for yeast fermentation (Allen et al. 2010).

3.3.2 Enzymatic Hydrolysis

Residual biomass from each of the three reactors from the same pretreatment time point were loaded with differing enzyme concentrations – low, moderate, or high – in order to study the interaction between pretreatment severity and enzyme loading on sugar yields. Glucose yields from enzymatic hydrolysis were seen to vary from 3-60% after three days of hydrolysis. Figure 3.3 summarizes the results for enzymatic hydrolysis glucose yields for the optimum condition. Glucose and xylose mopnomer only were detected following enzymatic hydrolysis, except for balsam which produced only glucose. Graphs of glucose

and xylose yields for all reactions conditions are available in Section B.2 of

Appendix B.
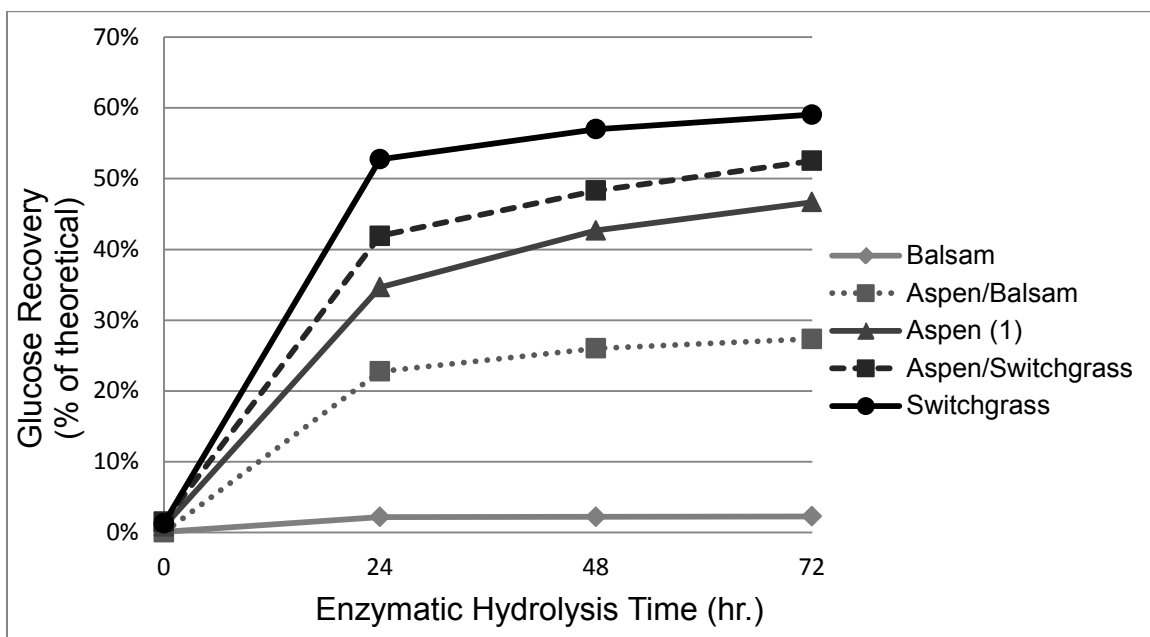
Figure 3.3: Glucose monomer recovery from enzymatic hydrolysis for aspen (1), balsam, switchgrass, 50:50 aspen:balsam blend, and 50:50 aspen:switchgrass blend at optimal conditions.

Enzymatic hydrolysis glucose yields at 72 hours of reaction were seen to be

strongly affected by both pretreatment severity and enzyme loading. The highest

enzyme loading for both the moderate and severe pretreatment condition always

yielded the highest glucose recovery, yet yields remained below 60% of

theoretical for all species.

Because of the relatively low glucose yields from enzymatic hydrolysis (<50% for

aspen at the optimal pretreatment time) the second aspen pretreatment

experiment was treated with Accelerase 1500 as previously described, but was

supplemented with additional β-glucosidase to investigate the possibility that the

enzyme mixture was limited by β-glucosidase activity. Both Accelerase BG (β-glucosidase supplement, Genencore) and Novozyme 188 (β-glucosidase from *Aspergillus niger*) were added at 1100 pNPG units per gram carbohydrate for Accelerase BG (one pNPG unit denotes 1 μmol of nitrophenol liberated from para-nitrophenyl-β-D-glucopyranoside per minute at 50°C and pH 4.8), and 120 β-glucosidase units per gram carbohydrate for Novozyme 188.

The additional β-glucosidase resulted in approximately 15% more glucose recovered at the highest enzyme loading and optimal pretreatment time (Figure 3.4). For Aspen with supplemental β-glucosidase nearly 100% of xylose was recovered following dilute acid pretreatment and enzymatic hydrolysis. The maximum glucose recovery for any experiment was 70% for switchgrass without supplemental β-glucosidase.
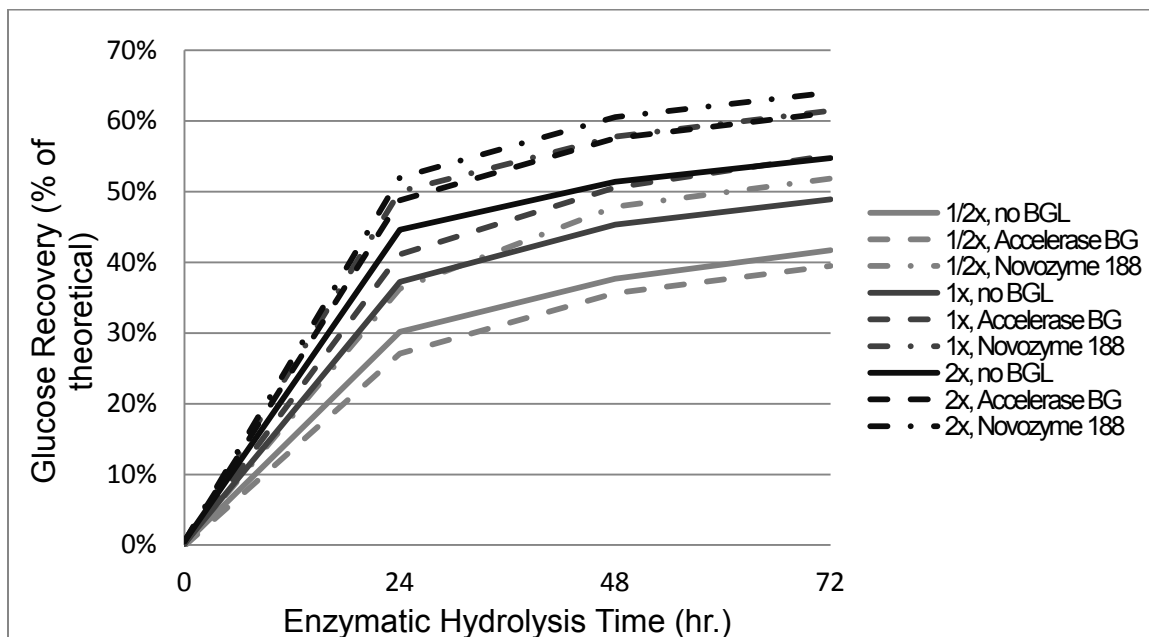
Figure 3.4: Glucose monomer recovery from enzymatic hydrolysis for aspen with supplemental β-glucosidase (aspen (2) from Table 3.3).

3.3.3 Total Sugar Recovery

Table 3.3 shows that the majority of sugar released during dilute acid

pretreatment is xylose, while glucose is the major sugar released from enzymatic

hydrolysis. Galactose, arabinose, and mannose are also generated, but in

amounts generally smaller than glucose (except for balsam where mannose was

a significant sugar product). A small amount of glucose is also released during

dilute acid pretreatment, presumably both from the hemicellulose and from a

small amount of cellulose hydrolysis. Although the most severe pretreatment

condition always resulted in higher glucose recovery after enzymatic hydrolysis,

the total sugar recovery is ultimately less due to the degradation of xylose at later

time points.

3.3.4 Sugar Recovery Model for Biomass Mixtures

A linear interpolation model was used to predict sugar recoveries for the

mixtures. Model predictions are presented in Table 3.4 below.

Table 3.4
Predicted and experimental sugar recoveries for mixed feedstocks.

	Dilute Acid Pretreatment		Enzymatic Hydrolysis		Total Sugars	
	Xylose[1]	*Glucose[2]*	*Xylose[1]*	*Glucose[2]*	*Xylose[1]*	*Glucose[2]*
Aspen/Balsam	81.7%	5.0%	8.2%	27.3%	89.8%	32.3%
Predicted	*82.6%*	*5.2%*	*2.2%*	*24.5%*	*85.9%*	*29.3%*
Aspen/Switchgrass	78.8%	5.4%	3.8%	52.5%	82.6%	58.0%
Predicted	*80.9%*	*7.6%*	*4.8%*	*53.0%*	*86.8%*	*60.2%*

[1] Predicted xylose recovery based on the aspen (2) pretreatment which had better recovery and lower variability in the data.
[2] Predicted glucose recovery based on the aspen (1) enzyme hydrolysis without supplemental β-glucosidase.

This model predicts sugar recovery within about 2% of the experimental data based on theoretical yields for both glucose and xylose from dilute acid hydrolysis. Glucose recovery from enzymatic hydrolysis can be predicted well to within 3%. The only exception to the good performance of this linear model is xylose recovery from enzymatic hydrolysis for the aspen/balsam blend which lies outside the expected range compared to what is predicted by the model. However, because xylose released from enzymatic hydrolysis makes only a minor contribution to the total sugars recovered, this does not significantly alter the results for total sugar recovery. Total sugar recovery is accurately predicted by this model to within 4%.

The recoveries achieved in this study for xylose from aspen and switchgrass are similar to previous work and results reported elsewhere in the literature. Yields of 70-85% recovery for hemicellulose sugars have been reported for switchgrass by Dien et al. (Dien et al. 2006) for similar pretreatment conditions. Glucose recoveries in this study are slightly lower than those reported elsewhere. Chung et al. (Chung et al. 2005) report 70% glucose recovery for aspen and 90% for switchgrass under enzymatic hydrolysis conditions of 60 FPU/g cellulose for eight days. Dien also reports glucose recoveries near 90% with 50 FPU/g cellulose and supplemental β-glucosidase. These results indicate that higher loadings of Accelerase 1500 and supplementation with β-glucosidase will be required to approach theoretical recoveries of glucose from cellulose for fermentation. Very few reports for balsam hydrolysis are reported in the literature

due to its low suitability for biochemical processing. Results for this study are comparable to previous work in our laboratory for balsam hydrolysis (Jensen et al. 2008; Morinelly et al. 2009; Jensen et al. 2010).

3.4 CONCLUSIONS

Mixtures of hardwood, softwood, and herbaceous feedstocks can be processed together by dilute acid and enzymatic hydrolysis with predictable results. Linear interpolation gives accurate results for total sugar recovery for mixtures to within 4% based on the pure species performance.   Enzyme loading and pretreatment severity are the most important factors for glucose recovery from enzymatic hydrolysis. Supplemental β-glucosidase increased glucose yields an average of 15%. Nearly 100% of xylose is recoverable for aspen with supplemental β-glucosidase. Maximum glucose recovery for any experiment was 70% for switchgrass without supplemental β-glucosidase, suggesting that it is less recalcitrant compared to aspen or balsam.

Chapter 4: Expression and Characterization of Three *Trichoderma reesei* Cellulose Hydrolases in *Kluyveromyces lactis* Yeast

4.1 INTRODUCTION

Enzymatic hydrolysis is currently the only viable means of converting cellulose to glucose for fermentation with high yield. In a commercial biorefinery, cellulase enzyme production is among the top three contributors to both operating costs (Humbird et al. 2011) and environmental burdens (MacLean and Spatari 2009). It is also one of the best opportunities for improving process economics. While pretreatment is the single biggest contributor to process costs, it is largely based around well-established technology and provides little room for improvement. Enzymatic hydrolysis is still a developing technology and holds greater potential for improvement (Wyman 2007).

Although the activity of many different enzymes on a solid substrate is a complex phenomenon and does not always follow traditional Michaelis-Menten kinetics (Mosier et al. 1999), a few generalizations can be made. Three categories of enzymes are important for the complete and efficient hydrolysis of cellulose to glucose; endoglucanases, exoglucanases, and β-glucosidases. Endoglucanases cleave cellulose chains at random internal location, generally acting most quickly on the amorphous regions, producing a greater number of chain ends for the exoglucanases to act on. Exoglucanases, also called cellobiohydrolases, attach

to chain ends and proceed along the chain cleaving off two-glucose units called cellobiose. β-glucosidases then hydrolyze the soluble cellobiose molecules into two molecules of glucose. The overall activity of the system can be modeled and has been shown to be dependant not only on enzyme and substrate concentration, but also on the degree of polymerization of the cellulose substrate and the fraction of β-glucosidic bonds accessible to the enzymes (Zhang and Lynd 2006b). In addition, product inhibition is also known to effect the activity of the enzyme system, with cellobiose strongly inhibiting the activity of exoglucanases, which is generally the rate-limiting step in natural systems. This makes the inclusion of β-glucosidases important for the optimization of synthetic enzyme systems, as glucose is a much weaker inhibitor of exoglucanase activity.

The filamentous fungus *Trichoderma reesei* one of the best studied cellulolytic microorganisms. It is a prodigious producer of cellulolytic enzymes, and has had its complete genome sequenced by the Department of Energy's Joint Genome Institute (Martinez et al. 2008), which is available online at http://genome.jgi-psf.org/Trire2/Trire2.home.html. In addition, *T. reesei* cellulases are among the most efficient free cellulases produced, making it an excellent starting point for cellulase enzyme studies.

In order to study individual cellulases it is desirable to produce them in a heterologous system without any native cellulose hydrolases. Although heterologous expression may involve some loss of enzyme activity due to non-

native processing, it does allow us to focus on a single enzyme and eliminates the problem of cross-contamination from other cellulolytic enzymes. Expression in other filamentous fungi such as *Aspergillus oryzae* is difficult due to lengthy transformation protocols (Takashima et al. 1998; Takashima et al. 1999). Expression in bacteria such as *Escherichia coli* is impractical because necessary post-translational modifications are not properly carried out by bacteria (Laymon et al. 1996; Okada et al. 1998b). Expression in the yeast *Saccharomyces cerevisiae* generally results in over-glycosylation of the enzymes and loss of enzyme function (Van Arsdell et al. 1987; Penttilä et al. 1988; Zurbriggen et al. 1990; Cummings and Fowler 1996; Saloheimo et al. 1997; Okada et al. 1998b). Expression has been performed in other yeast such as *Pichia pastoris* and *Schizosaccharomyces pombe* with limited success (Okada et al. 1998a; Okada et al. 1998b; Godbole et al. 1999; Boer et al. 2000).

For this study we have chosen to investigate the expression of *T. reesei* cellulases in the yeast expression system *Kluyveromyces lactis*. Although prior to beginning this study no reports had been made in the literature of *T. reesei* cellulases expressed in *K. lactis*, some work with xylanases and chitinases indicated the potential for success (Swinkels et al. 1993; Müller et al. 1998; van Ooyen et al. 2006).

4.2 MATERIALS AND METHODS

4.2.1 Trichoderma reesei Cultures

T. reesei strain QM 9414 was obtained from the American Type Culture

Collection and maintained on agar plates containing either potato dextrose media

(ATCC media 336) or 2 wt% malt extract, and subcultured weekly. A spore

suspension inoculum was prepared by rinsing a one week old sporulating plate

with 5 ml of sterile distilled water. The spore suspension was divided in 1 ml

aliquots and stored at -20°C for use inoculating liquid cultures. Spore

concentration in the suspension was determined to be approximately 2.5×10^6

spores per milliliter using an Improved Newbauer counting chamber.

4.2.2 T. reesei Induction Media

A minimal media with filter paper as the sole carbon source was adapted from

literature sources (Mandels and Weber 1969; Chen et al. 1987; Penttila et al.

1987) for growth and induction of the *T. reesei* cellulase system. A 100x trace

salt solution was prepared by dissolving 0.0500 g $FeSO_4 \cdot 7H_2O$, 0.0156 g

$MnSO_4 \cdot H_2O$, 0.0141 g $ZnSO_4 \cdot 7H_2O$, and 0.0290 g $CoCl_2 \cdot 6H_2O$ in 100 ml distilled

water and adjusting the pH to 2.0 with concentrated hydrochloric acid (to aid

dissolution). A 2x basal salt solution was prepared by dissolving 2 g KH_2PO_4, 1.4

g $(NH_4)_2SO_4$, 0.3 g urea, 0.3 g $MgSO_4 \cdot 7H_2O$, and 0.37 g $CaCl_2 \cdot 2H_2O$, plus 10 ml

of the trace salt solution, in distilled water to make a final volume of 500 ml. The

pH was adjusted to 5.0 with 10N NaOH and the solution was filter sterilized

through a 0.45 mm membrane filter. The induction media was prepared by combining one 7 cm diameter Whatman #1 filter paper (0.35 g), 0.05 g Bacto peptone, and 25 mL distilled water in a 250 mL Erlenmeyer flask, covering with aluminum foil, and autoclaving at 121°C for 15 minutes. After cooling to room temperature, 25 ml of the 2x basal salt solution was added aseptically to the filter paper peptone mixture to produce the induction media.

4.2.3 RNA Isolation and cDNA Synthesis

Fifty milliliters of rich media containing 1 wt% yeast extract, 2 wt% peptone, and 2 wt% dextrose (YPD) was inoculated with 50 μl of spore suspension and allowed to grow for two days at room temperature (25-30°C) with 150 RPM agitation on an orbital shaking platform. After two days growth, the mycelia were separated from the media by centrifugation at 2800 gravities RCF for five minutes and pouring off the supernatant. The mycelia were gently resuspended in 50 ml sterile distilled water, then centrifuged at 5500 gravities RCF for five minutes. The mycelia were then resuspended in 25 ml of 2x basal salt solution and added to 25 ml of filter paper peptone mixture to induce cellulase transcription. The cultures were allowed to grow for an additional three days at room temperature and 150 RPM agitation.

After three days growth on filter paper media the mycelia were harvested. Large pieces of filter paper which remained in the media were removed with tweezers, and the mycelia and media were transferred to a 50 ml centrifuge vial and

centrifuged at 19000 g relative centrifugal force (RCF) for two minutes. The

supernatant was poured off and 0.5 g of wet mycelia were removed, blotted with

a filter paper to remove excess moisture, and rapidly transferred to a mortar

chilled under liquid nitrogen and ground to a fine powder. 0.2 g ground mycelia

was transferred to a chilled 2 ml microcentrifuge tube and total RNA was

extracted using the RNeasy Plant Mini kit (Qiagen) according to manufacturer

directions. RNA yield was approximately 30 mg total RNA at a concentration of

500 ng/μl as determined by absorbance measurements taken at 260 nm on a

Nanodrop ND-1000 spectrophotometer (Thermo Scientific). First strand cDNA

synthesis was performed using the Omniscript Reverse Transcription kit (Qiagen)

according to manufacturer directions with a 15 residue oligo-dT primer.

4.2.4 Gene Amplification

Target genes were amplified from the first strand of cDNA using GoTaq Flexi

DNA polymerase (Promega) and the following primers and reaction conditions:

Primers

Cel6A sense 5'-TATGTTCTATGGTACCCAAGCTTGCTCAAGCGTCTGGGGC-3'

Cel6A antisense 5'-GGCCGCGCTTAATTAATTACAGGAACGATGGGTTTGCGTT-3'

Cel7A sense 5'-AATACTCGAGAAAAGACAGTCGGCCTGCACTCTCCAATCG-3'

Cel7A antisense 5'-GGACGCGCTTAATTAATTACAGGCACTGAGAGTAGTAAGG-3'

Cel7B sense 5'-ATCTACAGCAAGATCTCAGCAACCGGGTACCAGCACCCCC-3'

Cel7B antisense 5'-TTAGTTATGCGGCCGCCTAAAGGCATTGCGAGTAGTAGTC-3'

81

<u>PCR Reaction Conditions</u>

10 ng template DNA

400 nM each primer

0.5 unit Taq DNA polymerase

3 mM MgCl$_2$

200 mM dNTP

<u>10 µl 5x buffer</u>

50 µl total reaction volume

Initial Denaturation – 95°C for 2:00 minutes

Anneal – 60°C for 0:30

Extension – 72°C for 1:30   Repeat 40 times

Denature – 95°C for 0:30

Final Extension – 72°C for 5:00 minutes

Primers were designed to amplify each target gene without the native signal sequence which would normally direct secretion of the enzyme in *T. reesei*. Native signal sequences were determined to be 24 amino acids for Cel6A (Chen et al. 1987), 17 amino acids forCel7A (Shoemaker et al. 1983), and 22 amino acids for Cel7B (Penttila et al. 1986). Forward and reverse primers were also designed to allow for in-frame cloning of each gene into the pKLAC1 plasmid vector (New England Biolabs) at specific locations; KpnI/PacI for Cel6A, XhoI/PacI for Cel7A, and BglII/NotI for Cel7B. The genes were fused to the α-

mating factor signal sequence directing secretion of the protein in *K. lactis* yeast, and placed under the control of the inducible LAC4-PBI promoter. The use of the XhoI restriction site for cloning of Cel7A into the pKLACI plasmid resulted in a secreted protein with a native N-terminal sequence. The use of KpnI for Cel6A and BglII for Cel7B resulted in a secreted protein with a non-native N-terminal; the amino acid sequence EAEARRARSPRGT precedes the native N-terminal sequence for Cel6A, and EAEARRARS precedes the native N-terminal sequence for Cel7B. The use of the XhoI cloning site for these two proteins was not possible due to the existence of a XhoI restriction site within the cDNAs of those two proteins.

PCR reaction mixtures (GoTaq Flexi, Promega) were separated by electrophoresis on low gelling temperature agarose (Amresco) and visualized under UV light with ethidium bromide staining to confirm amplification of DNA products with appropriate molecular weights, Bands were excised with a clean scalpel and DNA was purified using the QIAquick Gel Extraction kit (Qiagen). Cel6A and Cel7B were further amplified through a second round of PCR as above, using the first round product as the template. After purification, all fractions of the same gene product were pooled for further use.

4.2.5 Plasmid Construction

Plasmid pKLAC1 (shown in Figure 4.1 below) was obtained from New England Biolabs for use with their *K. lactis* expression system.
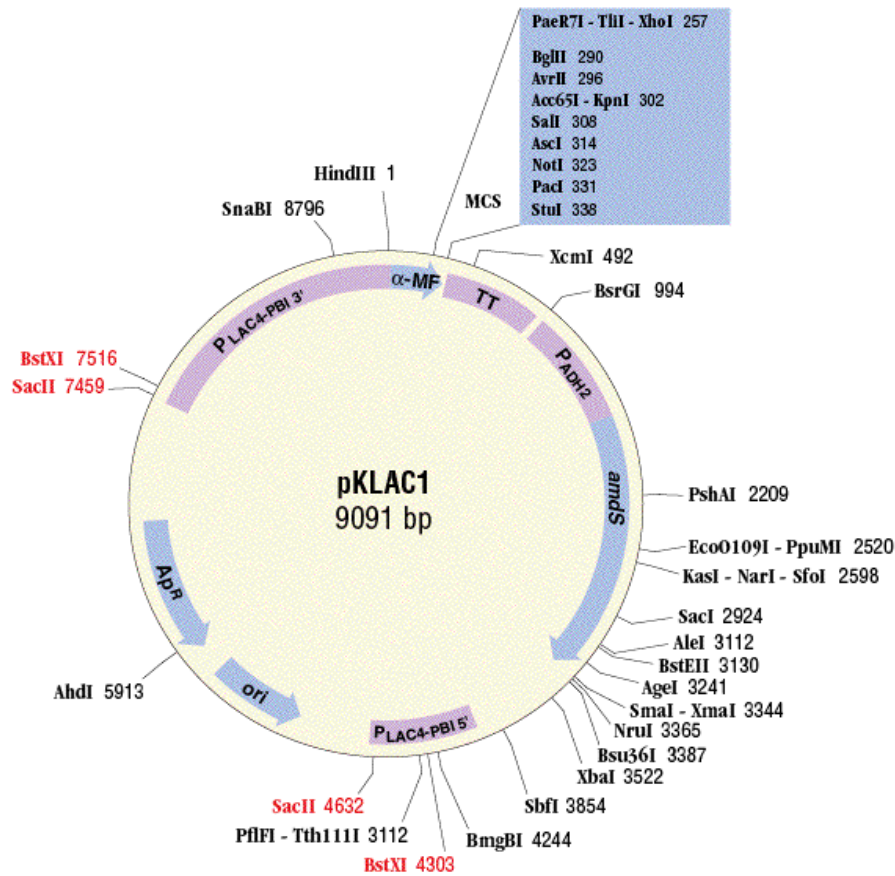
Figure 4.1: Plasmid pKLAC1 from New England Biolabs.

The pKLAC1 plasmid is a multi-functional plasmid containing regions for selection and propagation in *E. coli*, as well as an expression cassette for insertion into the *K. lactis* genome. The expression cassette uses the $P_{LAC4-PBI}$ promoter to induce enzyme expression in the presence of galactose. It contains portions of both the 5' and 3' region in order to target insertion into the LAC4 region of the *K. lactis* genome. In addition it possesses the α-mating factor (α-MF) signal sequence to direct secretion of the enzyme product, a multiple cloning site (MCS) for insertion of the gene under study, and the transcription termination (TT) region. Also on the expression cassette is a constitutive promoter driving

constant expression of the acetamidase gene for selection of positive transformants. The acetamidase gene confers the ability to utilize acetamide as a nitrogen source for cell growth.

Plasmid and gene products were each digested with appropriate restriction enzymes (New England Biolabs) to prepare for ligation. Five hundred nanograms (500 ng) of insert DNA and 500 ng of plasmid DNA were digested overnight for each ligation reaction. Digested DNA was then purified on low gelling temperature agarose as above. Following purification a ligation reaction containing 75 ng of vector DNA and 25 ng of insert DNA (for a 2:1 molar excess of insert:vector DNA) was incubated with T4 DNA ligase in a total reaction volume of 10 µl at 16°C overnight. Following ligation, 5 µl of the reaction mixture was used to transform NEB Turbo competent *E. coli* (New England Biolabs) with the plasmid construct DNA according to manufacturer's protocol. *E. coli* were grown on LB agar containing 100 µg/ml ampicillin to select for positive transformants. Colonies were selected and grown in suspension cultures in LB media with ampicillin. Cells were pelleted by centrifugation and plasmid DNA was purified using the Plasmid MaxiKit from Qiagen. Plasmid DNA samples were sent to Nevada Genomics for sequencing, and the gene products were found to be identical to published sequences each gene.

4.2.6 K. lactis Transformation

K. lactis strain GG799 competent cells were also obtained from New England Biolabs for enzyme expression studies. Purified plasmid was linearized with SacII and separated on low gelling temperature agarose. The larger fragment (containing the gene expression cassette) was excised and purified as above. Purified DNA was then inserted into the *K. lactis* competent cells by chemiporation and thermal shock according to the manufacturer's protocol. Transformants were grown under nitrogen limiting conditions and selected for the utilization of acetamide as a nitrogen source. After two rounds of selective growth, positive transformants were grown in rich media containing galactose (YPGal) and screened for enzymatic activity.

4.2.7 Enzyme Purification and Concentration

Size exclusion chromatography (SEC) was performed using Toyopearl HW-40F (Tosoh Biosciences), a hydroxylated methacrylic polymer resin with a molecular weight cut-off of 10 kDa, to de-salt the cell culture supernatant for further processing. A column approximately 2.5 cm diameter by 60 cm length was packed with resin giving a bed volume of 295 ml. This produces a working volume (approximately 40% of the bed volume) sufficient to process 100 ml of cell culture supernatant at a time. The resin was equilibrated with 20 mM Tris-Cl, pH 8.0 and the cell culture supernatant was pumped through the column at a rate of 1-10 ml/min. with fractions collected using a Gradi-Frac fraction collector (Pharmacia Biotech).

Anion exchange chromatography (AEX) was performed using Toyopearl DEAE-650M (Tosoh Biosciences). One to ten milliliters of AEX resin was packed by gravity flow into a column at least twice that volume and equilibrated with 20 mM Tris-Cl, pH 8.0. The de-salted cell culture supernatant from SEC was allowed to flow by gravity through the AEX resin bed for the protein of interest to bind. Proteins were then eluted from the AEX resin using gradient elution from 0-0.5 M NaCl in 20 mM Tris-Cl, pH 8.0. One to two milliliter fractions were collected and stored at -20 C. Protein concentration in the fractions was determined by the Bradford method (Bradford 1976) prior to performing assays for protein specific activity.

4.2.8 Enzyme Activity Assays

Endoglucanase Cel7B specific activity assays were performed on carboxymethyl cellulose (CMC, from TCI America) as the substrate, using the dinitrosalicylic acid (DNS) method for estimating reducing sugars as described by Wood and Bhat (Wood and Bhat 1988). Briefly, enzyme is incubated with 1% CMC in 50 mM citrate buffer, pH 4.8, 1 ml total reaction volume, at 50°C with 250 RPM shaking. After 30 minutes 3 ml of DNS reagent are added to stop the reaction, and the solution is boiled for five minutes. Finally, 20 ml of distilled water is added, the absorbance is measured at 540 nm, and compared to a glucose standard curve in the range of 0-1 mg/ml glucose.

Exoglucanase activity was determined on Avicel PH-101 (Fluka) powdered cellulose as the substrate using an Agilent 1200 series HPLC for detection of cellobiose production. Avicel at 4 mg/ml was suspended in distilled water by magnetic stirring, and 250 µl was pipette into a 2 ml screw-top microcentrifuge tube. 1 M citrate buffer pH 4.5, tetracycline 10 mg/ml, cycloheximide 10 mg/ml, enzyme, and distilled water were added to give the following final reaction condition: 50 mM citrate pH 4.8, tetracycline 0.04 mg/ml, cycloheximide 0.03 mg/ml, Avicel 1 mg/ml, enzyme 0.1 mg/ml, total reaction volume 1 ml. Microcentrifuge tubes were capped and taped flat to the surface of a Lab-Line Environ-Shaker orbital shaker table, and incubated at 50°C and 250 RPM. Samples were taken at 0, 24, 48, and 72 hours to track the production of cellobiose over three days. To take samples, reaction tubes were centrifuged at maximum speed in a Galaxy 16 benchtop microcentrifuge (VWR) for two minutes. One hundred microliter samples were pipeted from the surface and transferred to 250 µl HPLC vials, then the reaction vials were vortex mixed and returned to the incubated shaker table. Cellobiose concentration was determined by refractive index measurements following sugar separation on a Bio-Rad Aminex HPX-87P analytical column.

4.3 RESULTS AND DISCUSSION

Amplification of the genes by PCR was confirmed by agarose gel electrophoresis and ethidium bromide staining to visualize DNA products and determine their size (Figure 4.2 below).
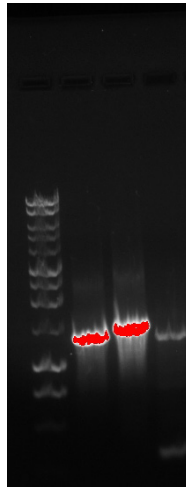
Figure 4.2: DNA amplified by PCR with gene-specific primers, stained with ethidium bromide and pictured under UV light. Lane 1, 1 kb molecular weight marker (Promega); Lane 2, Cel6A amplification product; Lane 3, Cel7A; Lane 4, Cel7B.

Gene product sizes were calculated to be 1376 bp for Cel6A, 1526 bp for cel7A, and 1337 bp for Cel7B, based on the size of the gene sequence being amplified plus the addition of non-overlapping portions of the forward and reverse primers.. After purification, DNA concentration was determined by spectrophotometry on a Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and calculated to be 61 ng/µl for Cel6A, 103 ng/µl for Cel7A, and 12 ng/µl for Cel7B.

Following transformation, positive transformants of *K. lactis* containing

chromosomally integrated copies of Cel6A, Cel7A, Cel7B, or a maltose binding

protein (positive control) were taken from plates containing selective media and

inoculated into 50 ml of rich media containing 1% yeast extract, 2% peptone, and

2% galactose (YPGal). Cultures were grown for two days at 28-30°C and 250

RPM shaking. Following growth, cells were separated by centrifugation and the

supernatant was filtered through a 0.2 μm membrane filter. Fifty microliters of

cell culture supernatant was spotted onto a plate containing 0.5% carboxymethyl

cellulose (CMC), 0.7% agar, 50mM citrate pH 4.8, and allowed to incubate

overnight at room temperature. The following morning the plate was stained with

0.3% Congo Red solution for 30-60 minutes, then destained by flooding the plate

with 1 M NaCl and allowing it to soak for 15-30 minutes to permit the excess

Congo Red stain to diffuse out of the solid agar media. Destaining was

performed 3-5 times. Results are shown below in Figure 4.3.
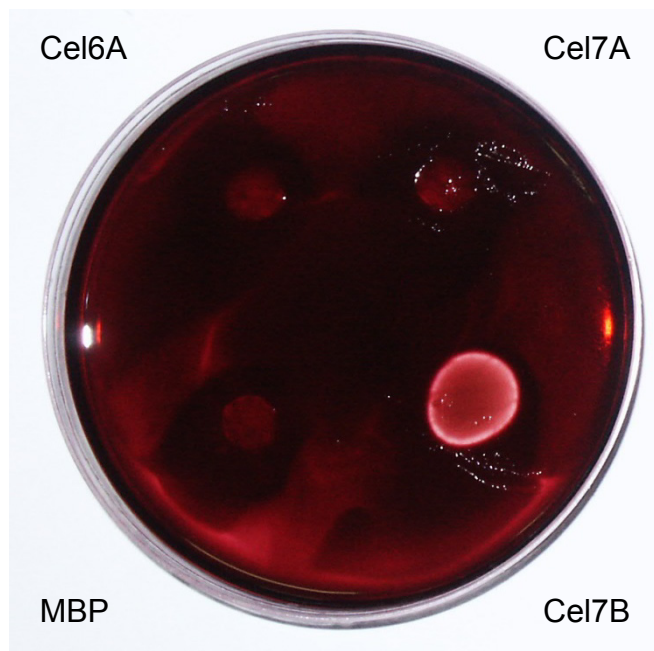
Figure 4.3: 50 µl of cell culture supernatant spotted onto a plate containing 0.5% carboxymethyl cellulose (CMC), 0.7% agar, 50mM citrate, pH4.8. Clockwise from upper left; Cel6A, Cel7A, Cel7B, and maltose binding protein (MBP). The plate was stained with 0.3% Congo Red solution and rinsed several times with 1M NaCl.

The pronounced clearing around Cel7B indicates endoglucanase activity. Endoglucanases are capable of hydrolyzing CMC, reducing the binding affinity of CMC and Congo Red, resulting in the observed clearing zone in areas where Cel7B has been active. Exoglucanases are not capable of hydrolyzing CMC due to steric hindrances, making this activity assay specific to endoglucanases.

Cel7B was purified and concentrated by SEC and AEX as described above in the methods section. Elution fractions 7-16 from AEX were visualized by sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE) with Coomassie Brilliant Blue staining, shown below in Figure 4.4.
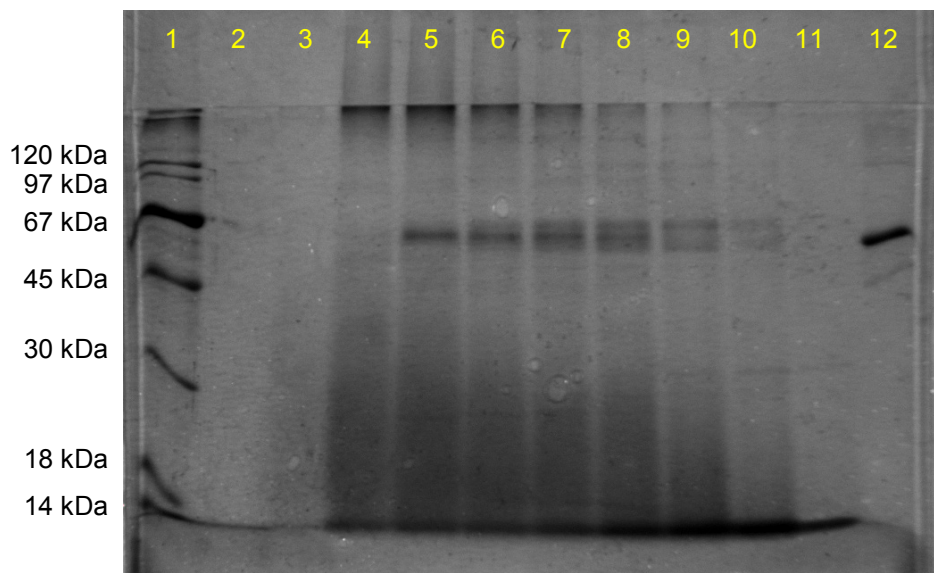
Figure 4.4: SDS-PAGE of Cel7B elution fractions. Lane 1 at the left is a 14-120 kDa molecular weight marker (Rockland Inc.) with bands at 14, 18, 30, 45, 67, 97, and 120 kDa. Lanes 2-11 are Cel7B elution fractions 7-16, and lane 12 is a bovine serum albumin standard at 67 kDa.

Recombinant Cel7B shows a molecular weight near 67 kDa, heavier than the molecular weight of 54 kDa for the native protein reported in the literature (Shoemaker et al. 1983; Penttila et al. 1986). Recombinant Cel7B also appears to be produced in a tight but continuous distribution of molecular weights, which are separable by sodium chloride gradient elution from AEX. The differences in molecular weight are most likely due to the differing glycosylation patterns seen in yeast, which have been shown to effect enzyme activity. Earlier AEX fractions show a single band at the middle of the molecular weight distribution, with later elution fractions enriched in both the lighter and heavier molecular weight species.

Protein concentration in each elution fraction was quantified by the Bradford

method (Bradford 1976). Specific activity was determined by CMC-DNS assay

as described above in the methods section. Results are summarized in Table

4.1 below.

Table 4.1
CMC/DNS assays of Cel7B AEX fractions

Elution Fraction	Concentration (mg/ml)	Specific Activity (M glucose.g^{-1}.h^{-1})
7	0.07	N/D
8	0.28	N/D
9	1.36	0.542
10	6.29	0.345
11	6.90	0.286
12	12.18	0.137
13	14.71	0.057
14	17.48	N/D
15	17.75	N/D
16	14.95	N/D

The highest specific activity was found in the earlier elution fractions, which

corresponds to the mid-molecular weight species. The highest specific activity

measured here was approximately one-fifth of the value reported for the native

protein by Karlsson et al. (Karlsson et al. 2001).

SEC and AEX were performed on recombinant Cel6A in the same manner as for

Cel7B. Various purification fractions were visualized by SDS-PAGE, shown

below in Figure 4.5.
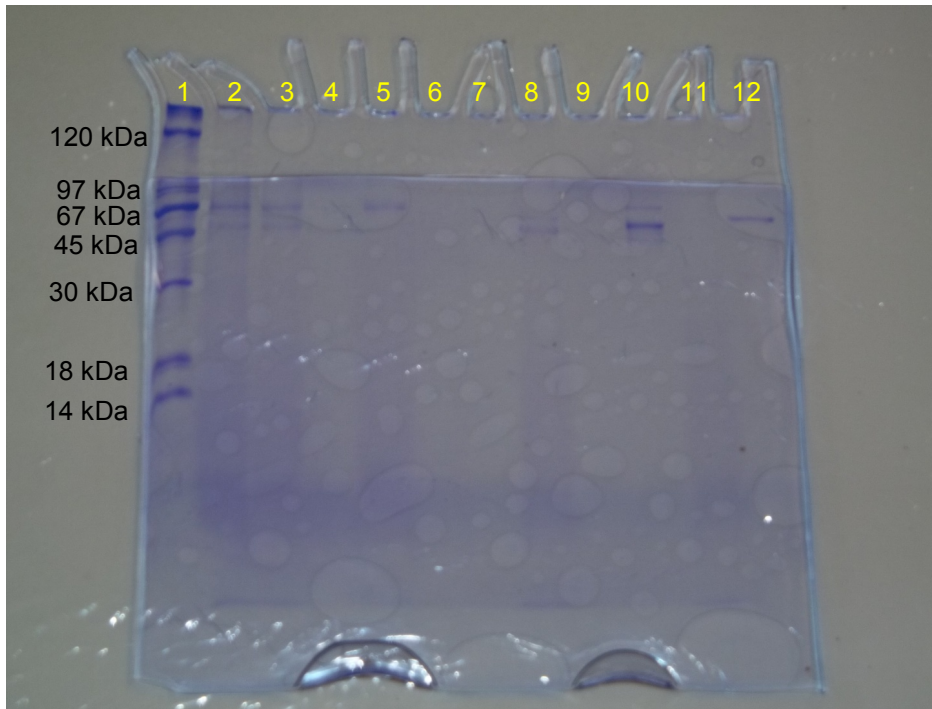
Figure 4.5: SDS-PAGE of Cel6A purification fractions. Lane 1 is a molecular weight marker (Rockland Inc.), lane 2 is the crude cell culture supernatant, lane 3 is the desalted cell culture supernatant fromSEC, lane 5 is the flowthrough from AEX, lane 8 is the eluant from AEX, lane 10 is Accelerase 1500, and lane 12 is a bovine serum albumin standard at 67 kDa.

Unlike Cel7B, recombinant Cel6A is produced by *K. lactis* in three distinct

molecular weight bands at approximately 50, 60, and 65 kDa. This is relatively

close to the 53 kDa molecular weight reported in the literature for the native

protein (Bhikhabhai et al. 1984; Sandgren 2003). Also, it was discovered that

only the 50 and 60 kDa species bound to the AEX resin under these conditions.

The heaviest 65 kDa molecular weight species was found to be in the

flowthrough from AEX. Exoglucanase activity assays were performed on the

Cel6A concentrated by AEX but no activity was detected. The flowthrough from

AEX was to dilute to assay directly.

Flowthrough from AEX containing the 65 kDa molecular weight species of Cel6A

was concentrated by ultrafiltration in a VivaCell 70 protein concentrater (Sartorius

Stedim) with a 10 kDa molecular weight cutoff membrane. Cell culture

supernatants from colonies producing Cel6A, Cel7A, and Cel7B were also

concentrated in the same manner. Exoglucanase activity assays were

performed on all samples and the results are shown below in Figure 4.6.
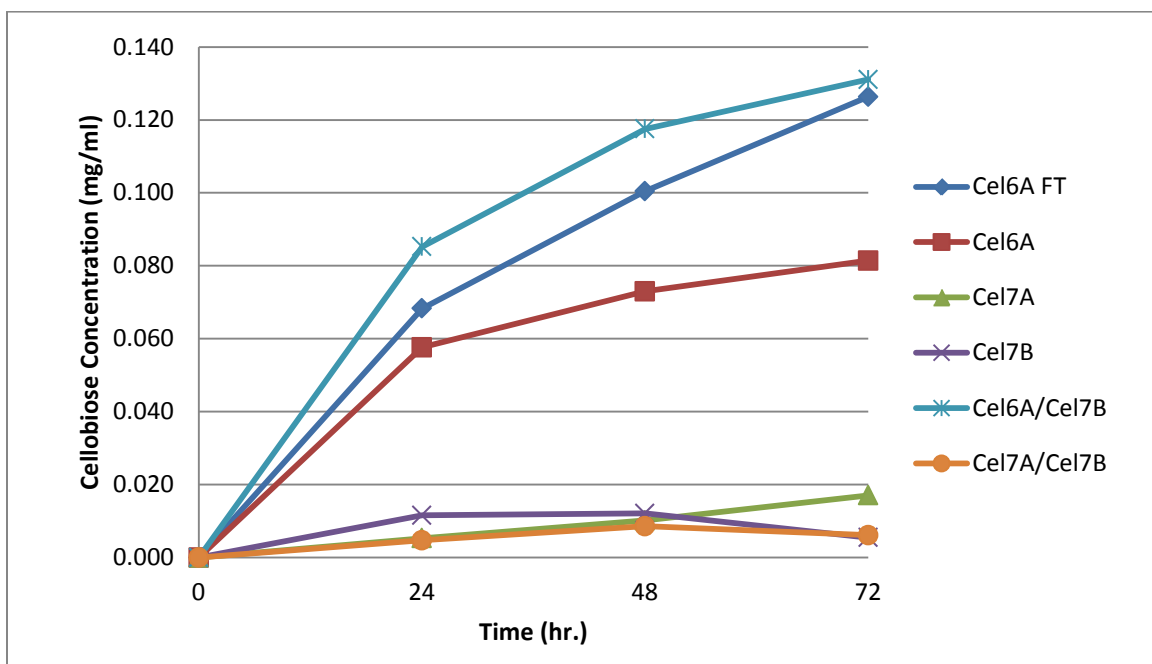
Figure 4.6: Exoglucanase activity assays on Avicel powdered cellulose.

The highest production of cellobiose was found for the combination of

Cel6A:Cel7B at a mass ratio of 10:1. This is due to synergies between the

endoglucanase and exoglucanase resulting in more efficient hydrolysis. The

next highest production of cellobiose was found for Cel6A flowthrough, which

contains only the heaviest 65 kDa molecular weight species. Cel6A cell culture

supernatant containing all three molecular weight species also showed cellobiose

production, but at a lower rate per milligram total protein than the AEX

flowthrough. The endoglucanase Cel7B by itself showed virtually no cellobiose

production, nor did the other exoglucanase Cel7A either alone or in combination

with Cel7B. Assays were repeated and standard deviations calculated to

determine the statistical signifacance of different reaction conditions, and the

results are shown below in Figure 4.7.
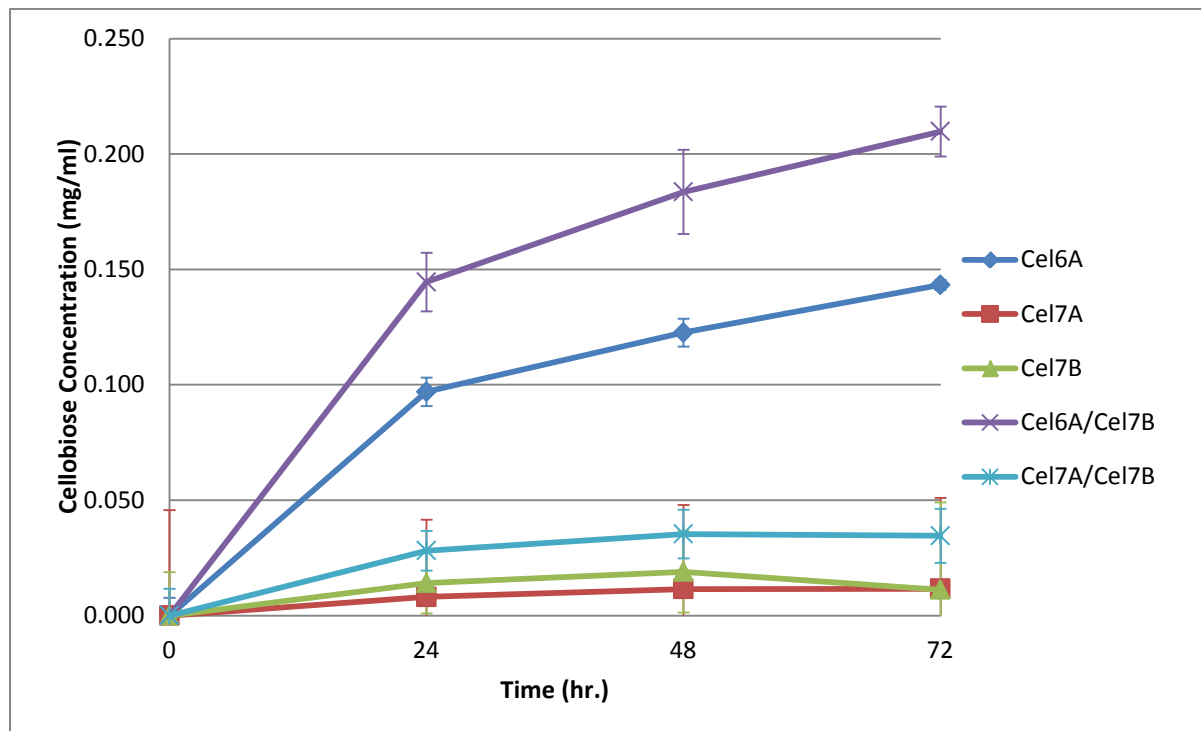
Figure 4.7: Repeatability of exoglucanase activity on Avicel powdered cellulose. Error bars represent one standard deviation; for single enzymes N=4, for mixtures N=2.

Assays were performed for cell cultures supernatants from two different

transformed colonies of *K. lactis*. Assays for the individual enzymes were

performed in duplicate, giving four replicates for individual enzyme assays.

Assays for combinations of enzymes were performed singly due to limitations in

the amount of enzyme available.   These results show that the combination of

Cel6A and Cel7B performs significantly better than Cel6A alone. They also show that *T. reesei* Cel7A is inactive when expressed in *K. lactis* yeast.

4.4 CONCLUSIONS AND RECOMMENDATIONS

In this study we have shown that *K. lactis* yeast is a suitable expression host for some *T. reesei* cellulases. We have demonstrated endoglucanase activity of recombinant Cel7B on the artificial cellulose analog carboxymethyl cellulose (CMC) in a plate screening assay using Congo Red staining, which is a valuable tool for rapid screening of large numbers of endoglucanases. We have also measured the specific activity of Cel7B on CMC using dinitrosalicylic acid (DNS) detection, and determined that the specific activity of the recombinant enzyme is on the same order of magnitude as the native enzyme, though several fold lower specific activity was detected. We have also demonstrated activity of Cel7B on powdered cellulose by showing increased cellobiose production from exoglucanase activity in the presence of a small amount of the endoglucanase Cel7B. Cel7B appears to be produced in a narrow but continuous distribution of differing molecular weights, with some forms of the enzyme showing higher specific activity than others, most likely due to differing patterns of glycosylation affecting the molecular weight and activity of the enzyme.

T. reesei Cel6A has also been shown to be produced in an active form by *K. lactis*. Unlike Cel7B, Cel6A is produced in three distinct molecular weight

species, with only the largest form displaying any enzymatic activity. Enzymatic activity of Cel6A has been demonstrated by cellobiose production on Avicel powdered cellulose, with detection and quantification by HPLC analysis.

The *T. reesei* exoglucanase Cel7A has also been expressed in *K. lactis* and was seen to be produced in a narrow but continuous distribution of differing molecular weights like Cel7B by SDS-PAGE (data not shown). However no enzymatic activity was detected for the recombinant Cel7A by the assay methods described in this paper.

Future work to optimize the ratio of Cel6A to Cel7B would be valuable to determine whether the optimum ratio for the heterologously expressed proteins is similar to what is found for the native enzyme system. Additionally, there exist four other endoglucanases from *T. reesei*, and two β-glucosidases, which contribute to efficient cellulose hydrolysis in the native system, and would be valuable to add to our lab system to more rigorously quantify the benefit of each for a commercial biorefinery. Knowledge of the extent of glycosylation of these proteins would be beneficial, and may offer some explanation for the lack of enzymatic activity seen in the heterologously expressed Cel7A. Finally, the ability to express and measure activity for individual proteins offers the opportunity for mutagenesis studies to potentially create and identify an enzyme with improved activity for industrial use.

Chapter 5: Co-expression of *Trichoderma reesei* Cellulose Hydrolases in *Kluyveromyces lactis* Yeast for Rapid Enzyme System Studies

5.1 INTRODUCTION

Lignocellulosic biomass is a complex substrate which requires the action of many different enzymes for complete and efficient hydrolysis.   Cellulase enzyme production in a commercial biorefinery is among the major contributors to both operating costs (Humbird et al. 2011) and environmental burdens (MacLean and Spatari 2009), and also one of the best opportunities for process improvement (Wyman 2007).

Because of the large number of enzymes involved, the number of possible enzyme combinations is truly astronomical, and therefore a high-throughput system will be necessary for effective study. However, many barriers exist to conventional high-throughput analysis of biomass deconstruction: the activity of any individual enzyme is difficult to detect on natural substrates, artificial substrates exist for some enzymes but not all, and the activity on artificial substrates is not necessarily representative of their activity on natural substrates (Zhang et al. 2006a). In addition, many different types of sugars are released, and no high-throughput sugar detection method is available that works well for all of them. Finally, for consolidated bioprocessing (CBP) which is the ultimate goal of this research, there is the issue of sugar uptake and utilization by the

fermenting microorganism, which is not even considered by most hydrolysis studies (Lynd et al. 2005).

Previous research on improving enzymatic hydrolysis rates generally follow one of two lines; blending and optimization studies of enzyme systems, and improvements in individual enzyme activity. Research into enzyme blend optimization is progressing with new microplate techniques developed for high throughput enzymatic hydrolysis studies (Chundawat et al. 2008; King et al. 2009), such as the GENPLAT system at the Great Lakes Bioenergy Research Center <www.glbrc.org> (Walton et al. 2011). Robotic multi-channel pipetting of powdered cellulose slurries and varying enzyme blends, coupled with automatic data acquisition on microtiter plate readers allows for accuracy, repeatability, and high assay throughput. However, microtiter plate assays require high-purity single enzyme stock solutions for blending, which are difficult and time-consuming to produce. Also, only existing enzymes can be blended, which does not address the fundamental issue of low enzyme specific activity.

Studies into increasing the specific activity of individual enzymes have been hampered by the lack of a useful assay of individual enzymatic activity on natural substrates. Individual enzymes have little effect on complex natural substrates; it is the synergistic action among the enzymes which allows for efficient hydrolysis of lignocellulosic biomass. Artificial substrates do exist which can detect some enzyme activities, however they are not good indicators of enzyme activity on

natural substrates (Zhang et al. 2006a). Enzyme improvement research has therefore focused primarily on improving enzyme thermostability as measured by activity on artificial substrates. This is a characteristic which is more easily defined, and therefore more easily measured. Voutilainen et al. (Voutilainen et al. 2007) have used random mutagenesis and robotic screening to develop a *Melanocarpus albomyces* cellobiohydrolases with increased residual activity on the small soluble substrate methylumbelliferyllactoside (MULac) at 70°C, compared to the native enzyme under the same conditions. However the novel enzyme still had a lower activity at both 50°C and 70°C than the native enzyme at 50°C. Two groups (Sandgren et al. 2005; Heinzelman et al. 2009) have used structure-guided recombination to develop enzymes with improved performance at higher temperatures. Also, one group at Shenzhen University in China (Liu et al. 2006) created a *T. reesei* endoglucanase – Cel61A – with higher catalytic activity, by creating a fusion protein containing two catalytic domains attached to a single carbohydrate binding molecule. Although structure-guided recombination is a systematic approach to enzyme engineering, it has a major drawback in that it requires extensive structural and functional information for every enzyme.

In order to study, and eventually improve hydrolysis of lignocellulosic biomass, we would like to be able to study individual enzymes in the context of the entire enzyme system. It is also desirable to be able to produce them in a heterologous expression system without any native cellulose hydrolases. This allows us to

focus on a single enzyme and eliminates the problem of cross-contamination from other cellulolytic enzymes. An ideal heterologous expression system would be one that can express multiple enzymes simultaneously for rapid evaluation of enzyme systems, and one in which we can also modify a single enzyme individually to study its effect on the overall system performance.

For this research we will be using the *Kluyveromyces lactis* yeast expression system. *K. lactis* is a well-characterized non-*Saccharomyces* yeast used in the food and biotechnology industry, enzymes produced in K. lactis have GRAS (generally regarded as safe) status with the FDA, and protocols for genetic transformation and protein expression in *K. lactis* are well-established (Romanos et al. 1992; Swinkels et al. 1993; van Ooyen et al. 2006). It has previously been used to express three cellulases and two xylanases from other sources (van Ooyen et al. 2006). Prior work (Brodeur-Campbell and Shonnard in prep) has also demonstrated that it is capable of producing *Trichoderma reesei* cellulases in active forms, without the extreme degree of hyperglycosylation often seen in *Saccharomyces*-expressed cellulases

One of the unique aspects to research reported here is the idea of simultaneously co-expressing multiple enzymes in a single yeast strain for laboratory research purposes, and determine enzymatic activity on Avicel powdered cellulose as the substrate. Although Avicel is a highly crystalline form of cellulose, it still provides enough reactive locations for endoglucanase activity

to demonstrate synergy between the enzymes (Zhang and Lynd 2006b), and can be dispensed with good repeatability by pipetting as a slurry. The *K. lactis* expression system also has available to it two very similar research plasmids with two different selection mechanisms; acetamidase activity which confers the ability to utilize acetamide as a nitrogen source for cell growth, and antibiotic resistance which allows the yeast to grow in the presence of G418 antibiotic. The two different selection mechanisms behave very differently in the yeast, with significant consequences for how they can be employed. Acetamide utilization favors multi-copy insertion events with a copy-number mode at three copies, most likely because the increased ability to utilize acetamide confers increased growth ability up to three copies, after which the metabolic burden of increased acetamidase production becomes a significant hindrance to growth.   Antibiotic resistance selection favors single-copy insertion events, probably due to little improvement in growth ability from a high level of expression of the antibiotic degrading enzyme. This allows us to perform either single- or multi-gene insertions, and to do them sequentially or simultaneously. It has previously been demonstrated that with acetamide selection, multiple different genes can be inserted simultaneously into the genome (Read et al. 2007). This provides us with the ability to create a multi-enzyme cellulolytic system in the yeast in a single step. Following this with a second round of transformation and selection using antibiotic resistance will allow us to insert one additional gene to study the effect it has on overall enzyme system performance.

A second unique aspect to this research is the idea to use yeast growth rate as the primary indicator of cellulose hydrolysis and sugar utilization.   This has the potential to eliminate complicated or time-consuming enzyme purification and sugar detection methods. Additional advantages include the just-in-time production and utilization of enzymes at the time and place where they are needed for optimal cell growth. This opens up the possibility of capturing enzyme synergies not realizable in separate hydrolysis and fermentation (SHF) scenarios, such as the ability to screen lignin peroxidases, which have been shown to degrade activity of other enzymes over extended incubation times, and to expand our search beyond traditional cellulolytic enzymes into enzymes with other functions which may be beneficial for the yeast, such as transporter proteins to increase sugar uptake and utilization. All of these factors are important considerations for the ultimate goal of consolidated bioprocessing, that of being able to create a single microorganism which can hydrolyze biomass and convert it to useful products, such as ethanol and high value chemicals.

K. lactis is also a close relative of the yeast *Kluyveromyces marxianus*, which is one of the most promising candidates for consolidated bioprocessing (CBP) of lignocellulosic biomass. *K. marxianus* possesses many traits that make it an excellent candidate for lignocellulosic ethanol production. It has the ability to grow and ferment at temperatures above 40°C (Anderson et al. 1986), which is advantageous for increasing the activity of cellulase enzymes and decreasing the ability of contaminating microorganisms to flourish. It is resistant to common

fermentation inhibitors from biomass pretreatment, such as aldehydes and

aromatic compounds, and has a high rate of aldehyde uptake and conversion to

alcohols compared to *Saccharomyces cerevisiae* (Oliva 2003). It has a high

ethanol tolerance and high ethanol production rate, producing ethanol at 45°C at

a rate similar to *S. cerevisiae* at 30°C (Nonklang et al. 2008). And it is known to

utilize all the major sugars liberated from lignocellulosic biomass, including

cellobiose, xylose, mannose, galactose, and arabinose (Anderson et al. 1986;

Nonklang et al. 2008; Rodrussamee et al. 2011). Unfortunately little else is

known about *K. marxianus*; the genome has not been sequenced, no type strain

has been adopted, and there is little other accumulated information

(Rodrussamee et al. 2011). However, using *K. lactis* as a model system,

improvements made to biomass hydrolysis and sugar utilization in *K. lactis*

should be highly relevant to the needs of *K. marxianus*.

The objective of this study is to show proof of concept for using the two different

selection mechanisms to simultaneously co-express multiple cellulase enzymes

in *K. lactis*. This research demonstrates the ability to create a basic cellulolytic

capacity in the yeast, and then to build an improved system on top of the basic

system, and show the ability to detect an improvement in hydrolytic performance

due to the insertion of one additional gene into the yeast genome.

5.2 MATERIALS AND METHODS

5.2.1 Strains, Vectors, and Media

In order to show growth of *K. lactis* associated with cellulose hydrolysis, yeast growth rates on cellobiose – the end product from cellulose hydrolysis in these experiments – will have to be shown to be differentiable from the cell growth rate on galactose, which must be present to induce recombinant enzyme production. *K. lactis* strain 22A295-1, described by Meyer et al.(Meyer et al. 1991), was obtained from the American Type Culture Collection (ATCC # MYA-2288, denoted hereafter as strain 2288). This strain has had the main galactose metabolic pathway disabled by interrupting the galactokinase gene, but it still retains the induction function in the presence of galactose. Strain 2288 is still capable of metabolizing galactose; however, the rate of utilization has been greatly reduced. This strain also has a requirement for supplemental uracil and adenine in the growth media, as it is incapable of synthesizing those two molecules autotrophically.

K. lactis strain GG799 (denoted hereafter as strain 799) was obtained under commercial license from New England Biolabs. This is essentially a wild-type strain capable of metabolizing galactose normally, and requiring no supplementation of the media. Plasmid pKLAC1 was obtained under the same license for use with the *K. lactis* expression system, as previously discussed in Chapter 4.

The research plasmid pGBN19 was generously provided by Dr. Jeremiah Read at New England Biolabs. The pGBN19 plasmid, described by Read et al.(Read et al. 2007), is similar to the pKLAC1 plasmid, except that the *amdS* acetamidase reporter gene has been replaced by the *Neo*[+] reporter gene which confers resistance to G418 antibiotic in yeasts.

YPD media is 1 wt% yeast extract, 2 wt% peptone, and 2 wt% dextrose. Rich media is YP (1 wt% yeast extract, 2 wt% peptone), and contains other carbohydrate sources as noted. Minimal media contains yeast nitrogen base without amino acids (VWR), adenine 0.1 mg/ml, uracil 0.04 mg/ml, and other carbohydrate source as noted.

5.2.2 Plasmid Construction

The endoglucanase Cel7B gene was excised from the previously constructed pKLAC1.Cel7B plasmid as a BglII/NotI fragment, amplified by PCR using the specific primers described in Chapter 4. The PCR reaction mixture was separated by electrophoresis on low gelling temperature agarose (Amresco), and visualized under UV light with ethidium bromide staining. The Cel7B band was excised with a clean scalpel and the DNA was purified using the QIAquick Gel Extraction kit (Qiagen).

Plasmid pGBN19 and gene Cel7B were each digested with BglII and NotI restriction enzymes (New England Biolabs) to prepare for ligation. Five hundred

nanograms (500 ng) of insert DNA and 500 ng of plasmid DNA were digested

overnight for each ligation reaction. Digested DNA was then purified on low

gelling temperature agarose as above. Following purification a ligation reaction

containing 75 ng of vector DNA and 25 ng of insert DNA (a 2:1 molar excess of

insert:vector DNA) was incubated with T4 DNA ligase in a total reaction volume

of 10 µl at 16°C overnight. Following ligation 5 µl of the reaction mixture was

used to transform NEB Turbo competent *E. coli* (New England Biolabs) with the

plasmid pGBN19.Cel7B according to manufacturer's protocol. *E. coli* were

grown on LB agar plates containing 100 µg/ml ampicillin to select for positive

transformants. Colonies were selected and grown in suspension cultures in LB

media with ampicillin. Cells were pelleted by centrifugation and plasmid DNA

was purified using the Plasmid MiniKit from Qiagen.

5.2.3 Transformation of K. lactis

K. lactis strain 799 expressing *T. reesei* Cel6A (denoted hereafter as the parent

strain, see Chapter 4) was prepared for electroporation according to the method

described by Johnson et al.(Johnson et al. 2010) Briefly, cells were inoculated

into 50 ml of YPD media and allowed to grow for 2-4 doubling times. Cells were

harvested by centrifugation at 1500 g RCF at 4 °C. Cells were resuspended in

50 ml rich media with 200 mM HEPES buffer pH 8.0, 100 mM lithium acetate, 10

mM dithiothreitol, and allowed to incubate for 30 min. at 4 °C. They were then

centrifuged and resuspended in increasingly smaller volumes of ice-cold, sterile 1

M sorbitol; 50 ml, 25 ml, 2 ml, and 150 µl. Cells were then frozen at -80 °C until needed.

Plasmid pGBN19.Cel7B was linearized with SacII and separated on low gelling temperature agarose. The larger fragment (containing the gene expression cassette) was excised and purified using the QIAquick Gel Extraction kit as above.

Cells that were prepared for electroporation (75 µl) and plasmid DNA (3 µg in 15 µl sterile, distilled water) were mixed in a pre-chilled 2 mm gap-width electroporation cuvette (Eppendorf), and pulsed in a Bio-Rad MicroPulser on the Sc2 setting (5 ms, 1.5 kV, 25 µF, 200 Ω). Following electroporation, 1 ml ice-cold YPD media containing 1 M sorbitol was added to the electroporation cuvette and the cells were allowed to rest on ice for five minutes. Cells were then transferred to a sterile culture tube and incubated at 30 °C for 2 hours. They were then centrifuged at 5000 g RCF, the supernatant was removed, the cells were resuspended in 1 ml sterile distilled water, centrifuged again and resuspended in 250 µl sterile distilled water. Cells were then divided in half and plated either onto YPD media with 200 µg/ml G418, or onto doubly-selective media which was yeast carbon base with acetamide (New England Biolabs) plus 200 µg/ml G418 antibiotic.

5.2.4 Growth Rate Studies

To investigate the usefulness of growth rate screening for cellobiose utilization, several studies were performed to characterize *K. lactis* growth on various substrates. For growth rate studies, untransformed cell lines of strains 799 and 2288 were inoculated into minimal media containing galactose (0.5 wt%), cellobiose (4.5 wt%) or both galactose and cellobiose (0.5 wt% and 4.5 wt% respectively). Cell growth was followed spectrophotometrically by measuring light scattering at 600 nm over a period of three days. For rapid screening tests, growth of untransformed 799 and 2288 cells was observed on rich media in Petri dishes with various carbohydrate substrates added as noted.

5.2.5 Endoglucanase Plate Screening

For endoglucanase activity screening, cells were selected from doubly-selective media and plated onto rich media containing 2 wt% galactose, 2 wt% agar, 1 wt% CMC, and 50 mM citrate pH 4.8. Cells were allowed to grow for 2 days after which time they were gently rinsed off of the plate with distilled water. Plates were then flooded with 10 ml of 0.3 wt% Congo Red stain for half an hour. Excess Congo Red was poured off, and the plate was destained for 15-30 minutes with 1 M NaCl. Destaining was performed 3-5 times. Congo Red binds to CMC, but not to areas which have been hydrolyzed by endoglucanase activity, showing clearing zones in areas which have been affected by enzyme activity.

5.2.6 Cellulose Hydrolysis Assays

Exoglucanase activity was determined on Avicel PH-101 (Fluka) powdered cellulose as the substrate using HPLC detection of cellobiose for measurement. Avicel at 4 mg/ml was suspended in distilled water by magnetic stirring, and 250 µl was pipette into a 2 ml screw-top microcentrifuge tube. Citrate buffer (1 M, pH 4.5), tetracycline 10 mg/ml, cycloheximide 10 mg/ml, recombinant enzyme, and distilled water were added to give the following final reaction condition: 50 mM citrate pH 4.8, tetracycline 0.04 mg/ml, cycloheximide 0.03 mg/ml, Avicel 1 mg/ml, enzyme, and distilled water up to a final reaction volume of 1 ml. Microcentrifuge tubes were capped and taped flat to the surface of an incubated rotary shaker and incubated at 50°C and 250 RPM. To take samples, reaction tubes were centrifuged at maximum speed (16,000 g RCF) in a benchtop microcentrifuge for two minutes. One hundred microliter samples were pipeted from the surface, then the reaction vials were vortex mixed and returned to the incubated shaker table. Samples were taken at 0, 24, 48, and 72 hours to track the production of cellobiose over three days. Cellobiose was detected using an Agilent 1200 HPLC with refractive index detector following sugar separation on a Bio-Rad Aminex HPX-87P analytical column. The process of concentrating recombinant enzyme, where necessary, was performed by ultrafiltration in a VivaCell 70 protein concentrater (Sartorius Stedim) with a 10 kDa molecular weight cutoff membrane.

5.2.7 Determination of Enzyme Concentration

To determine the total amount of enzyme secreted by transformed cells,

enzymes were concentrated (as stated immediately above) and protein

concentration of the concentrated solution was determine by the Bradford

method (Bradford 1976). The concentration of enzyme in the original cell culture

supernatant was calculated from this using the starting and ending volumes to

determine the concentration factor. This approach was taken because Bradford

assays cannot be performed on the crude cell culture supernatant due to the low

level of enzyme being secreted, and the high background reading from to the

peptone in the media.

5.3 RESULTS AND DISCUSSION

To determine if cell growth on cellobiose could be differentiated from cell growth

on galactose, experiments were performed to measure cell growth rates on the

two different sugars separately, and in combination. Untransformed cell lines of

both strains 799 and 2288 were inoculated into minimal media with the two

sugars, as described in the methods section. Growth was measured

spectrophotometrically by light scattering measurements at 600 nm, and the

results are presented in Figure 5.2 below.
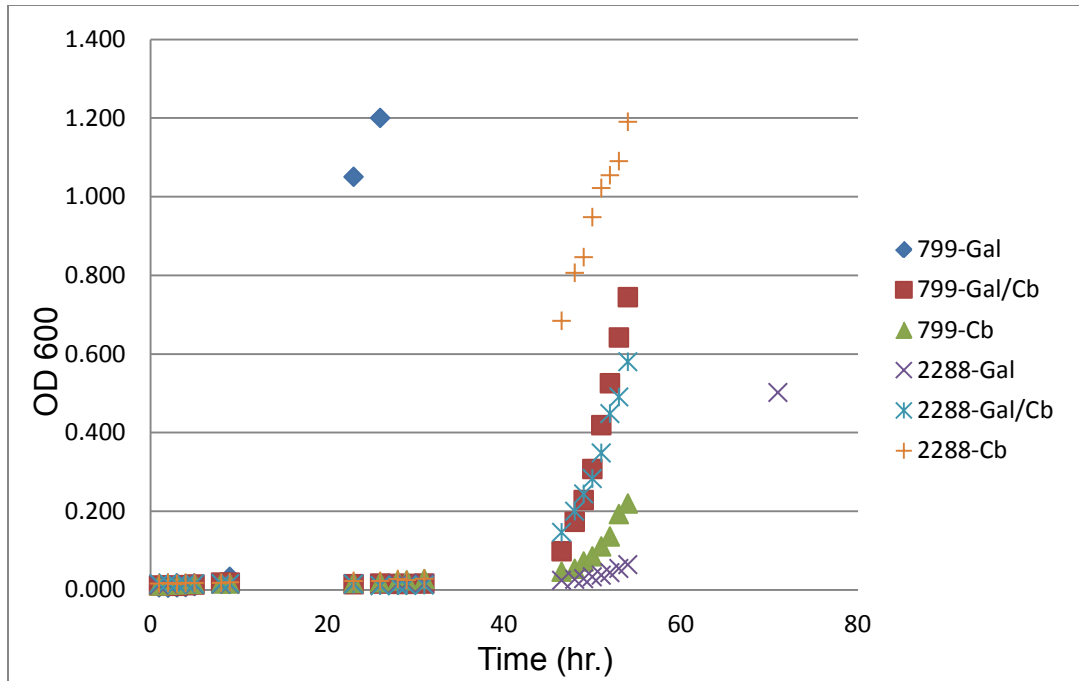
Figure 5.1: K. lactis growth rates in minimal media on different sugar sources. Sugar concentrations are 0.5% galactose (Gal), and 4.5% cellobiose, (Cb).

This plot shows that both strains are capable of growth on both substrates. Strain 799 grows most rapidly on galactose, moderately on galactose and cellobiose mixed, and slowest on cellobiose. Strain 2288 grows fastest on cellobiose, moderately on galactose and cellobiose mixed, and slowest on galactose. Doubling times, determined from the exponential growth period, for the different yeast strains and sugars were calculated from the optical density data and are presented in Table 5.1 below.

Table 5.1
Doubling times, in hours, of untransformed 799 and 2288 cells on minimal media with various carbohydrate sources.

	799	2288
Galactose	1.20	5.34
Galactose/Cellobiose	2.56	3.74
Cellobiose	3.13	3.45

Figure 5.2 and Table 5.1 clearly show that both strains are capable of metabolizing either sugar, although galactose utilization by strain 2288 is greatly reduced. Most importantly, the data show that differences in growth rates of strain 2288 on galactose and cellobiose can clearly be seen, and that growth on cellobiose in the presence of galactose is clearly differentiable from growth on galactose alone. This will be necessary to demonstrate yeast growth from cellulose hydrolysis.

Direct spectrophotometric measurement of cell growth in the presence of lignocellulosic biomass is complicated by light scattering due to the solid substrate particles. Additionally, for high-throughput screening methods, growth screening in Petri dishes is preferable for the rapid evaluation of large numbers of candidates.

Meanwhile, other work in our laboratory (Alshoug, unpublished) demonstrated that while *K. lactis* yeast are capable of growth on minimal media and using only ammonium sulfate as a nitrogen source, heterologous enzyme production is below detectable limits without supplementation of the media with a more complex nutrient source, such as peptone or casamino acids. Therefore, cell growth rate screening studies in Petri dishes were performed on rich media. These results are pictured below in Figure 5.3, and presented more quantitatively in Table 5.2, following.
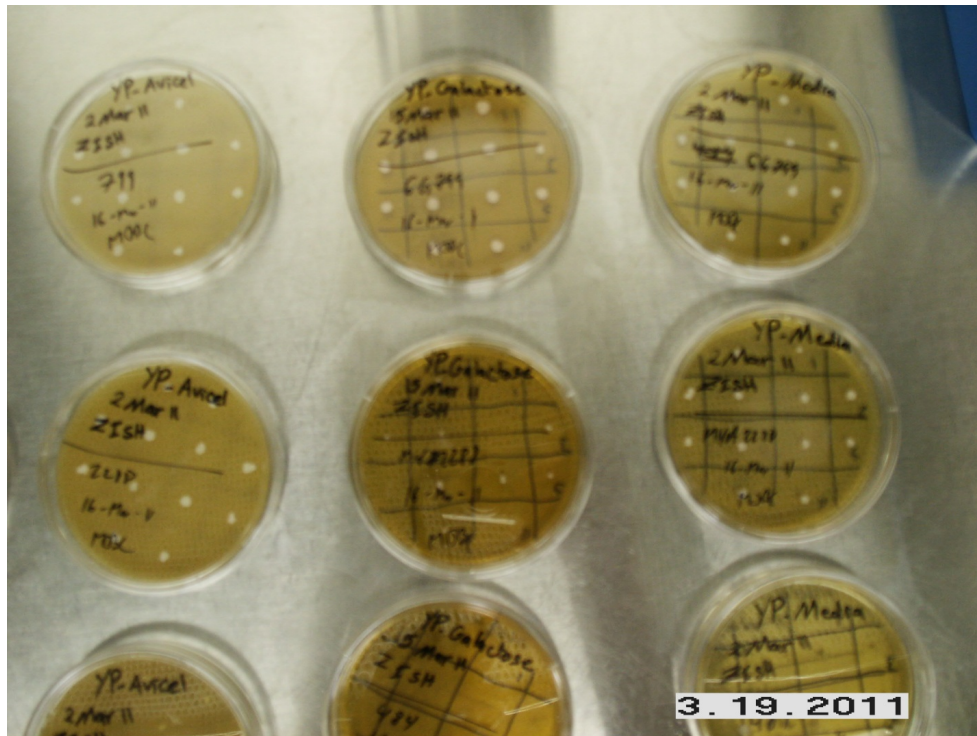
Figure 5.2: K. lactis growth in Petri dishes on media containing yeast extract and peptone (YP). Top row is strain 799, second row is strain 2288. Left column is YP with Avicel 2 wt%, middle column is YP with galactose 0.1 wt%, right column is YP with no carbohydrate source.

Table 5.2
Quantification of growth rates on agar media in petri dishes after five days.

	799	2288
YP	++	++
YP.Gal2	+++	+
YP.Avicel	++	++
YP.Avicel.Gal01	+++	+
YP.Avicel.Gal2	+++	+

YP is 1 w% yeast extract and 2 wt% peptone, Avicel is 2 wt%, Gal2 is 2 wt%, and Gal01 is 0.1 wt%. (+) indicates colonies <2 mm diameter, (++) indicates colonies 2-5 mm diameter, (+++) indicates colonies >5 mm diameter.

These results show that *K. lactis* is capable of growth on YP alone with no other carbohydrate source present. Strain 2288 is seen to be inhibited by the presence of galactose, and grows more slowly on plates containing galactose

than on others. Avicel was seen to settle to the bottom of the media before solidification, making no difference to the end results.

It was determined that a quantitative measurement of differential cell growth rates from cellulose hydrolysis in Petri dishes would not be feasible under these conditions. The research focus therefore returned to the use of strain 799 with HPLC detection of cellobiose production as the measure of cellulose hydrolysis rates.

The parent strain (a cell line of *K. lactis* strain 799 expressing *T. reesei* exoglucanase Cel6A on plasmid pKLAC1 and selected by acetamide utilization) was transformed with the gene for the endoglucanase Cel7B on plasmid pGBN19, and plated onto either singly-selective media (YPD with 200 µg/ml G418), or onto doubly-selective media (yeast carbon base with acetamide (NEB) plus 200 µg/ml G418). The number of colonies which grew on the singly-selective media was roughly equal to the number of colonies which grew on the doubly-selective media. This observation of the number of colonies indicates that the second transformation does little to interrupt gene expression from the first transformation, despite the fact that the two transformation events are targeted towards the same region of the yeast genome. The colonies which grew on the doubly-selective media were seen to be smaller than those which grew on the singly-selective media (YPD with 200 µg/ml G418), but this has been

observed to be true of yeast cell growth on yeast carbon base with acetamide in general.

Ten positive transformants were selected from the doubly-selective media and patched onto plates for endoglucanase activity assays using CMC as the substrate, as described in the methods section. The results are shown below in Figure 5.4.
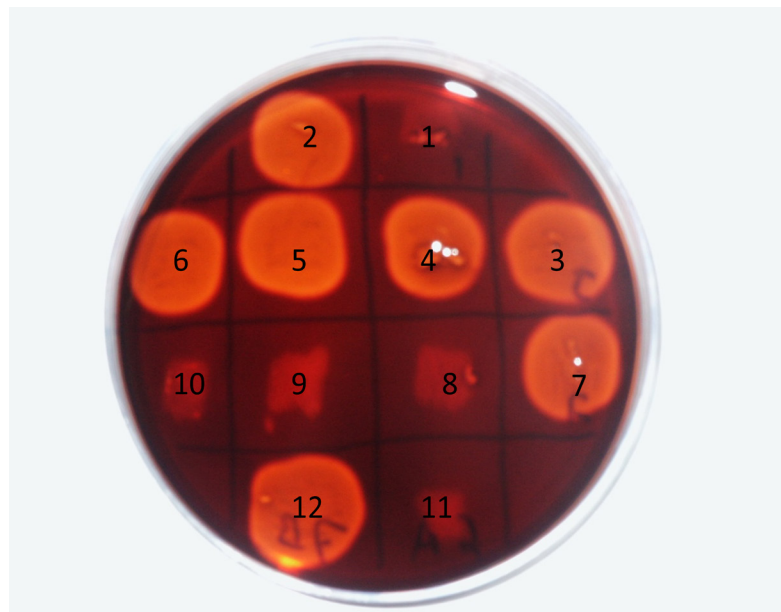
Figure 5.3: YPGal/CMC/citrate with Congo Red staining. Samples 1-10 are the product of two sequential rounds of transformation and selection, and grown on doubly-selective media. The parent strain expresses the exoglucanase Cel6A, and was selected by acetamide utilization. The daughters have been further transformed with the endoglucanase Cel7B, and selected by both acetamide utilization and G418 resistance. Sample 11 is the parent strain, showing no native endoglucanase activity. Sample 12 is a strain expressing only Cel7B from previous work.

CMC/Congo red staining is specific to endoglucanase activity, with clearing zones showing areas where endoglucanase enzymes have been active on the CMC substrate. The results from the plate screening clearly show that some of

the positive transformants which display antibiotic resistance also display endoglucanase activity (2-7), while some of the positive transformants which display antibiotic resistance do not display endoglucanase activity (1, 8-10). Cell lines 1-8 were selected for further study of cellulose hydrolysis rates from the enzymes expressed in the cell culture supernatant. The goal was to measure hydrolysis rates from the enzymes secreted into the cell culture supernatant, and to see if measurable differences could be observed for the parent strain, transformants displaying endoglucanase activity (2-7), and transformants not displaying endoglucanase activity (1,8).

Enzymes in the cell culture supernatant were concentrated by ultrafiltration for standard enzyme activity assays on Avicel powdered cellulose. The total protein concentration in solution was determined by the Bradford method (Bradford 1976). Cellulose hydrolysis assays were performed as described in the methods section with a standard enzyme loading of 0.1 mg/mL. The results are presented in Figure 5.5 below.
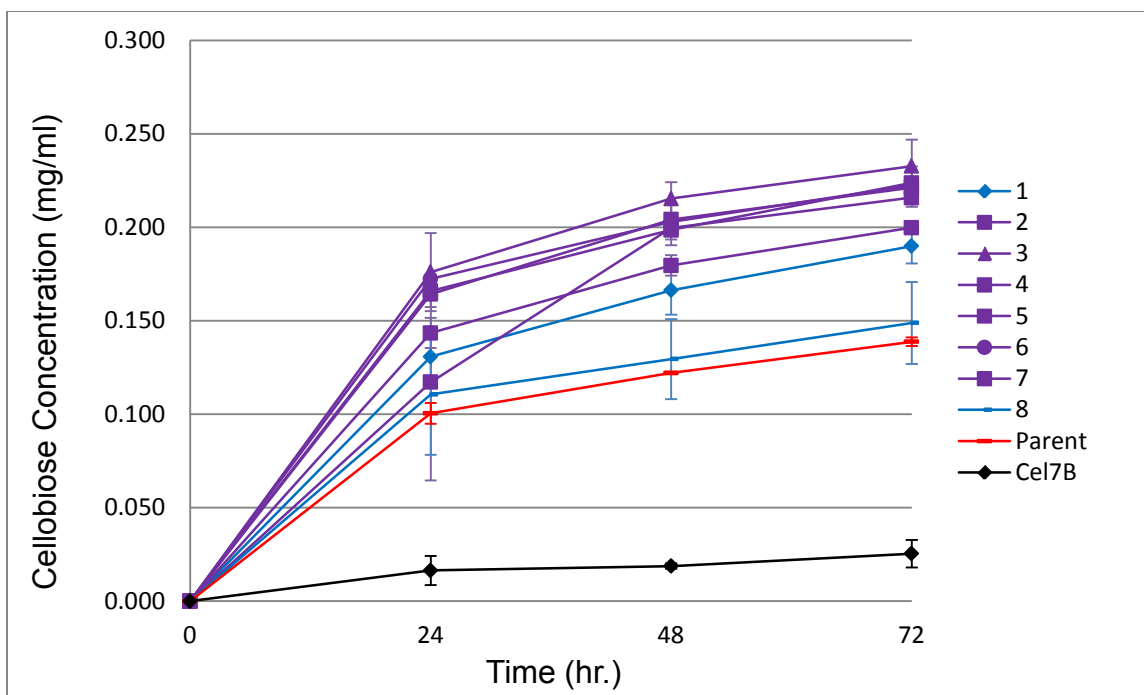
Figure 5.4: Cellobiose production of concentrated cell culture supernatants on Avicel powdered cellulose, averages and standard deviations of two experiments. Assay conditions are 1 mg/ml Avicel, 0.1 mg/ml enzyme, 50 mM citrate pH 4.8, 50°C.

These results show that double transformants displaying endoglucanase activity by the CMC/Congo Red plate screening assay (purple lines) can be clearly differentiated from the parent strain (red line), showing nearly twice the rate of cellobiose production as the parent strain (nearly twice the amount of cellobiose produced in the same period of time). Double transformants that do not show endoglucanase activity by the CMC/Congo Red plate screening assay (blue lines) still appear to perform slightly better than the parent strain, although worse than the double transformants that do display endoglucanase activity. These results are valuable because they demonstrate that the specific activity of the enzyme cocktail increases when multiple different enzymes are present, since the same enzyme loading of 0.1 mg/mL was used in each assay.

For faster throughput it is desirable to be able to perform assays on the crude cell culture supernatant, without enzyme concentration. To do this, cell culture supernatant was separated from cells by centrifugation, and then the supernatant was filtered through a 0.2 µm membrane filter in order to remove cells and cell debris, but still allow enzymes to flow through. Filtered cell culture supernatant (693 µl) was combined with the Avicel suspension, citrate buffer, tetracycline, and cycloheximide, as described in the methods section, with no distilled water added. With this method, the dilution factor of cellulases compared to the cell culture supernatent is 1/.693 = 1.44. The final reaction volume remains 1 ml. Results from these assays are presented in Figure 5.6 below.
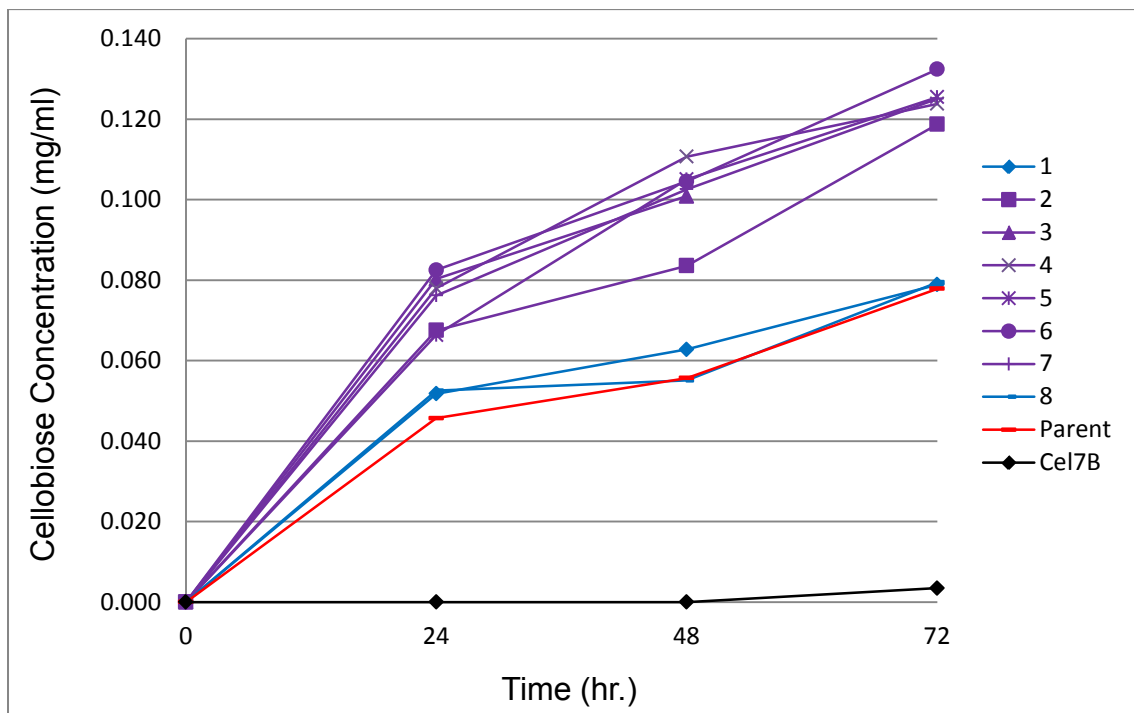
Figure 5.5: Cellobiose production of unconcentrated supernatants on Avicel

These results show an even clearer difference between double transformants displaying endoglucanase activity in one group, and the parent strain and double transformants not displaying endoglucanase activity in the other group. These results are one step closer to a truly high-throughput screening assay because they can be performed on slightly diluted crude cell culture supernatant without requiring a concentration step. This reduces the time, effort, and resources consumed for each assay, although the detection time is still limited by HPLC analysis, which limits us to approximately 20 samples per day.

One disadvantage of performing assays on the crude cell culture supernatant is that the results cannot directly show that the specific activity of the enzyme cocktail increases, because the total enzyme loading is not equal in each case. In order to determine total enzyme production and specific activity, enzyme concentrations in the cell culture were determined as described in the methods section. Results are given below in Table 5.3.

Table 5.3 shows that total enzyme production for the double transformants displaying endoglucanase activity (Samples 2-7) is greater than for either the parent strain or for the double transformants which do not display endoglucanase activity (Samples 1 and 8). This suggests that the double transformants which do not display endoglucanase activity are likely not producing any endoglucanase at all, as opposed to the other possibility that they are producing endoglucanase but that the endoglucanase is not active. These results also

indicate that some of the improved performance in the daughter cells is due

simply to a greater amount of enzyme being secreted, although the standard

activity assay with 0.1 mg/ml enzyme loading (Figure 5.5) shows that the specific

activity of the enzyme cocktail is also increased, due to having multiple different

enzymes produced.

Table 5.3
Enzyme concentration in the cell culture supernatant for sequential
transformants. Standard deviations are calculated based on duplicate samples.

Sample	Concentration (mg/ml)	Standard Deviation	%RSD
1	0.0190	0.0016	8.3%
2	0.0372	0.0011	2.9%
3	0.0271	0.0024	8.8%
4	0.0311	0.0028	9.1%
5	0.0261	0.0021	8.2%
6	0.0347	0.0019	5.6%
7	0.0378	0.0013	3.4%
8	0.0174	0.0010	5.5%
Parent	0.0126	0.0034	26.9%
Cel7B	0.0190	0.0004	2.3%

It can be determined from the values in Table 5.3, and the dilution factor (1.44),

that the enzyme loading in the crude cell culture supernatant activity assay is

approximately 3-5 times lower than the 0.1 mg/mL enzyme loading for the

standard activity assay. However, an observation of Figures 5.3.4 and 5.3.5

shows that cellobiose production only decreases by a factor of about two. This

suggests that enzyme:substrate interactions may be approaching saturation at

the standard enzyme loading level of 0.1 mg/ml. In order to investigate this

phenomenon, concentrated cell culture supernatant containing exoglucanase

Cel6A from the parent strain was blended with endoglucanase Cel7B in a ratio of

9:1 Exo:Endo, and serially diluted it to levels seen in the unconcentrated cell

culture supernatant for the double transformant. This ratio of Exo:Endo was

chosen because in the native *T. reesei* cellulolytic system, approximately 10% of

the secreted cellulase is endoglucanase Cel7B and approximately 80% of the

secreted cellulase is a combination of the two exoglucanases Cel6A and Cel7A,

with the remaining 10% of secreted cellulase being made by minor contributions

of other cellulolytic proteins (M. Sandgren 2003). These assays were performed

down to dilutions of one-twentieth of the standard assay, and the results are

shown below in Figure 5.7.
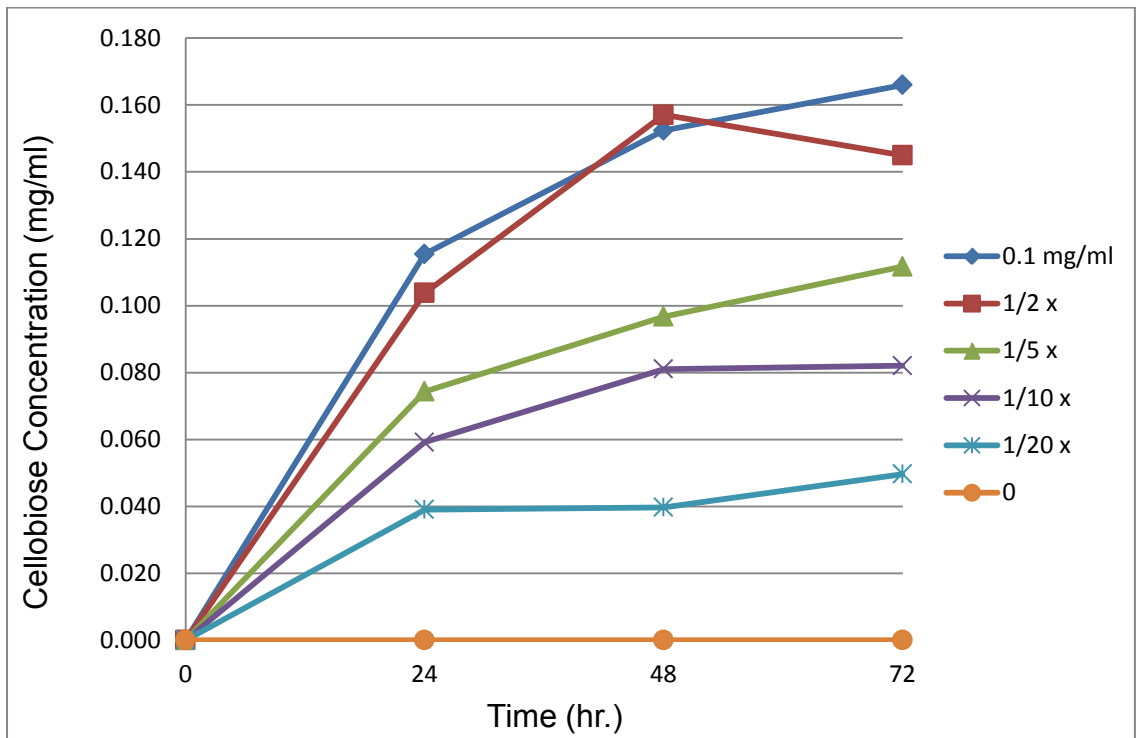
Figure 5.6:   Enzyme   activity   assays   with   increasingly   diluted   enzyme
concentration.   Enzyme   system   is   a   synthetic   blend   of   nine   parts   of   the
exoglucanase   Cel6A   (from   the   parent   strain)   mixed   with   one   part   of   the

endoglucanase Cel7B (from previous work), and loaded at the enzyme concentrations listed.

These data show that enzyme activity is detectable down to at least one-twentieth of the concentration of the standard enzyme activity assay. In order to determine enzyme specific activity more precisely, a quantative measure of the specific activity during the first 24 hours was defined by Equation 5.1 below

Eq. 5.1 $$Specific\ Activity = \frac{Cellobiose\ Amount\ (mg)}{Enzyme\ Amount\ (mg)*24(hr)}$$

Specific activities were plotted on a graph, shown below in Figure 5.8
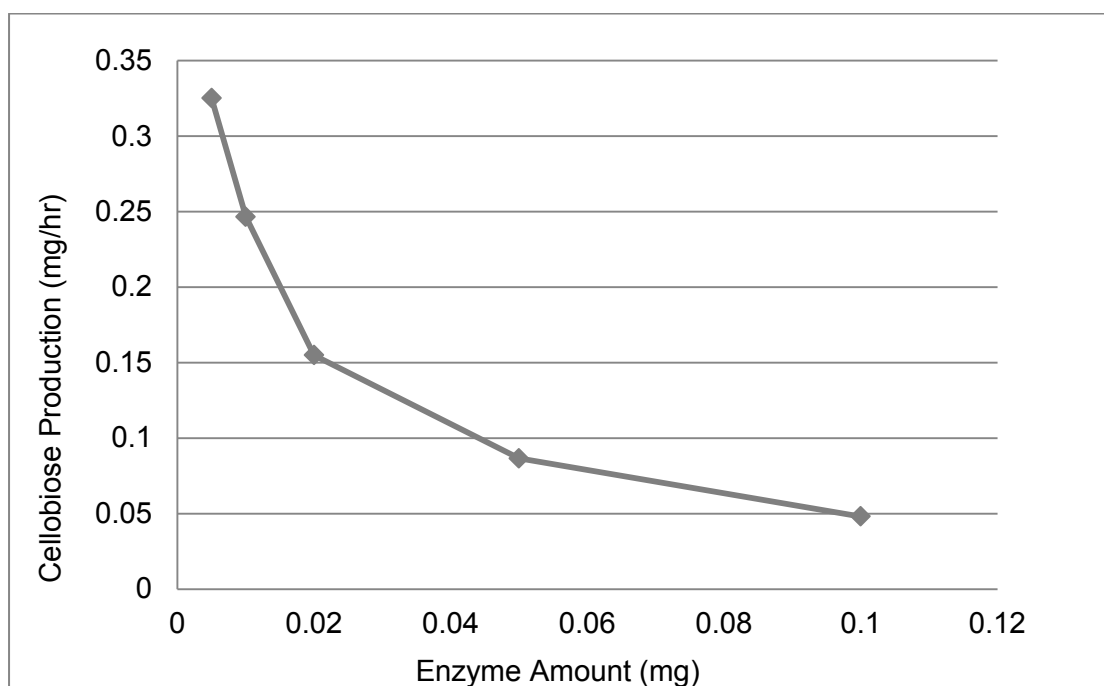
Figure 5.7: Specific activity of blended enzyme solution at varying concentrations.

Although activity of multiple enzymes on a solid substrate is a complex

phenomenon, it can clearly be seen that the enzyme loading of 0.1 mg/ml in the

standard assay is approaching saturation, with diminishing returns for increasing

enzyme loading. Part of the reason may be that all reactive sites on the cellulose

are occupied by enzyme, and further enzyme additions have no sites to act upon.

Another contribution to the diminishing returns may be the lower cellobiose

production at lower enzyme concentration, which would reduce the product

inhibition on the enzyme and therefore result in increased enzyme activity at

lower concentrations.

To compare the specific activity of the double transformants to the blended

enzyme system (Figure 5.7), specific activities were calculated according to

Equation 5.3.1, and are given here in Table 5.4.

Table 5.4
Enzyme concentration in the assays, and specific activity, of double
transformants at differing enzyme loading levels

	Enzyme conc. in assay mg/ml	Specific Activity at Low Concentration mg CB/(mg Enz*hr)	Specific Activity at High Concentration mg CB/(mg Enz*hr)
2	0.0257	0.1093	0.0598
3	0.0188	0.1784	0.0734
4	0.0216	0.1506	0.0643
5	0.0181	0.1527	0.0692
6	0.0240	0.1430	0.0719
7	0.0262	0.1213	0.0685
Average	0.0224	0.1425	0.0678
St. Dev.	0.0035	0.0245	0.0050
%RSD	15.5%	17.2%	7.40%

The assay from Figure 5.7 at 1/5x concentration (0.02 mg/ml) is about the same as in the assay of the unconcentrated supernatant (0.0224), and the specific activity for that point is approximately 0.15 as seen on the graph in Figure 5.8., comparable to the average for the assays of unconcentrated cell culture supernatant (0.1425). These data suggest that the enzyme system secreted by the double transformants is similar in composition to the synthetically blended enzyme system at a ratio of 9:1 exo:endo.

5.4 CONCLUSIONS AND RECOMMENDATIONS

In this work it has been shown that *K. lactis* is capable of simultaneously expressing multiple *T. reesei* cellulases, and that the total and specific activity of the secreted enzymes is increased both through greater enzyme production, and synergies of enzyme activity. It was demonstrated that it is possible to determine enzyme activity in the cell culture supernatant without the need for enzyme purification and concentration. Two different selection methods have been employed in order to first create a basic cellulolytic enzyme system in *K. lactis*, and then to improve its performance through the insertion of an additional gene into the yeast genome. Improvements in cellulose hydrolysis rates are detectable in the daughter cells following the second transformation, and directly attributable to the specific genetic modification. Finally, successful screening for endoglucanase activity in Petri dishes has also been demonstrated, providing an opportunity to use this high-throughput screening technique to increase

screening efficiency on applicable enzyme systems, which include an endoglucanase.

Measureable differences in growth rates on galactose and cellobiose, individually and in combination, have been demonstrated for both galactose-utilizing (799) and galactose-incompetent (2288) strains of *K. lactis*. However, high-throughput growth rate screening of colonies in Petri dishes was found to be impractical under the conditions of the petri dish assay. The growth rate screening technique using the galactose-incompetent strain 2288 could be revisited, employing a green fluorescent protein (GFP) tag on the cells for easier spectrophotometric quantification of cell growth. This would allow for cell growth in suspension culture while reducing the influence of light scattering due to interference of the biomass particles. An improvement to any high throughput petri dish assay would need to address the separation of the Avicel particles from the growing colony, which was caused by gravitational settling during the curing of the agar. This could possibly be accomplished by spreading a thin layer of Avicel in soft agar on the top of the so that a thin layer of Avicel sits on top of the agar and in contact with the cells.

Future work should focus on evaluating hydrolysis rates of actual pretreated lignocellulosic biomass instead of Avicel powdered cellulose, for more relevant quantification of enzyme system performance on natural substrates. The number and types of carbohydrate active enzymes evaluated should be expanded to

include hemicellulases, ligninases, and swollenins. This would provide a large, important, but still limited set of enzymes to evaluate, giving an opportunity to refine the techniques while still permitting HPLC detection.

Finally, if growth rate screening or another high-throughput sugar detection technique can be devised, the sequential transformation and selection technique provides a valuable tool which could be employed in random mutagenesis and directed evolution studies to evaluate a large number of mutations in a single gene to determine individual contributions to enzyme system performance. Whole genome screening would also be possible to look for other ancillary proteins, such as those which might increase protein production, improve inhibitor tolerance, or otherwise enhance overall enzyme system performance.

References

110th Congress. 2007. Energy Independence and Security Act of 2007. United States Government Printing Office.

Aden A, Ruth M, Ibsen K, Jechura J, Neeves K, Sheehan J, Wallace B, Montague L, Slayton A, Lukas J. 2002. Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. Golden: National Renewable Energy Laboratory.

Adler PR, Del Grosso SJ, Pargon WJ. 2007. Life-Cycle Assessment of Net Greenhouse-Gas Flux for Bioenergy Cropping Systems. Ecological Applications 17(3):675-691.

Allen SA, Clark W, McCaffery JM, Cai Z, Lanctot A, Slininger PJ, Liu ZL, Gorsich SW. 2010. Furfural induces reactive oxygen species accumulation and cellular damage in Saccharomyces cerevisiae. Biotechnol. Biofuels 3:2.

Anderson PJ, McNeil K, Watson K. 1986. High-Efficiency Carbohydrate Fermentation to Ethanol at Temperatures above 40°C by *Kluyveromyces marxianus* var. *marxianus* Isolated from Sugar Mills. Applied and Environmental Microbiology 51(6):1314-1320.

Bare J. 2012. Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts (TRACI). Environmental Protection Agency.

Bhikhabhai R, Johansson G, Petterson G. 1984. Isolation of cellulolytic enzymes from *Trichoderma reesei* QM9414. Journal of Applied Biochemistry 6(5-6):336-345.

Boer H, Teeri TT, Koivula A. 2000. Characterization of *Trichoderma reesei* Cellobiohydrolase Cel7A Secreted from *Pichia pastoris* Using Two Different Promoters. Biotechnology and Bioengineering 69(5):486-494.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry 72(1-2):248-254.

Brander M, Tipper R, Hutchinson C, Davis G. 2009. Consequential and Attributional Approaches to LCA: a Guide to Policy Makers with Specific Reference to Greenhouse Gas LCA of Biofuels. Ecometrica Press.

Brodeur-Campbell MJ, Shonnard DR. in prep. Expression and Characterization of Three *Trichoderma reesei* Cellulose Hydrolases in *Kluyveromyces lactis* Yeast.

Chazdon RL, Colwell RK, Denslow JS, Guariguata MR. 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of Northeastern Costa Rica. Forest biodiversity research, monitoring and modeling: conceptual background and old world case studies. p. 285-309.

Chen MC, Gritzali M, Stafford DW. 1987. Nucleotide Sequence and Deduced Primary Structure of Cellobiohydrolase II from *Trichoderma reesei*. Bio/Technology 5:274-278.

Cherubini F, Jungmeier G. 2010. LCA of a biorefinery concept producing bioethanol, bioenery, and chemicals from switchgrass. International Journal of Life Cycle Assessment 15:53-66.

Chundawat SPS, Balan V, Dale BE. 2008. High-Throughput Microplate Technique for Enzymatic Hydrolysis of Lignocellulosic Biomass. Biotechnology and Bioengineering 99(6):1281-1294.

Chung YC, Bakalinsky A, Penner MH. 2005. Enzymatic saccharification and fermentation of xylose-optimized dilute acid-treated lignocellulosics. Appl. Biochem. Biotechnol. 121-124:947-961.

Cummings C, Fowler T. 1996. Secretion of *Trichoderma reesei* β-glucosidase by *Saccharomyces cerevisiae*. Current Genetics 29:227-233.

Curran MA. 1997. Life-Cycle Based Government Policies. International Journal of Life Cycle Assessment 2(1):39-43.

Dien BS, Jung H-JG, Vogel KP, Casler MD, Lamb JFS, Iten L, Mitchell RB, Sarath G. 2006. Chemical composition and response to dilute-acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. Biomass Bioenergy 30:880-891.

Greenhouse Gas Emissions: Transportation Sector Emissions [Internet]. 2010a. [updated 14-Jun. 2012, cited. Available from: http://epa.gov/climatechange/ghgemissions/sources/transportation.html

EPA. 2010b. Regulation of Fuels and Fuel Additives: Changes to Renewable Fuel Standard Program; Final Rule. Federal Register. p. 14669-15320.

Fargione JE, Hill J, Tilman D, Polasky S, Hawthorne P. 2008. Land Clearing and the Biofuel Carbon Debt. Sciencexpress.

Fargione JE, Cooper TR, Flaspohler DJ, Hill J, Lehman C, McCoy T, McLeod S, Nelson EJ, Oberhauser KS, Tilman D. 2009. Bioenergy and Wildlife:

Threats and Opportunities for Grassland Conservation. BioScience 59(9):767-777.

Forbes RD, Society of American Foresters. 1961. Forestry handbook. New York: Ronald Press Co.

Gasol CM, Gabarrell X, Anton A, Rigola M, Carrasco J, Ciria P, Rieradevall J. 2009. LCA of poplar bioenergy system compared with *Brassica carinata* energy crop and natural gas in regional scenario. Biomass and Bioenergy 33:119-129.

Godbole S, Decker SR, Nieves RA, Adney WS, Vinzant TB, Baker JO, Thomas SR, Himmel ME. 1999. Cloning and Expression of *Trichoderma reesei* Cellobiohydrolase I in *Pichia pastoris*. Biotechnology Progress 15:828-833.

Grohmann K, Torget R, Himmel M. 1985. Optimization of dilute acid pretreatment of biomass. Biotechnology and Bioengineering Symposium 15:59-80.

Grubb PJ. 1997. The Maintenance of Species-Richness in Plant Communities: The Importance of the Regeneration Niche. Biolgical Reviews 52(1):107-145.

Guinée JB, Heijungs R, Huppes G. 2010. Life cycle assessment: past, present and future. Environmental Science & Technology.

Guo G-L, Chen W-H, Chen W-H, Men L-C, Hwang W-S. 2008. Characterization of dilute acid pretreatment of silvergrass for ethanol production. Bioresour. Technol. 99:6046-6053.

Heinzelman P, Snow CD, Wu I, Nguyen C, Villalobos A, Govindarajan S, Minshull J, Arnold FH. 2009. A family of thermostable fungal cellulases created by structure-guided recombination. Proceedings of the National Academy of Science 106(14):5610-5615.

Heller MC, Keoleian GA, Volk TA. 2003. Life cycle assessment of a willow bioenergy cropping system. Biomass and Bioenergy 25:147-165.

Hsu DD, Inman D, Heath GA, Wolfrum EJ, Mann MK, Aden A. 2010. Life Cycle Environmental Impacts of Selected U.S. Ethanol Production and Use Pathways in 2022. Environmental Science & Technology 44(13):5289-5297.

Humbird D, Davis R, Tao L, Kinchin C, Hsu D, Aden A. 2011. Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol. Golden CO: National Renewable Energy Laboratory.

IPCC. 2000. Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories.

IPCC. 2006. 2006 Guidelines for National Greenhouse Gas Inventories, Prepared by the National

Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). IGES, Japan.

IPCC. 2007a. Summary for Policymakers. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, M.Tignor, Miller HL, editors. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

IPCC. 2007b. Transport and its infrastructure. In: Kahn Ribeiro S, Kobayashi S, Beuthe M, Gasca J, Greene D, Lee DS, Muromachi Y, Newton PJ, Plotkin S, Sperling D et al. , editors. Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Jensen J, Morinelly J, Aglan A, Mix A, Shonnard DR. 2008. Kinetic Characterization of Biomass Dilute Sulfuric Acid Hydrolysis: Mixtures of Hardwoods, Softwood, and Switchgrass. AIChE J. 54(6):1637-1645.

Jensen JR, Morinelly JE, Gossen KR, Brodeur-Campbell MJ, Shonnard DR. 2010. Effects of dilute acid pretreatment conditions on enzymatic hydrolysis monomer and oligomer sugar yields for aspen, balsam, and switchgrass. Bioresour. Technol. 101:2317-2325.

Johnson SC, Yang M, Murthy PN. 2010. Heterologous expression and functinal characterization of a plant alkaline phytase in *Pichia pastoris*. Protein Expression and Purification 74:196-203.

Kabel MA, Bos G, Zeevalking J, Voragen AGJ, Schols HA. 2007. Effect of pretreatment severity on xylan solubility and enzymatic breakdown of the remaining cellulose from wheat straw. Bioresour. Technol. 98(2034-2042):2034.

Karlsson J, Saloheimo M, Siika-aho M, Tenkanen M, Penttilä M, Tjerneld F. 2001. Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. European Journal of Biochemistry 268:6498-6507.

Kemppainen AJ. 2003. Comparative Life Cycle Assessments for Biomass to Ethanol Production from Different Regional Feedstocks. [Houghton]: Michigan Technological University. p. 198.

King BC, Donnelly MK, Bergstrom GC, Walker LP, Gibson DM. 2009. An Optimized Microplate Assay System for Quantitative Evaluation of Plant Cell Wall-Degrading Enzyme Activity of Fungal Culture Extracts. Biotechnology and Bioengineering 102(4):1033-1044.

Kline K, Dale VH, Lee R, Leiby P. 2009. In Defense of Biofuels, Done Right. Issues in Science and Technology 76:75-84.

Laymon RA, Adney WS, Mohagheghi A, Himmel ME, Thomas SR. 1996. Cloning and Expression of Full-Length *Trichoderma reesei* Cellobiohydrolase I cDNAs in *Escherichia coli*. Applied Biochemistry and Biotechnology 57/58:389-397.

Liska AJ, Yang HS, Bremer VR, Klopfenstein TJ, Walters DT, Erickson GE, DCassman KG. 2009. Improvements in Life Cycle Energy Efficiency and Greenhouse Gas Emissions of Corn-Ethanol. Journal of Industrial Ecology 13(1):58-74.

Liu G, Tang X, Tian S-L, Deng X, Xing M. 2006. Improvement of the cellulolytic activity of Trichoderma reesei endoglucanase IV with an additional catalytic domain. World Journal of Microbiology and Biotechnology 22:1301-1305.

Lloyd TA, Wyman CE. 2005. Combined sugar yields for dilute sulfuric acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. Bioresour. Technol. 96:1967-1977.

Lynd L, Grethlein H. 1987. Hydrolysis of dilute acid pretreated mixed hardwood and purified microcrystalline cellulose by cell-free broth from Clostridium thermocellum. Biotechnol. Bioeng. 29(1):92-100.

Lynd LR, Weimer PJ, van Zyle WH, Pretorius IS. 2002. Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiology and Molecular Biology Reviews 66(3):506-577.

Lynd LR, van Zyle WH, McBride JE, Laser M. 2005. Consolidated bioprocessing of cellulosic biomass: an update. Current Opinion in Biotechnology 16:577-583.

MacLean HL, Spatari S. 2009. The contribution of enzymes and process chemicals to the life cycle of ethanol. Environmental Research Letters 4:1-10.

Mandels M, Weber J. 1969. The Production of Cellulases. Advances in Chemistry 95:391-414.

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D et al. . 2008. Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina). Nature Biotechnology 26:553-560.

Martínez JM, Reguant J, Montero MA, Montané D, Salvadó J, Farriol X. 1997. Hydrolytic pretreatment of softwood and almond shells. Degree of polymerization and enzymatic digestibility of the cellulose fraction. Ind. Eng. Chem. Res. 36:688-696.

McBride AC, Dale VH, Baskaran LM, Downing ME, Eaton LM, Efroymson RA, Garten CT, Jr., Kline KL, Jager HI, Mulholland PJ et al. . 2011. Indicators to support environmental sustainability of bioenergy systems. Ecological Indicators 11:1277-1289.

McMillan JD. 1994. Pretreatment of Lignocellulosic Biomass. In: Himmel ME, Baker JO, Overend RP, editors. Enzymatic Conversion of Biomass for Fuels Production. American Chemical Society. p. 292–324.

Meyer J, Walker-Jonah A, Hollenberg C. 1991. Galactokinase encoded by GAL1 is a bifunctional protein required for induction of the GAL genes in Kluyveromyces lactis and is able to suppress the gal3 phenotype in Saccharomyces cerevisiae. Molecular Cell Biology 11:5454-5461.

Mitchell D. 2008. A Note on Rising Food Prices. In: Working Paper No. 4682, editor. World Bank Policy Research Working Paper Series. World Bank.

Morinelly JE, Jensen JR, Browne M, Co TB, Shonnard DR. 2009. Kinetic Characterization of Xylose Monomer and Oligomer Concentrations during Dilute Acid Pretreatment of Lignocellulosic Biomass from Forests and Switchgrass. Ind. Eng. Chem. Res. 48(22):9877-9884.

Mosier NS, Hall P, Ladisch CM, Ladisch MR. 1999. Reaction Kinetics, Molecular Action, and Mechanisms of Cellulolytic Proteins. Advances in Biochemical Engineering / Biotechnology 65:23-40.

Mu D, Seager T, Rao PS, Zhao F. 2010. Comparative Life Cycle Assessment of Lignocellulosic Ethanol Production: Biochemical Versus Thermochemical Conversion. Environmental Management 46(4):565-578.

Müller S, Sandal T, Kamp-Hansen P, Dalbøge H. 1998. Comparison of Expression Systems in the Yeasts Saccharomyces cerevisiae, Hansenula

polymorpha, Klyveromyces lactis, Schizosaccharomyces pombe and Yarrowia lipolytica. Cloning of Two Novel Promoters from Yarrowia lipolytica. Yeast 14:1267-1283.

Nonklang S, Abdel-Banat B, Cha-aim K, Moonjai N, Hoshida H, Limtong S, Yamada M, Akada R. 2008. High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. Applied and Environmental Microbiology 74(24):7514-7521.

Biomass Program: Feedstock Composition and Property Database [Internet]. 2004. U.S Department of Energy. [updated 14-May 2004, cited. Available from: http://www.afdc.energy.gov/biomass/progs/search1.cgi

Okada H, Sekiya T, Yokoyama K, Tohda H, Kumagai H, Morikawa Y. 1998a. Efficient secretion of *Trichoderma reesei* cellobiohydrolase II in *Schizosaccharomyces pombe* and characterization of its products. Applied Microbiology and Biotechnology 49:301-308.

Okada H, Tada K, Sekiya T, Yokoyama K, Takahashi A, Tohda H, Kumagai H, Morikawa Y. 1998b. Molecular Characterization and Heterologous Expression of the Gene Encoding a Low-Molecular-Mass Endoglucanase from Trichoderma reesei QM9414. Applied and Environmental Microbiology 64(2):555-563.

Oliva JMS, Felicia; Ballesteros, Ignacio; González, Alberto; Negro, Maria José; Manzanares, Paloma; Ballesteros, Mercedes. 2003. Effect of Lignocellulosic Degradation Compounds from Steam Explosion Pretreatment on Ethanol Fermentation by Thermotolerant Yeast Kluyveromyces marxianus. Applied Biochemistry and Biotechnology 105-108:141-153.

Penttila M, Lehtovaara P, Nevalainen H, Bhikhabhai R, Knowles J. 1986. Homology between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. Gene 45:253-263.

Penttila M, Nevalainen H, Ratto M, Salminen E, Knowles J. 1987. A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61:155-164.

Penttilä ME, André L, Lehtovaara P, Bailey M, Teeri TT, Knowles JKC. 1988. Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. Gene 63:103-112.

Read JD, Colussi PA, Ganatra MB, Taron C. 2007. Acetamide Selection of *Kluyveromyces lactis* Cells Transformed with an Integrative Vector Leads

to High-Frequency Formation of Multicopy Strains. Applied and Environmental Microbiology 73(16):5088-5096.

Reis L. 2009. Inputs in a Logging Operation, for a SimaPro Model. Houghton: Michigan Technological University.

Rodrussamee N, Lertwattanasakul N, Hirata K, Suprayogi, Limtong S, Kosaka T, Yamada M. 2011. Growth and ethanol fermentation ability on hexose and pentose sugars and glucose effect under various conditions in thermotolerant yeast *Kluyveromyces marxianus*. Applied Microbiology and Biotechnology 90:1573-1586.

Romanos MA, Scorer CA, Clare JJ. 1992. Foreign gene expression in yeast: a review. Yeast 8(6):423-488.

Saloheimo M, Nakari-Setala T, Tenkanen M, Penttilä M. 1997. cDNA cloning of a *Trichoderma reesei* cellulase and demonstration of endoglucanase activity by expression in yeast. European Journal of Biochemistry 249:584-591.

Sampson R. 2007. Switchgrass Production in Ontario: A Management Guide. Resource Efficient Agricultural Production (REAP) - Canada.

Samuel R, Pu Y, Raman B, Ragauskas AJ. 2010. Structural Characterization and Comparison of Switchgrass Ball-milled Lignin Before and After Dilute Acid Pretreatment. Applied Biochemistry and Biotechnology 162(1):62-74.

Sandgren M. 2003. Structural and functional studies of glycoside hydrolase family 12 enzymes from *Trichoderma reesei* and other cellulolytic microorganisms [Dissertation]. [Uppsala]: Uppsala University. p. 68.

Sandgren M, Ståhlberg J, Mitchinson C. 2005. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. Progress in Biophysics and Molecular Biology 89:246-291.

Sannigrahi P, Ragauskas AJ, Tuskan GA. 2010. Poplar as a feedstock for biofuels: A review of compositional characteristics. Biofuels, Bioproducts, and Biorefining 4:209-226.

Serapiglia MJ, Cameron KD, Stipanovic AJ, Smart LB. 2009. Analysis of Biomass Composition Using High-Resolution Thermogravimetric Analysis and Percent Bark Content for the Selection of Shrub Willow Bioenergy Crop Varieties. BioEnergy Research 2(1-2):1-9.

Sewalt VJH, Glasser WG, Beauchemin KA. 1997. Lignin Impact on Fiber Degradation. 3. Reversal of Inhibition of Enzymatic Hydrolysis by

Chemical Modification of Lignin and by Additives. J. Agric. Food Chem. 45:1823-1828.

Sheehan JJ. 2009. Biofuels and the conundrum of sustainability. Current Opinion in Biotechnology 20:318-324.

Shoemaker S, Schweickart V, Ladner M, Gelfand D, Kwok S, Myambo K, Innis M. 1983. Molecular Cloning of Exo-Cellobiohydrolase I Derived from *Trichoderma reesei* Strain L27. Bio/Technology 1:691-696.

Simpson EH. 1949. Measurements of diversity. Nature 163:688.

Skrigan AI, Murashkevich TV, Pershina GN, Khorevskaya NM, Pimenova SV. 1967. The chemical composition of the wood of some species of willow. Chemistry of Natural Compounds 3(6):328-330.

Spartari S, Zhang Y, Maclean HL. 2005. Life Cycle Assessment of Switchgrass- and Corn Stover-Derived Ethanol-Fueled Automobiles. Environmental Science & Technology 39(9750-9758):9750.

Swinkels BW, van Ooyen AJJ, Bonekamp FJ. 1993. The yeast *Kluyveromyces lactis* as an efficient host for heterologous gene expression. Antonie van Leeuwenhoek 64:187-201.

Takashima S, Iikura H, Nakamura A, Hidaka M, Masaki H, Uozumi T. 1998. Overproduction of recombinant *Trichoderma reesei* cellulases by *Aspergillus oryzae* and their enzymatic properties. Journal of Biotechnology 65:163-171.

Takashima S, Nakamura A, Hidaka M, Masaki H, Uozumi T. 1999. Molecular Cloning and Expression of the Novel Fungal β-Glucosidase Genes from *Humicola grisea* and *Trichoderma reesei*. Journal of Biochemistry 125:728-736.

Tengborg C, Stenberg K, Galbe M, Zacchi G, Larsson S, Palmqvist E, Hahn-Hägerdal B. 1998. Comparison of SO2 and H2SO4 impregnation of softwood prior to steam pretreatment on ethanol production. Appl. Biochem. Biotechnol. 70-72:3-15.

Tilman D. 1996. Biodiversity: Population versus ecosystem stability. Ecology 77:350-363.

Tilman D, Hill J, Lehman C. 2006. Carbon-Negative Biofuels from Low-Input High-Diversity Grassland Biomass. Science 314:1598-1600.

U.S. DOE. 2011. U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry. Oak Ridge, TN.

U.S. DOE. 2006. Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda. U.S. Department of Energy Office of Science and Office of Energy Efficiency and Renewable Energy

Van Arsdell J, Kwok S, Schweickart VL, Ladner MB, Gelfand DH, Innis MA. 1987. Cloning, Characterization, and Expression in *Saccharomyces cerevisiae* of Endoglucanase I from *Trichoderma reesei*. Bio/Technology 5:60-64.

van Ooyen AJJ, Dekker P, Huang M, Olsthoorn MMA, Jacobs DI, Colussi PA, Taron CH. 2006. Heterologous protein production in the yeast *Kluyveromyces lactis*. FEMS Yeast Research 6:381-392.

Voutilainen SP, Boer H, Linder MB, Puranen T, Rouvinen J, Vehmaanperä J, Koivula A. 2007. Heterologous expression of Melanocarpus albomyces cellobiohydrolase Cel7B, and random mutagenesis to improve its thermostability. Enzyme and Microbial Technology 41:234-243.

Walton J, Banerjee G, Car S. 2011. GENPLAT: and Automated Platform for Biomass Enzyme Discovery and Cocktail Optimization. Journal of Visualized Experiments 56:1-5.

Wang M, Wu M, Huo H. 2007. Life-cycle energy and greenhoues gas emission impacts of different corn ethanol plant types. Environmental Research Letters 2:1-13.

Webster CR, Flaspohler DJ, Jackson RD, Meehan TD, Gratton C. 2010. Diversity, productivity and landscape-level effects in North American grasslands managed for biomass production. Biofuels 1(3):451-461.

Webster CR, Flaspohler D, Pawson S, Brockerhoff E. 2012. Ecologically Sustainable Bioenergy Cropping Systems: Species Selection and Habitat Considerstions. In: Kole C, Joshi CP, Shonnard DR, editors. Handbook of Bioenergy Crop Plants. Boca Raton: CRC Press. p. 99-117.

Wiselogel A, Tyson S, Johnson D. 1996. Biomass feedstock resources and composition. In: Wyman CE, editor. Handbook of Bioethanol: Production and Utilization. Washington DC: Taylor & Francis.

Wood TM, Bhat KM. 1988. Methods for Measuring Cellulase Activities. Methods in Enzymology 160:87-112.

Wright MM, Brown RC, Boateng AA. 2008. Distributed processing of biomass to bio-oil for subsequent production of Fischer-Tropsch liquids. Biofuels, Bioproducts, and Biorefining 2:229-238.

Wyman CE, Dale BE, Elander RT, Holtzapple M, Ladisch MR, Lee YY. 2005. Coordinated development of leading biomass pretreatment technologies. Bioresource Technology 96(18):1959-1966.

Wyman CE. 2007. What is (and is not) vital to advancing cellulosic ethanol. TRENDS in Biotechnology 25(4):153-157.

Yat SC, Berger A, Shonnard DR. 2008. Kinetic characterization for dilute sulfuric acid hydrolysis of timber varieties and switchgrass. Bioresourse Technology 99:3855-3863.

Zamagni A. 2012. Life cycle sustainability assessment. International Journal of Life Cycle Assessment 17:373-376.

Zhang Y-HP, Himmel ME, Mielenz JR. 2006a. Outlook for cellulase improvement: Screening and selection strategies. Biotechnology Advances 24:452-481.

Zhang Y-HP, Lynd LR. 2006b. A Functionally Based Model for Hydrolysis of Cellulose by Fungal Cellulase. Biotechnology and Bioengineering 94(5):888-898.

Zurbriggen B, Bailey MJ, Penttilä ME, Poutanen K, Linko M. 1990. Pilot scale production of a heterolgous *Trichoderma reesei* cellulase by *Saccharomyces cerevisiae*. Journal of Biotechnology 13:267-278.

Appendix A: Land Use Calculations, Feedstock Composition Raw Data, Feedstock Production Inventory, and Feedstock Conversion Inventory

A.1 LAND USE CHANGE CARBON CALCULATIONS

A.1.1 Above Ground Biomass
The annual change in biomass carbon stocks on a dry matter (DM) basis was calculated by to the gain-loss method according to the Equation A.1:

Eq. A.1 $$\Delta C_B = \Delta C_G + \Delta C_{Conversion} - \Delta C_L$$

Where:
ΔC_B = Annual change in biomass carbon stocks on land converted to another category, in tonne-C/yr.
ΔC_G = Annual increase in biomass carbon stocks on land converted to another category, in tonne-C/yr.
$\Delta C_{Conversion}$ = Initial change in biomass carbon stocks on converted land, in tonne-C/yr.
ΔC_L = Annual decrease in biomass carbon stocks on land converted to another category, in tonne-C/yr.

For the Tier 1 approach ΔC_G and ΔC_L are assumed to be zero, leaving $\Delta C_{Conversion}$ the only term to calculate. $\Delta C_{Conversion}$ is described by Equation A.2:

Eq. A.2 $$\Delta C_{Conversion} = \sum \left\{ \left(B_{After} - B_{Before} \right) * CF \right\}$$

Where:
B_{After} = Biomass stocks on land after conversion, in tonne-DM/ha
B_{Before} = Biomass stocks on land before conversion, in tonne-DM/ha
CF = Carbon fraction of dry matter, in tonne-C/tonne-DM

Because all above ground biomass is assumed to be harvested for biofuel production, B_{After} is zero for all cases. Since all land conversion is assumed to be from an identical state of abandoned agricultural land, B_{Before} is the same for all cases. In order to calculate B_{Before} the IPCC default value for above ground biomass in a cold temperate wet grassland was selected (2.4 tonnes DM per hectare, Table 6.4 in the original IPCC document) with a carbon fraction of 0.47 tonnes of carbon per tonne of dry matter. Averaged over 20 years, the result is -0.056 tonnes of carbon per hectare per year. This represents the standing

biomass that was cleared but not utilized in order to establish the plantation. This biomass is mineralized and becomes a carbon emission to the atmosphere.

A.1.2 Below Ground Biomass

$\Delta C_{Conversion}$ for below ground biomass was calculated in the same manner using Equation 2 above. Below ground biomass for the grassland prior to conversion was determined from the default value for above ground biomass (2.4 tonnes per hectare), multiplied by the ratio of below ground biomass to above ground biomass (R) from the IPCC document.

Below ground biomass for the four biomass feedstocks was calculated in the same manner from the above ground biomass at harvest, using the previously stated productivity per hectare per year and harvest period in Table 1.1. R values were taken from IPCC estimates, except for switchgrass which was calculated from Sampson (Sampson 2007). Calculation results are summarized in Table A.1 below.

Table A.1

Above ground biomass, above to below ground ratios, and below ground biomass for conversion from degraded grassland to bioenergy plantation.

	Degraded Ag	Poplar	Willow	Switchgrass	Prairie
Productivity (tonne/(ha.yr)	N/A	13.5	12	9	5
Harvest Cycle (yr)	N/A	7	4	1	1
AGB at harvest (tonne/ha)	2.4	94.5	48	9	5
R	4	0.23	0.23	3.23	4.00
CF	0.47	0.48	0.48	0.47	0.47
Below Ground Carbon (tonne/ha)	9.6	21.74	11.04	29.07	20
ΔCarbon (tonne/(ha.yr)	N/A	0.294	0.037	0.455	0.238

Note that using the stated productivity and harvest cycles for poplar and willow gives an above ground biomass value in the range of 50-300 tonnes per hectare which is indicated as reasonable from Table 4.8 of the IPCC guidelines.

The carbon fraction used for poplar and willow is 0.48, and 0.47 was used for switchgrass and prairie. These below ground biomass and carbon changes are averaged over the initial 20 years of plantation operation gives the results shown in Table A.1 above.

A.1.3 Soil Organic Carbon

Soil organic carbon was calculated according to equation A.1.3 below

Eq. A.1.3 $$\Delta C_{Soils} = \Delta C_{Mineral} - L_{Organic} + \Delta C_{Inorganic}$$

Where:

ΔC_{Soils} is the annual change in carbon stock for the soil in tonnes C per year

$\Delta C_{Mineral}$ is the change in carbon in mineral soils

$L_{Organic}$ is annual loss of carbon from drained organic soils

$\Delta C_{Inorganic}$ is the change in inorganic carbon stocks in soils

$\Delta C_{Inorganic}$ is assumed to be zero except for Tier 3 calculations, and $L_{Organic}$ typically only applies to peat and bog land, leaving only $\Delta C_{Mineral}$ to be determined using Equations A.1.4 and A.1.5 below.

Eq. A.1.4 $$\Delta C_{Mineral} = \frac{SOC_0 - SOC_{0-T}}{T}$$

Eq. A.1.5 $$SOC = SOC_{ref} * F_{LU} * F_{MG} * F_I$$

Where:

SOC_0 is the soil organic carbon in the last year of the inventory period

SOC_{0-T} is the soil organic carbon at the beginning of the inventory period

T is the inventory period (20 years)

SOC_{ref} is the reference carbon stock

F_{LU} is the carbon stock change factor for land use

F_{MG} is the carbon stock change factor for management

F_I is the carbon stock change factor for inputs

The reference soil organic carbon given in the IPCC document is taken to a depth of 30 cm. For cold temperate moist ecosystems with sandy soils, typical of the upper Midwest region, the value is 71 tonnes carbon per hectare. For Tier 1 calculations of forest soils it is assumed that soil carbon stocks do not change from the default value, therefore all carbon stock change factors (F_{LU}, F_{MG}, F_I) are set to 1 for poplar and willow scenarios.

For grasslands, land use carbon stock change factors are set to 1 for a Tier 1 calculation. The stock change factors are taken from Table 6.2 in the IPCC document. The management stock change factor is set to 0.95 for degraded grassland reflecting the loss of carbon in the soil from poor management. For both switchgrass and prairie scenarios the management carbon stock change factor is set to 1.14 representing the improvement due to planting fast-growing,

native perennial grasses. Furthermore, for switchgrass the input carbon change factor is set to 1.11 representing the additional input of fertilizer in that feedstock production scenario. Inputs and results are summarized in Table below.

Table A.2

Soil organic carbon.

	Degraded Ag.	Poplar	Willow	Switchgrass	Prairie
SOC_{Ref}	71	71	71	71	71
F_{LU}	1	1	1	1	1
F_{MG}	0.95	1	1	1.14	1.14
F_I	1	1	1	1.11	1
SOC	67	71	71	90	81
T (yr.)	20	20	20	20	20
ΔSOC (tonne/(ha.yr)	N/A	0.178	0.178	1.142	0.697

A.1.4 Dead Organic Matter

Tier 1 methods assume that the net carbon fluxes in the dead organic matter pool are zero.

A.2 FEEDSTOCK COMPOSITION RAW DATA

Table A.3

Average feedstock composition by percent on a dry weight basis.

	Poplar[a]	**Willow**[b]	**Switchgrass**[c]	**Prairie**[d]	**Logging res.**[e]
Extract	6.72	4.81	11.17	13.96	4.81
Cellulan	41.16	39.84	33.58	30.48	46.91
Xylan	13.22	22.58	22.17	18.57	15.66
Galactan	0.89	1.54	1.02	1.15	0.99
Arabinan	0.82	0.61	2.75	3.07	1.02
Mannan	1.92	1.42	0.36	0.37	3.97
Lignin	24.59	22.54	20.26	18.11	23.75
Ash	1.50	2.00	5.90	7.60	0.30
Acetate	2.51	4.83	3.00	2.33	3.12
Protein	1.00	1.00	3.00	3.00	1.00
Total	94.32	101.17	103.20	98.65	101.52

[a]Poplar values are averaged from Mu et al. (Mu et al. 2010), and EERE samples #1 and #13 (Office of Energy Efficiency and Renewable Energy 2004), with Forbes (Forbes and Society of American Foresters 1961) included for acetate content.
[b]Willow values are taken from Serapiglia et al. (Serapiglia et al. 2009), with Skrigan et al. (Skrigan et al. 1967) included for ash content.

cSwitchgrass values are averaged from Samuel et al. (Samuel et al. 2010) and EERE samples #75, #80, #85, and #90 (Office of Energy Efficiency and Renewable Energy 2004), with Wiselogel et al. (Wiselogel et al. 1996) included for acetate and protein content.
dPrairie values are equal weighted averages for EERE switchgrass (above), EERE sample #131 for Big Bluestem, and EERE sample #141 for Tall Fescue.
eLogging residue composition is taken from Kemppainen (Kemppainen 2003).

A.3 FEEDSTOCK PRODUCTION INVENTORY

A.3.1 Poplar

o   Feedstock production profile was based on a study of SRF poplar growth in a plantation (Gasol et al. 2009)
o   16year plantation cycle, 3-5year rotations
o   13.5 o.d.t/ha*yr, 216 o.d.t/ha, average over plantation cycle, metric tonne

Table A.4
Field operations schedule for poplar plantation.

Year	Operation	Fuel (kg/ha)
1	Existing mow vegetation	39.95
	Base Fertilization	4.25
	Cultivator	8.5
	Herbicide treatment	5.1
	Plantation	15.3
	Harvest (nails)	25.5
3	Top fertilization	2.55
4	Insecticide treatment	5.1
6	Harvest	34
7	Base fertilization	4.25
8	Fungi or insecticide treatment	5.1
9	Top fertilization	2.55
11	Harvest	34
12	Base fertilization	4.25
13	Fungi or insecticide treatment	5.1
14	Top fertilization	2.55
16	Harvest	34
	Stool killdown	41.82
	Stool collection	6.97
		280.84

Table A.5

Fertilizer schedule for poplar plantation.

Year	Fertilizer	Amount (kg/ha)	N (kg/ha)	P (kg/ha)	P, as P_2O_5 (kg/ha)	K (kg/ha)	K, as K_2O (kg/ha)
1	(9-18-27)	600	54.00	108	247.5	162	195.1
3	(33.5-0-0)	250	83.75	0	0	0	0
7	(9-18-27)	600	54.00	108	247.5	162	195.1
9	(33.5-0-0)	250	83.75	0	0	0	0
12	(9-18-27)	600	54.00	108	247.5	162	195.1
14	(33.5-0-0)	250	83.75	0	0	0	0
Total			**413.25**	**324**	**742.5**	**486**	**585.3**

(9-18-27) and (33.5-0-0) blends assumed ammonium nitrate as N, Thomas meal:single superphosphate as P, potassium sulfate as K

Table A.6

Herbicide/pesticide schedule for poplar plantation.

Year	Material	SimaPro Proxy	Amount (kg/ha)
1	Herbicide Treatment	Glyphosate	6.816
4	Insecticide treatment	Generic Pest.	0.75
8	Fungi Treatment	Generic Pest.	0.75
8	Insecticide treatment	Generic Pest.	0.75
13	Fungi Treatment	Generic Pest.	0.75
13	Insecticide treatment	Generic Pest.	0.75

Table A.7

Elemental composition and nutrient requirements for poplar, adapted from Sannigrahi et al. (Sannigrahi et al. 2010).

	C	H	N	S	O	*P	*K
Ultimate Analysis (% o.d.m.)	50.29	6.12	0.42	0.03	41.52	0.06	0.21
Extracted over Plantation Cycle (o.d.t./ha)	108.63	13.22	0.91	0.06	89.68	**0.13**	**0.45**
Applied over Plantation Cycle (o.d.t./ha)	-	-	0.41	-	-	**0.32**	**0.49**

Table A.8

Direct emissions from poplar cultivation

Chemical	Origin	Environmental Compartment	Amount (kg/ha)
Carbon Dioxide	Fuel Combustion	Air	891
Dinitrogen Monoxide,	Fertilizer	Air	8.6
Nitrate, Eutrophication	Fertilizer	Water	549

EF of 3.172 kg CO_2/kg diesel, from stoicheometery. EF of 0.01325 kg N_2O/kg N in fertilizer. EF of 0.3 kg nitrate/kg N in fertilizer Eutrophication

A.3.2 Willow

o Feedstock production profile was produced from study of SRF willow growth in a plantation (Heller et al. 2003).
o 23year plantation cycle, 7-3year rotations
o 10 o.d.t/ha*yr for first 3year cycle, 13.6 o.d.t/ha*yr for remaining, 274.8 o.d.t./ha over plantation cycle, metric tonne

Table A.9
Field operations schedule for willow plantation.

Year	Operation	Fuel (kg/ha)	Oil (kg/ha)
0	Mow Existing Vegetation	11.17	0.04
	Apply Contact Herbicide	3.96	0.02
	Plow	24.31	0.09
	Disk	36.50	0.13
	Seed Broadcasting, Cropcover	0.92	0.00(2)
	Cultipack	110.11	0.07
1	Disk	30.36	0.13
	Cultipack	110.11	0.07
	Planting	48.08	0.15
	Apply Pre-emergent Herbicide	3.65	0.02
	Mechanical Weed Control(1)	1.44	0.01
	Mechanical Weed Control(2)	7.80	0.02
	Chemical Weed Control	0.18	0.00(1)
	1st Year Coppice	20.31	0.07
2	Fertilize	3.67	0.01
4	Harvest	84.11	0.18
5	Fertilize	3.67	0.01
7	Harvest	84.11	0.18
8	Fertilize	3.67	0.01
10	Harvest	84.11	0.18
11	Fertilize	3.67	0.01
13	Harvest	84.11	0.18
14	Fertilize	3.67	0.01
16	Harvest	84.11	0.18
17	Fertilize	3.67	0.01
19	Harvest	84.11	0.18
20	Fertilize	3.67	0.01
22	Harvest	84.11	0.18
23	Stool Elimination	3.96	0.02
	Plow	24.31	0.09
		1051.63	2.26

Crop cover used 59.25 kg/ha winter rye.

Table A.10
Nursery inputs for willow plantation.

Input	Amount (kg/ha)
Diesel Oil	6.47
LPG	1.01
Gasoline	18.78
Electricity	301.68 kWh
Heavy Fuel Oil (HFO)	70.80
Wood (for heat)	43.44
Carbaryl (insecticide)	0.22
Glyphosate (herbicide)	0.12
Fertilizer (15-15-15), Granular	110.25
Ammonium Sulfate Fertilizer	8.35
Urea Fertilizer	8.35
Surface Water	365,441

Fertilizer and pesticide inputs listed here and below were not double-counted, but included for display purposes.

Table A.11
Fertilizer schedule for willow plantation.

Year	Fertilizer	Amount (kg/ha)	N (kg/ha)	P (kg/ha)	P, as P_2O_5 (kg/ha)	K (kg/ha)	K, as K_2O (kg/ha)
Nursery	(15-15-15)	110.25	16.54	16.54	37.90	16.54	19.92
2	Ammonium Sulfate	471.79	100	0	0	0	0
5	Ammonium Sulfate	471.79	100	0	0	0	0
8	Ammonium Sulfate	471.79	100	0	0	0	0
11	Ammonium Sulfate	471.79	100	0	0	0	0
14	Ammonium Sulfate	471.79	100	0	0	0	0
17	Ammonium Sulfate	471.79	100	0	0	0	0
20	Ammonium Sulfate	471.79	100	0	0	0	0
			716.54	16.54	37.90	16.54	19.92

(15-15-15) blend was assumed ammonium nitrate as N, Thomas meal:single superphosphate as P, potassium sulfate as K

Table A.12

Herbicide/pesticide schedule for willow plantation

Year	Material	SimaPro Proxy	Amount (kg/ha)
Nursery	Carbaryl (insecticide)	Generic Pest.	0.22
	Glyphosate (herbicide)	Glyphosate	0.12
1	Glyphosate (herbicide)	Glyphosate	2.5
	Simazine (herbicide)	Generic Pest.	2.35
	Oxyfluorfen (herbicide)	Generic Pest.	1.12
23	Glyphosate (herbicide)	Glyphosate	2.5

Table A.13

Elemental composition and nutrient requirements for willow, adapted from Sannigrahi et al. (Sannigrahi et al. 2010).

	C	H	N	S	O	*P	*K
Ultimate Analysis (% o.d.m.)	49.40	6.01	0.45	0.05	42.90	0.05	0.18
Extracted over Plantation Cycle (o.d.t./ha)	135.75	16.52	1.24	0.14	117.89	0.14	0.05
Applied over Plantation Cycle (o.d.t./ha)	-	-	0.71	-	-	0.04	0.02

Table A.14: Direct emissions from willow cultivation.

Chemical	Origin	Environmental Compartment	Amount (kg/ha)
Carbon Dioxide	Fuel Combustion	Air	3348
	Fertilizer	Air	6.22
Dinitrogen Monoxide,	Fertilizer	Air	15
Nitrate, Eutrophication	Fertilizer	Water	960

EF of 3.172 kg CO_2/kg diesel, from stoicheometery. EF of 1.594 kg CO_2/kg N in urea fertilizer. EF of 0.01325 kg N_2O/kg N in fertilizer. EF of 0.3 kg nitrate/kg N in fertilizer Eutrophication

A.3.3 Switchgrass

o   Feedstock production profile was produced from a switchgrass production and management guide (Sampson 2007)

o   10year plantation cycle, yearly harvest

o   3 o.d.t/ha*yr for first year, 7 o.d.t/ha*yr for second year, 10 o.d.t./ha*yr for remaining, 90.0 o.d.t./ha over plantation cycle, metric tonne

Table A.15
Field operations for switchgrass production

Stage	Input/Operation	Input (kg/ha)
Establishment	Seed	9.0
	Planting (Ecoinvent)	*1*
	Rotary Harrow (Ecoinvent)	*2*
	Rotary Till (Ecoinvent)	*2*
Operation	Ammonium Sulfate, as N	550.0
	Fertilizing by Broadcaster (Ecoinvent)	*0.68*
	Mowing (Ecoinvent)	*10*
	Baling (Ecoinvent)	*129*
	Generic Pesticide	16
	Application of Plant Protectants by Field Sprayer (Ecoinvent)	*0.01*

Field operations were constructed from SimaPro EcoProfiles in the Ecoinvent Database. These profiles include all inputs, and are constructed on a 'per-ha' basis. Harrow and Tilling were assumed to use 2 passes/ha. Application of fertilizers and pesticides application area bases were scaled based on the capacity of the implements. Example: fertilizer broadcaster capacity is 500L, so spreading 339.5L fertilizer requires 0.679 of its capacity. This is directly applied to the input, 0.679 ha input for 1.0 ha basis.

Table A.16
Elemental composition and nutrient requirements for switchgrass, adapted from Sannigrahi et al. (Sannigrahi et al. 2010).

	*C	*H	*N	*S	*O	*P	*K
Ultimate Analysis (% o.d.m.)	47.75	5.75	0.74	0.08	42.37	0.05	0.07
Extracted over Plantation Cycle (o.d.t./ha)	42.98	5.18	0.67	0.07	39.13	0.05	0.06
Applied over Plantation Cycle (o.d.t./ha)			0.55			0.00	0.00

Table A.17
Direct emissions from switchgrass cultivation

Chemical	Origin	Environmental Compartment	Amount (kg/ha)
Carbon Dioxide	Fuel Combustion	Air	(Included in Ecoinvent)
Dinitrogen Monoxide,	Fertilizer	Air	11.5
Nitrate, Eutrophication	Fertilizer	Water	731

EF of 0.01325 kg N_2O/kg N in fertilizer. EF of 0.3 kg nitrate/kg N in fertilizer Eutrophication.

A.3.4 Prairie
o   Assumed to be identical to switchgrass, without management (fertilizer/pesticides) and half the productivity as indicated by Tilman et al. Tilman et al. (David Tilman et al. 2006).
o   10 yr plantation cycle
o   5 o.d.t./ha*yr productivity

Table A.18

Field operations for prairie production.

Stage	Input/Operation	Input (kg/ha)
Establishment	Seed	9.0
	Planting (Ecoinvent)	1
	Rotary Harrow (Ecoinvent)	2
	Rotary Till (Ecoinvent)	2
Operation	Mowing (Ecoinvent)	10
	Baling (Ecoinvent)	71.4

Table A.19

Direct emissions from prairie production.

Chemical	Origin	Environmental Compartment	Amount (kg/ha)
Carbon Dioxide	Fuel Combustion	Air	(Included in Ecoinvent)

A.3.5 Logging Residues

Based on a hardwood loggers survey carried out in previous work at Michigan Technological University (Reis 2009). The residues are considered to be a waste from existing harvesting operations, and therefore only includes inputs for collection/size reduction. Inputs assume 0.25 short ton/acre yield (0.56 tonne/ha).

Table A.20

Field operations for logging residue collection.

Input/Operation	Input (kg/ha)
Harvesting	1.87
Forwarding	1.04
Chipping	1.18
Machine Infrastructure	1.03E-5 p/ha

Table A.21

Direct emissions from logging residue collection.

Chemical	Origin	Environmental Compartment	Amount (kg/ha)
Carbon Dioxide	Fuel Combustion	Air	12.97

EF of 3.172 kg CO_2/kg diesel, from stoicheometery.

A.4 FEEDSTOCK CONVERSION INVENTORY

Process conversion inputs were obtained from Aspen simulations based on feedstock composition. Inputs are on a 1 MJ produced ethanol basis. Material flows are in kg unless otherwise noted.

Table A.22

Feedstock conversion results from SimaPro 7.2.

Conversion Stage	Poplar	Willow	Switchgrass	Prairie	Residues
Feedstock Requirement	1.38E-01	1.31E-01	1.45E-01	1.56E-01	1.29E-01
A200 – Pretreatment (g-CO2 eq)	*4.69*	*6.04*	*5.12*	*5.04*	*4.69*
Sulphuric acid	3.05E-03	2.90E-03	3.20E-03	3.43E-03	2.85E-03
Ammonia	2.06E-03	2.71E-03	2.25E-03	2.20E-03	2.07E-03
A300 – Fermentation (g-CO2 eq)	*1.38*	*1.33*	*1.45*	*1.55*	*1.30*
Corn Steep Liquor	1.91E-03	1.84E-03	2.01E-03	2.15E-03	1.80E-03
Diammonium phosphate, as N	2.36E-04	2.28E-04	2.48E-04	2.66E-04	2.23E-04
Sugar, from sugar beet	7.37E-05	7.05E-05	7.74E-05	8.31E-05	6.92E-05
A400 – Enzyme Production (g-CO2 eq)	*3.35*	*2.88*	*2.61*	*2.67*	*3.31*
Corn Steep Liquor	3.39E-04	2.91E-04	2.64E-04	2.71E-04	3.36E-04
Sugar, from sugar beet	4.99E-03	4.27E-03	3.88E-03	3.98E-03	4.94E-03
Ammonia	2.37E-04	2.03E-04	1.85E-04	1.89E-04	2.35E-04
Sulphur dioxide	3.49E-05	2.99E-05	2.71E-05	2.78E-05	3.45E-05
Soybean oil	2.75E-05	2.36E-05	2.15E-05	2.20E-05	2.73E-05
Ammonium sulphate, as N	9.51E-06	8.14E-06	7.40E-06	7.58E-06	9.40E-06
Phosphoric acid, 85% in H2O	5.47E-05	4.69E-05	4.26E-05	4.37E-05	5.42E-05
Potassium hydroxide	2.67E-05	2.28E-05	2.07E-05	2.13E-05	2.63E-05
Magnesium sulphate	9.69E-06	8.30E-06	7.55E-06	7.73E-06	9.59E-06
Calcium chloride	1.29E-05	1.11E-05	1.01E-05	1.03E-05	1.28E-05
Ethoxylated alcohols	6.46E-06	5.54E-06	5.03E-06	5.16E-06	6.39E-06
A600 – Waste Water Treatment (g-CO2 eq)	*8.81*	*11.67*	*9.64*	*9.35*	*8.90*
Sodium hydroxide, 50% in H2O	8.83E-03	1.17E-02	9.66E-03	9.38E-03	8.92E-03
A800 – Boiler (g-CO2 eq)	*0.95*	*0.88*	*1.09*	*1.23*	*0.89*
Lime, hydrated	1.26E-03	1.17E-03	1.45E-03	1.64E-03	1.18E-03
A900 – Utilities (g-CO2 eq)	*0.00*	*0.00*	*0.01*	*0.01*	*0.00*
Water, completely softened	1.84E-01	1.44E-01	2.75E-01	2.97E-01	1.56E-01
Total Conversion Process (g-CO2 eq)	*19.18*	*22.80*	*19.92*	*19.85*	*19.09*

Appendix B: Monomer and Oligomer Sugar Recoveries

The following material provides sugar monomer and oligomer recoveries following dilute acid pretreatment, and sugar monomer recoveries following enzymatic hydrolysis for each experiment. For dilute acid pretreatment, glucose, xylose, galactose, and combined arabinose plus mannose data are provided. Error bars represent the standard deviations for three experimental replicates at each time point. For enzymatic hydrolysis each time series represents pretreated solids from one individual reactor treated with enzymes at one of three different loading levels; 0.1, 0.25, or 0.5 ml enzyme solution per gram cellulose. Legend entries for the enzymatic hydrolysis results indicate pretreatment times relative to the optimal pretreatment time (1/2x, 1x, and 2x), and enzyme loading (low, medium, and high). Only glucose and xylose monomers were detected following enzymatic hydrolysis, except for balsam which produced only glucose.

B.1 DILUTE ACID PRETREATEMENT

The graphs appearing in this section show sugar recoveries during dilute acid pretreatment relative to initial carbohydrate content of the biomass expressed in monomer equivalents, from 0-100%. Xylose is the major sugar released from hemicellulose during dilute acid pretreatment, although mannose is a significant sugar in balsam experiments. Xylose monomer yields are generally lower at the earliest time point, achieve a maximum at the middle time point as xylose oligomers are converted to monomers, then decrease at the latest time point as xylose and other sugars are degraded into byproducts such as furfural, hydroxymethyl furfural, and tars. Xylose oligomers are highest at the earliest time point, then decrease at later time points as oligomers are converted to monomers, and then to degradation products. Galactose, arabinose, and mannose follow similar trends to xylose. Small amounts of glucose are also liberated during dilute acid pretreatment from hemicellulose and a small amount of cellulose degradation. Because glucose is much more resistant to acid hydrolysis, glucose monomer continues to accumulate in solution throughout the reaction, although it remains low, generally less than 5%. Galactose yield for the switchgrass experiment is reported as greater than 100% (figures S7 and S8). This is most likely due to incomplete separation of the galactose peak from the xylose peak, which elute close together in time. This indicates that HPLC maintenance was most likely necessary, although galactose represents only about 1% of the switchgrass feedstock, so small differences in measurement result in large differences when reported on a percent basis.
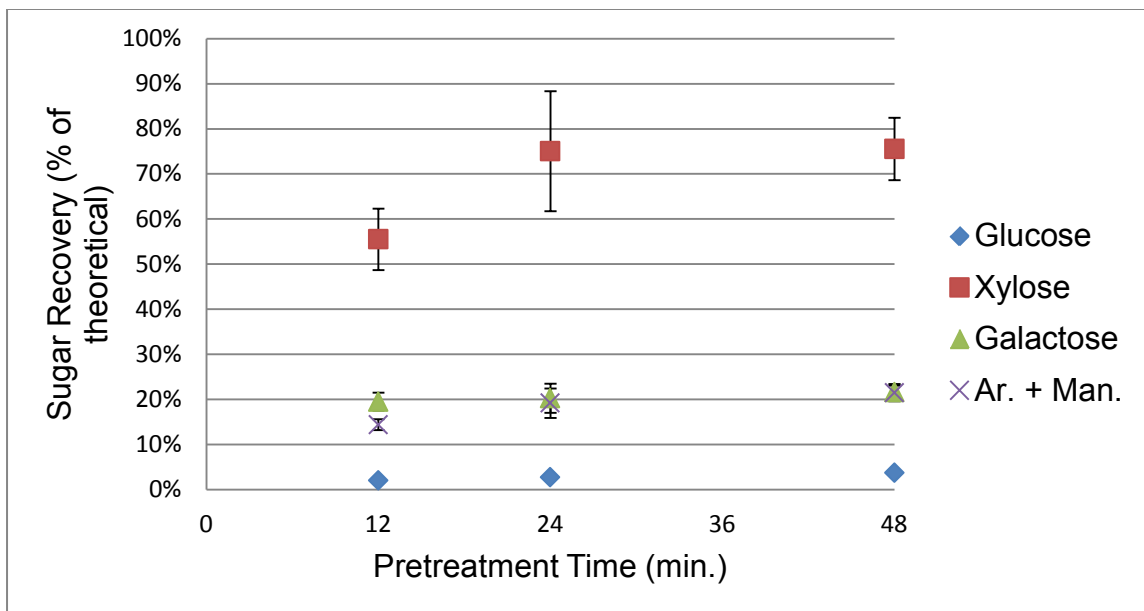
Figure B.1: Dilute acid pretreatment sugar monomer recoveries for the aspen (1) experiment, performed 13-Sep. 2010.
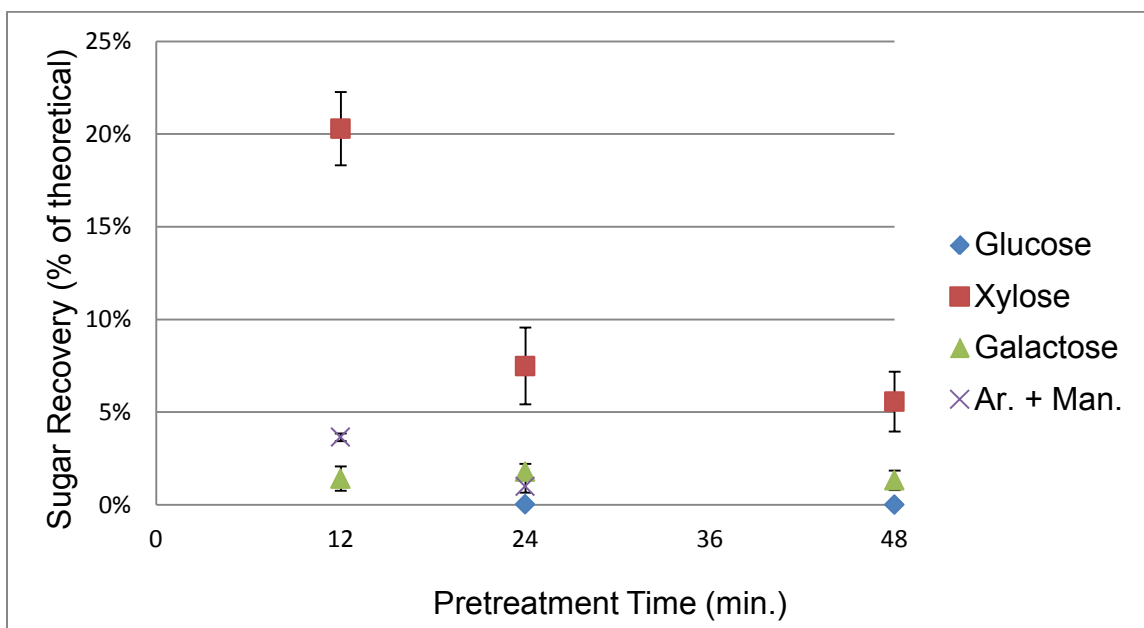
Figure B.2: Dilute acid pretreatment sugar oligomer recoveries for the aspen (1) experiment, performed 13-Sep. 2010.
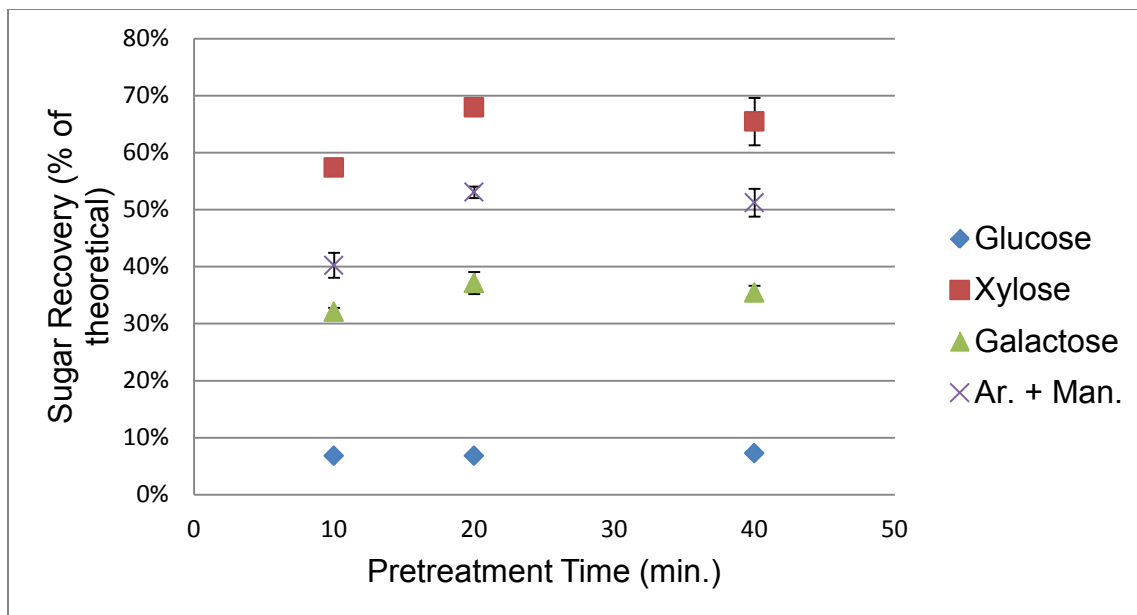
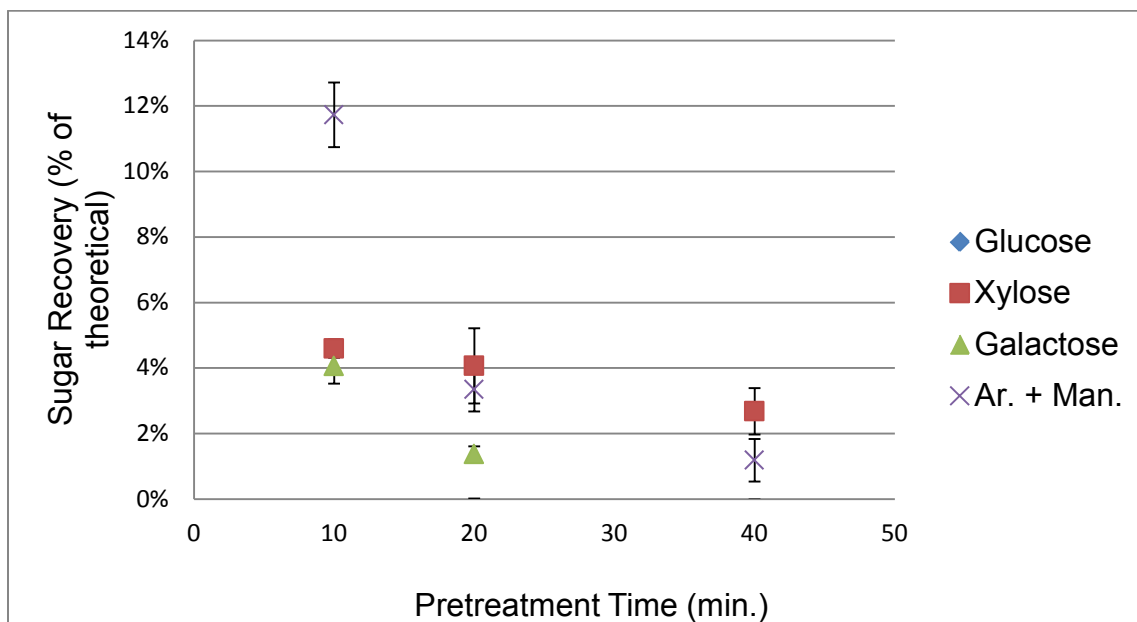**Figure B.3:** Dilute acid pretreatment sugar monomer recoveries for the balsam experiment, performed 28-Sep. 2010.
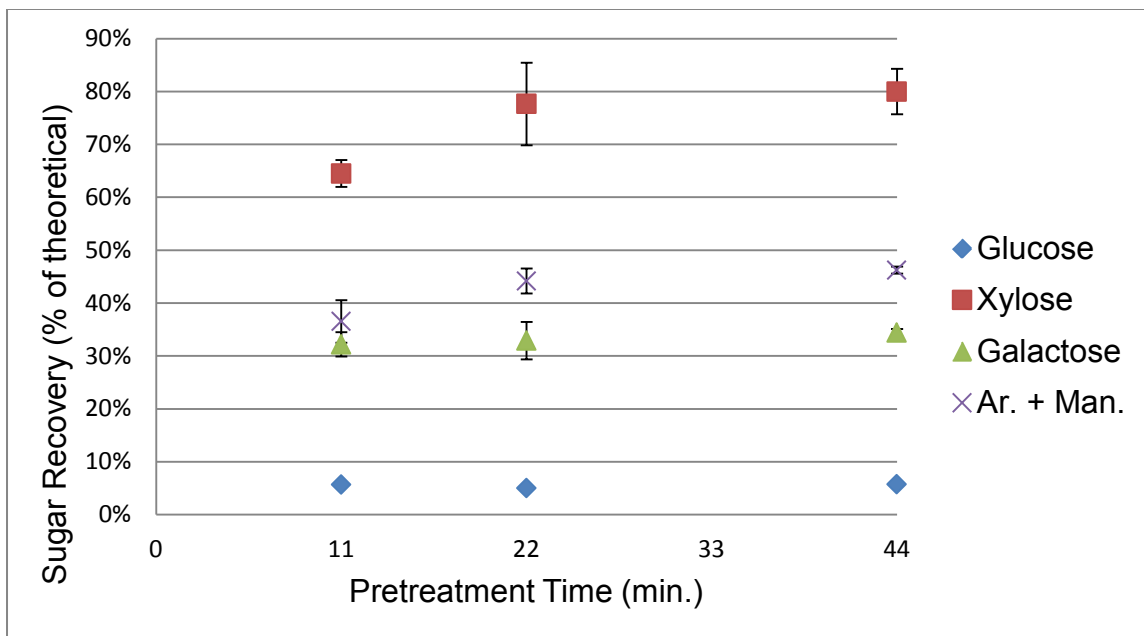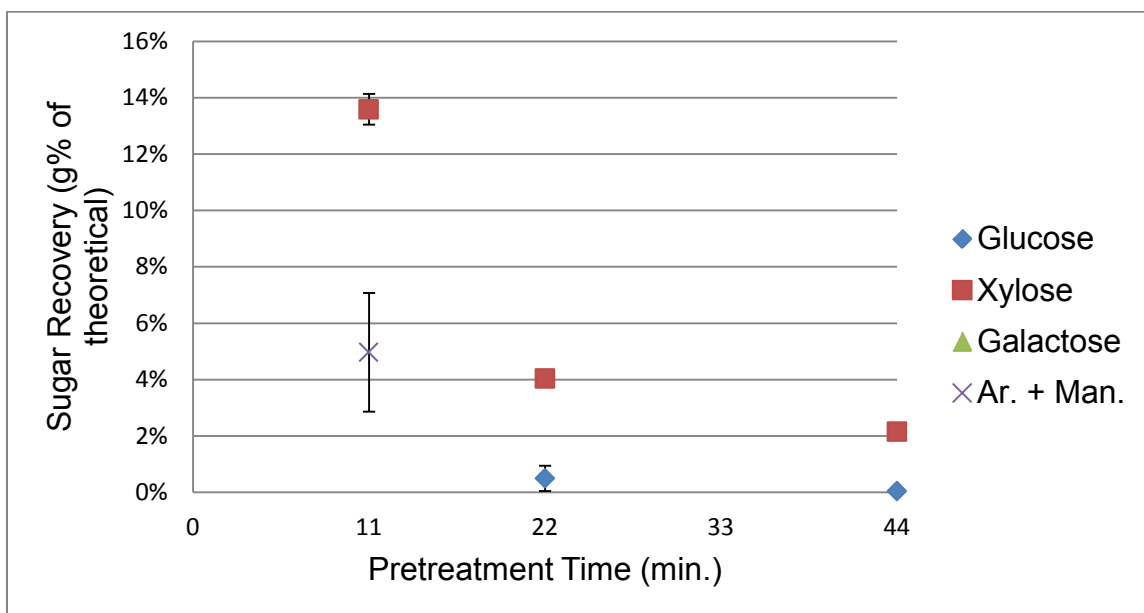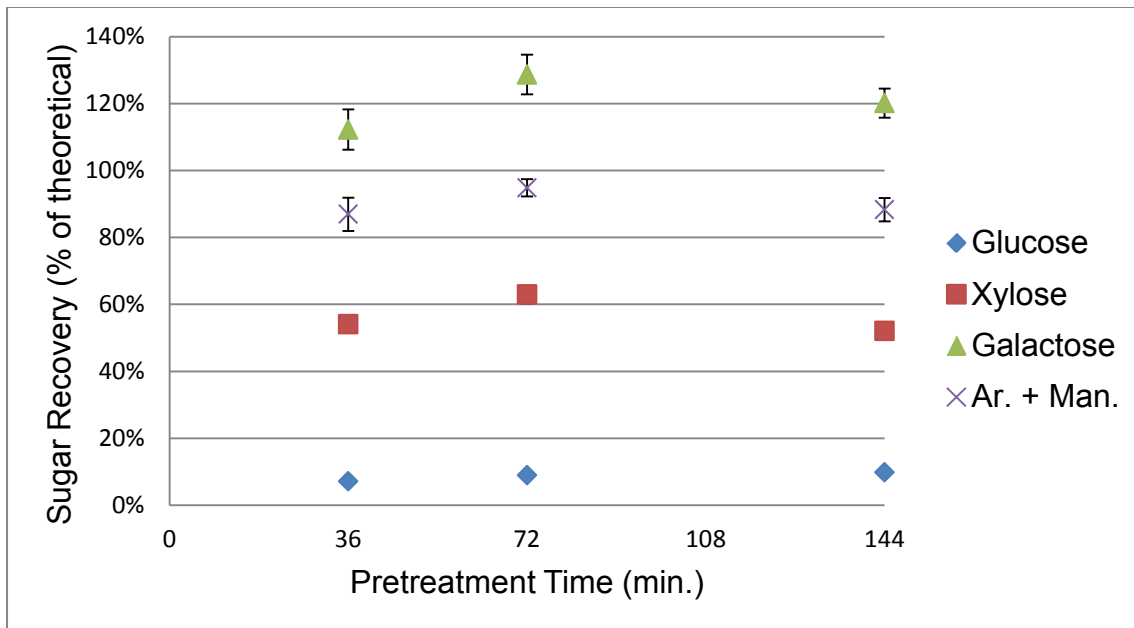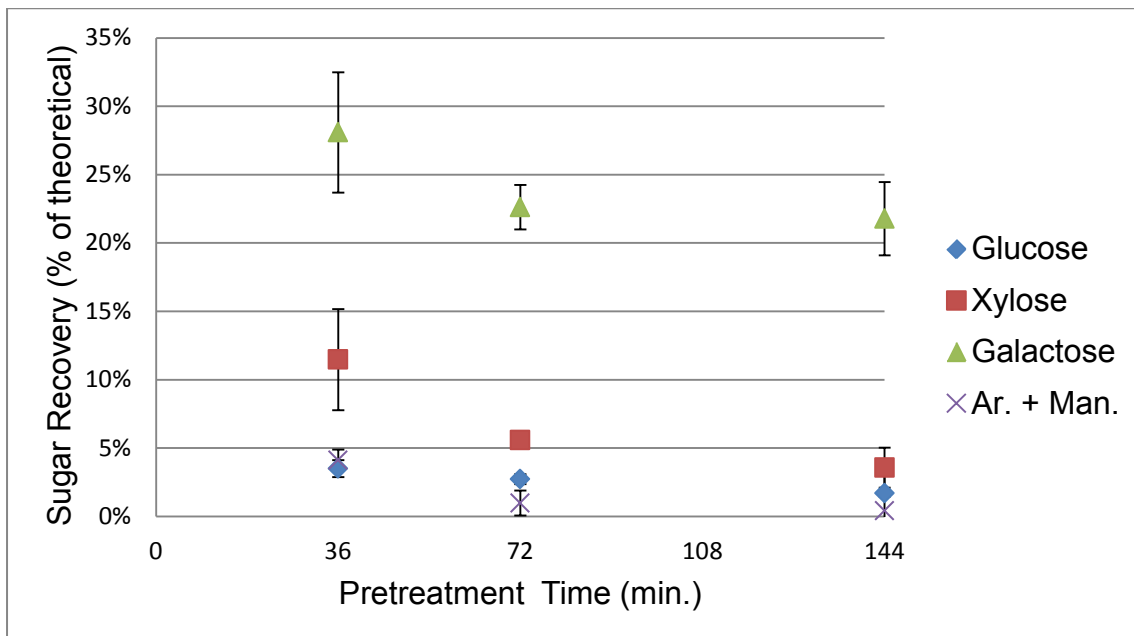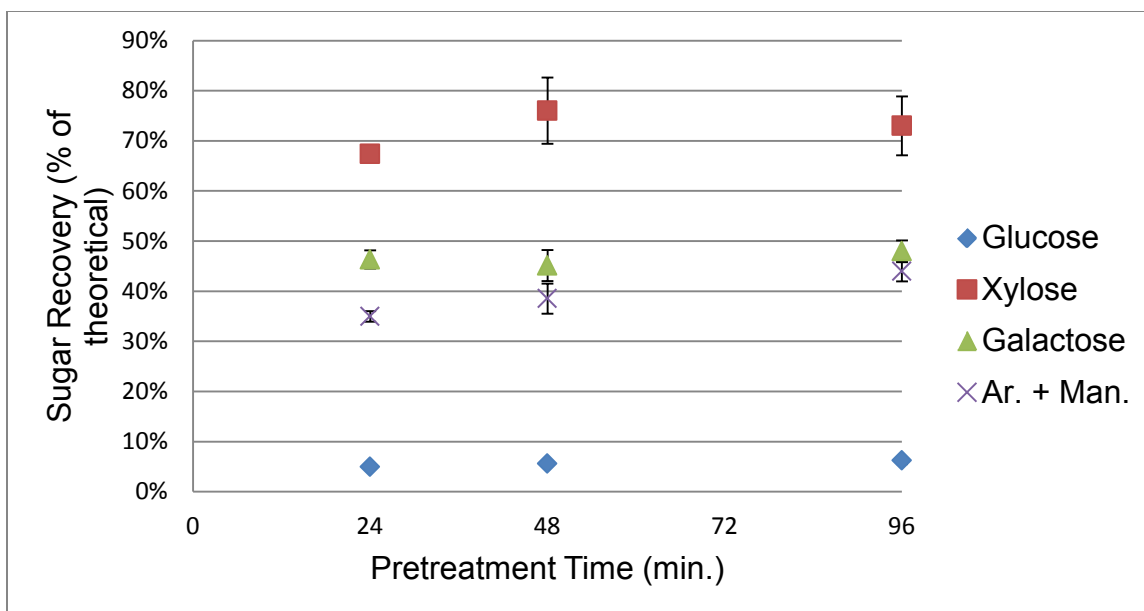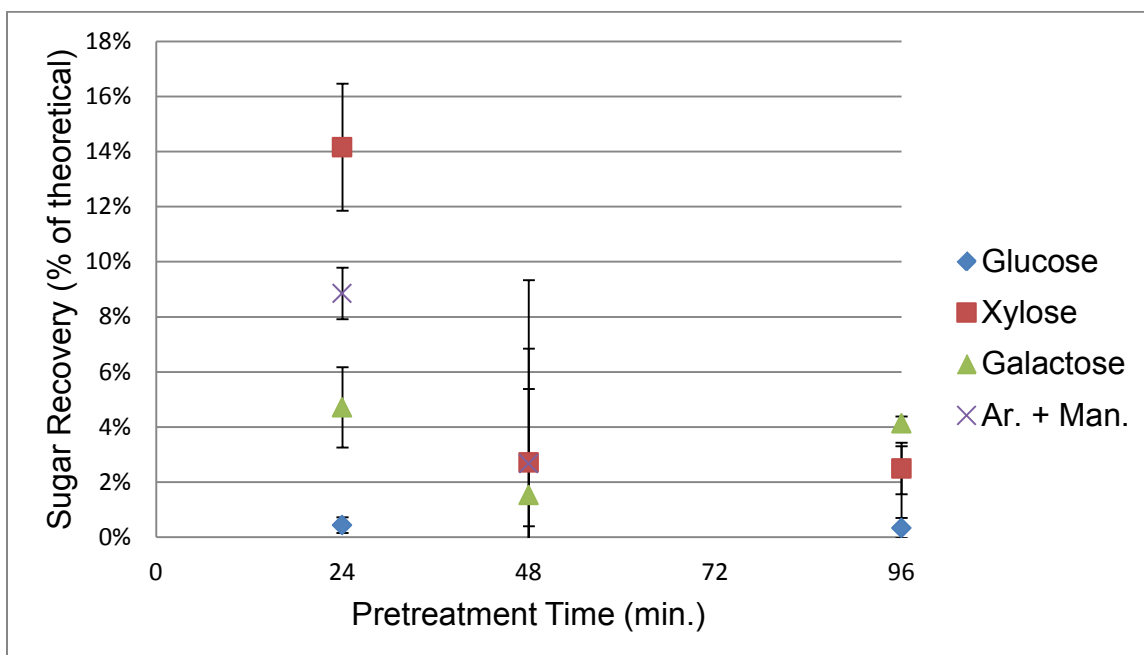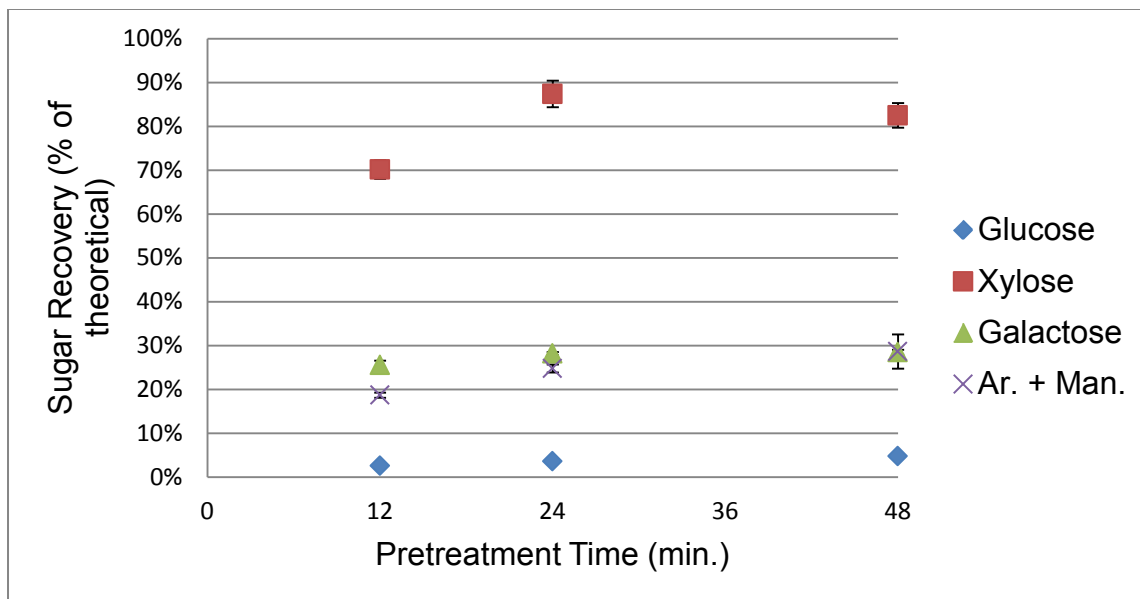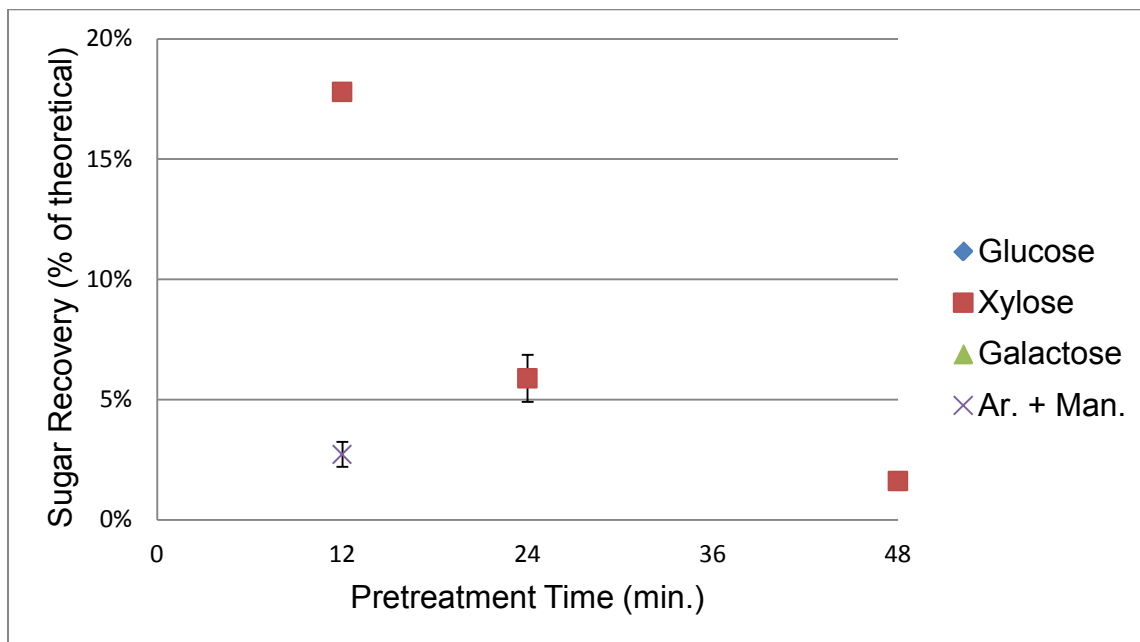

Figure B.4: Dilute acid pretreatment sugar oligomer recoveries for the balsam experiment, performed 28-Sep. 2010.

Figure B.5: Dilute acid pretreatment sugar monomer recoveries for the aspen/balsam experiment, performed 12-Oct. 2010.

Figure B.6: Dilute acid pretreatment sugar oligomer recoveries for the aspen/balsam experiment, performed 12-Oct. 2010.

Figure B.7: Dilute acid pretreatment sugar monomer recoveries for the switchgrass experiment, performed 19-Oct. 2010.

Figure B.8: Dilute acid pretreatment sugar oligomer recoveries for the switchgrass experiment, performed 19-Oct. 2010.

Figure B.9: Dilute acid pretreatment sugar monomer recoveries for the aspen/switchgrass experiment, performed 26-Oct. 2010.

Figure B.10: Dilute acid pretreatment sugar oligomer recoveries for the aspen/switchgrass experiment, performed 26-Oct. 2010.

Figure B.11: Dilute acid pretreatment sugar monomer recoveries for the aspen (2) experiment, performed 12-Dec. 2010.

Figure B.12: Dilute acid pretreatment sugar oligomer recoveries for the aspen (2) experiment, performed 12-Dec. 2010.

B.2 ENZYMATIC HYDROLYSIS

The graphs appearing in this section show sugar recoveries during enzymatic hydrolysis relative to initial carbohydrate content of the biomass expressed in monomer equivalents, from 0-100%. Glucose is the main sugar released from enzymatic hydrolysis, and glucose and xylose monomer only were detected in supernatants from enzymatic hydrolysis, except for balsam which produced only glucose. Enzymatic hydrolysis generally shows a continuous accumulation of sugars in solution over three days, with the fastest accumulation in the first 24 hours. Yields of 50-60% of theoretical glucose recovery were observed for aspen and switchgrass, with an additional 5-10% of xylose released. Yields were much lower for balsam-containing experiments as this species is known to not respond well to enzymatic hydrolysis. Enzymatic hydrolysis glucose yields at 72 hours of reaction were seen to be strongly affected by both pretreatment severity and enzyme loading, with the highest enzyme loading and most severe pretreatment condition producing the most glucose. However total sugar recovery was usually highest for the moderate pretreatment condition with the highest enzyme loading, due to monomer sugar degradation at the severest pretreatment condition. Xylose recovery was generally highest for the high enzyme loading and the lowest pretreatment severity. This is indicative of the incomplete hemicellulose hydrolysis from pretreatment, and again total sugar recovery was lower than for the optimum conditions. Both glucose and xylose recovery were significantly improved for aspen when supplemental β-glucosidase was added to the enzyme mixture, showing that enzyme systems must still be further optimized for commercial implementation.
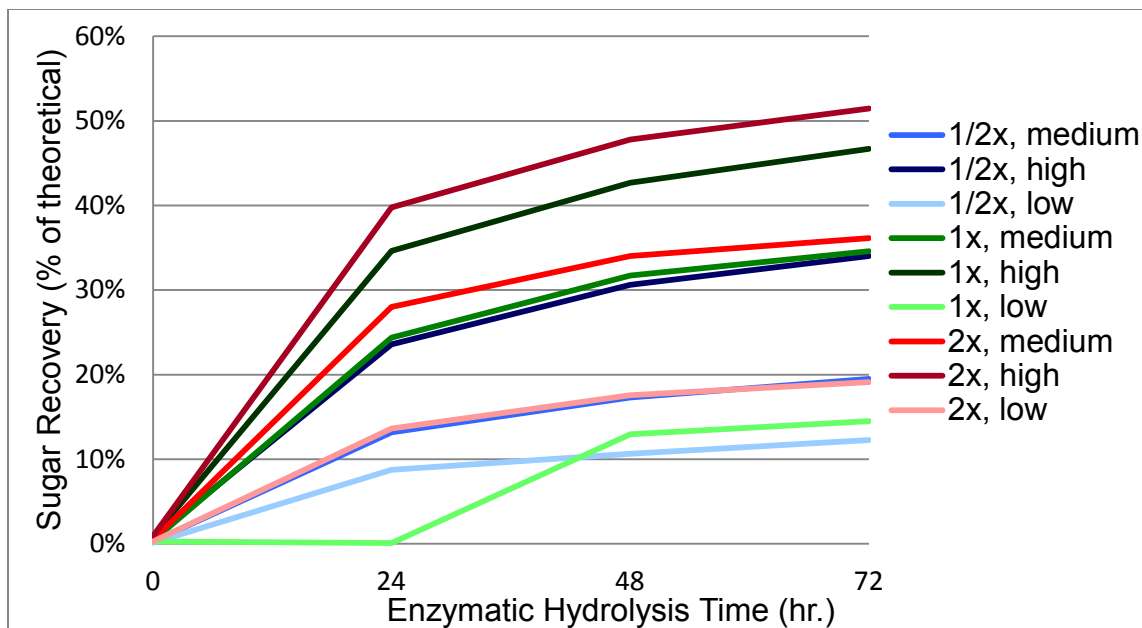
Figure B.13: Enzymatic hydrolysis glucose monomer recoveries for the aspen (1) experiment, performed 13-Sep. 2010.
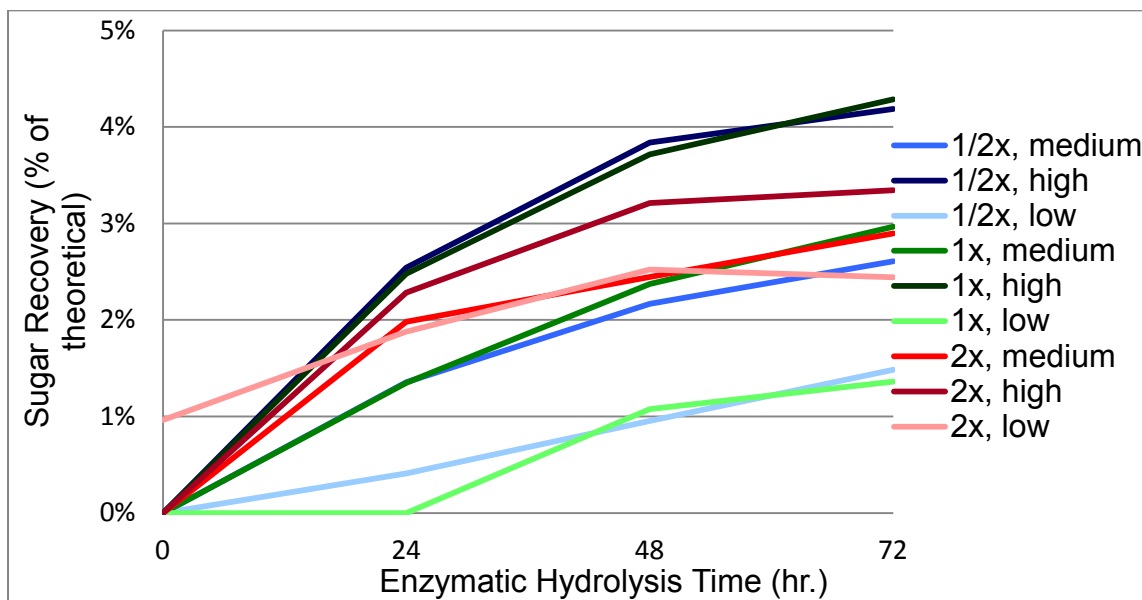

Figure B.14: Enzymatic hydrolysis xylose monomer recoveries for the aspen (1) experiment, performed 13-Sep. 2010.
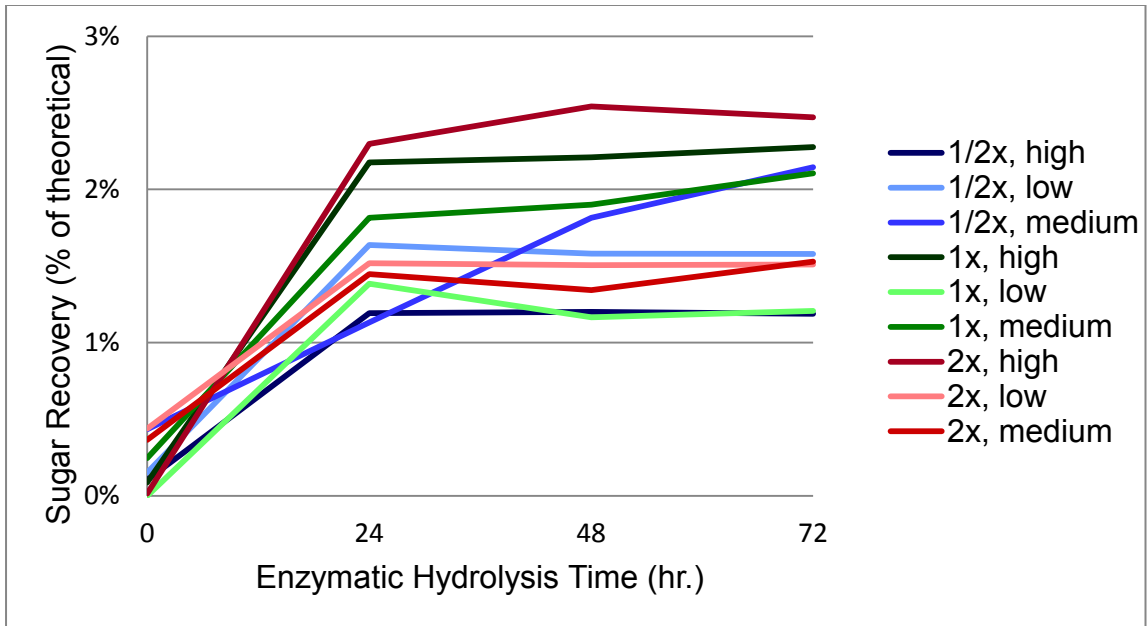
Figure B.15: Enzymatic hydrolysis glucose monomer recoveries for the balsam experiment, performed 28-Sep. 2010.
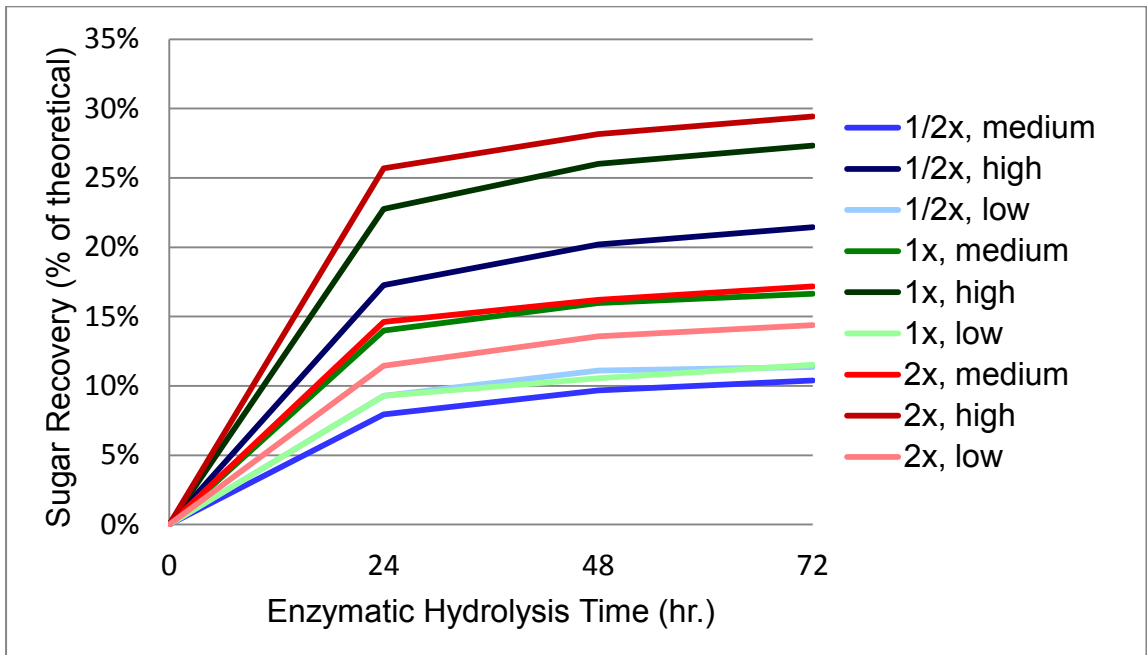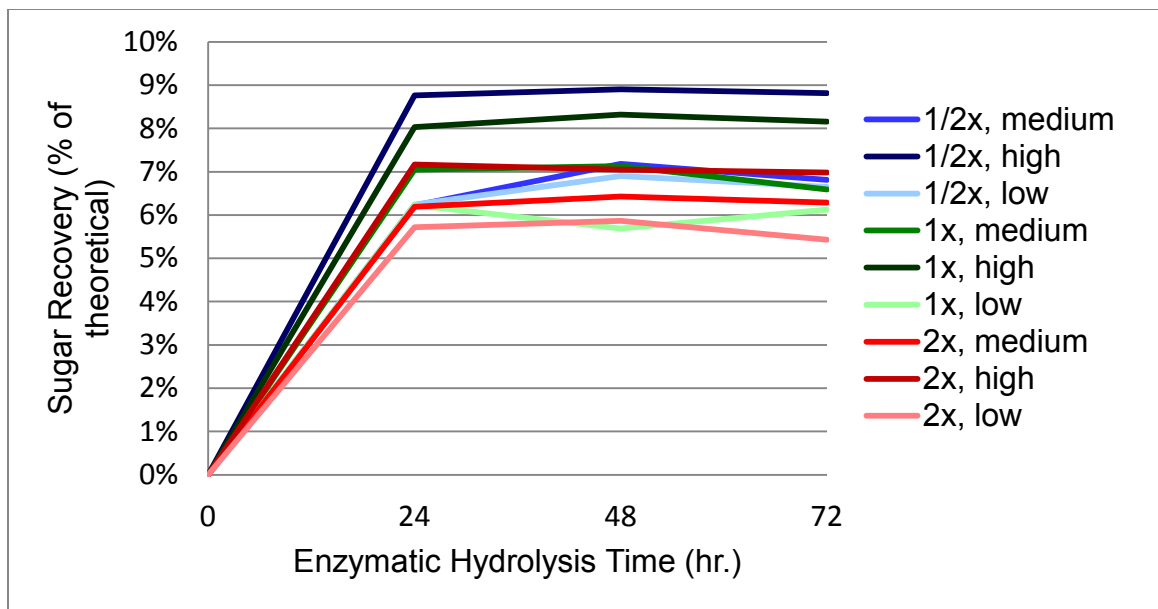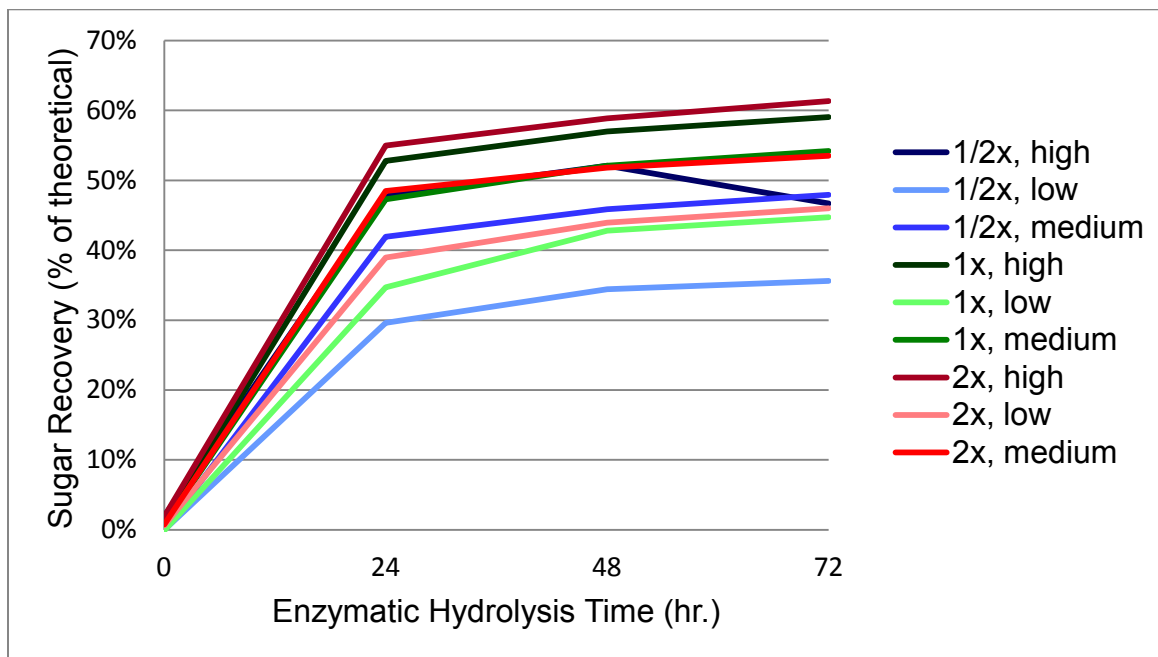

Figure B.16: Enzymatic hydrolysis glucose monomer recoveries for the aspen/balsam experiment, performed 12-Oct. 2010.

Figure B.17: Enzymatic hydrolysis xylose monomer recoveries for the aspen/balsam experiment, performed 12-Oct. 2010.

Figure B.18: Enzymatic hydrolysis glucose monomer recoveries for the switchgrass experiment, performed 19-Oct. 2010.
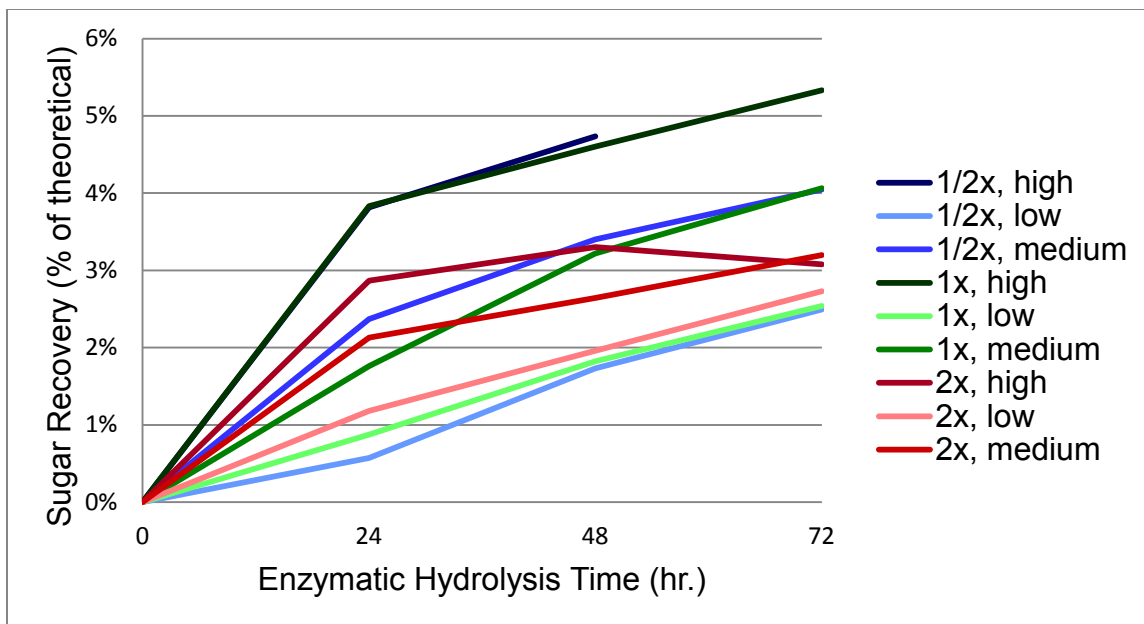
Figure B.19: Enzymatic hydrolysis xylose monomer recoveries for the switchgrass experiment, performed 19-Oct. 2010.
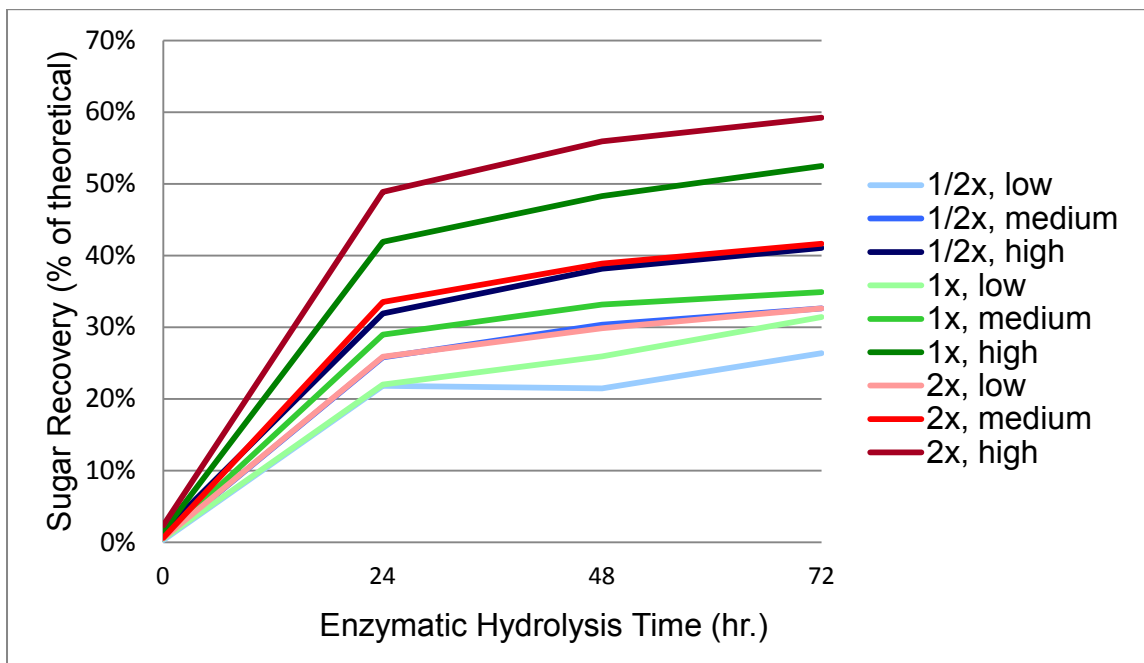

Figure B.20: Enzymatic hydrolysis glucose monomer recoveries for the aspen/switchgrass experiment, performed 26-Oct. 2010.
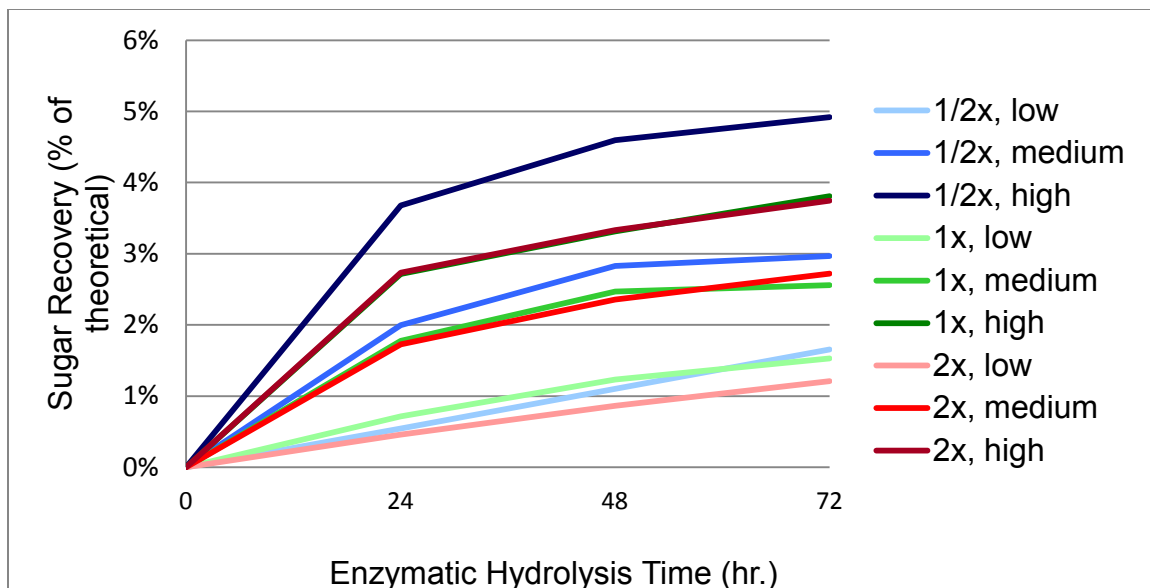
Figure B.21: Enzymatic hydrolysis xylose monomer recoveries for the aspen/switchgrass experiment, performed 26-Oct. 2010.
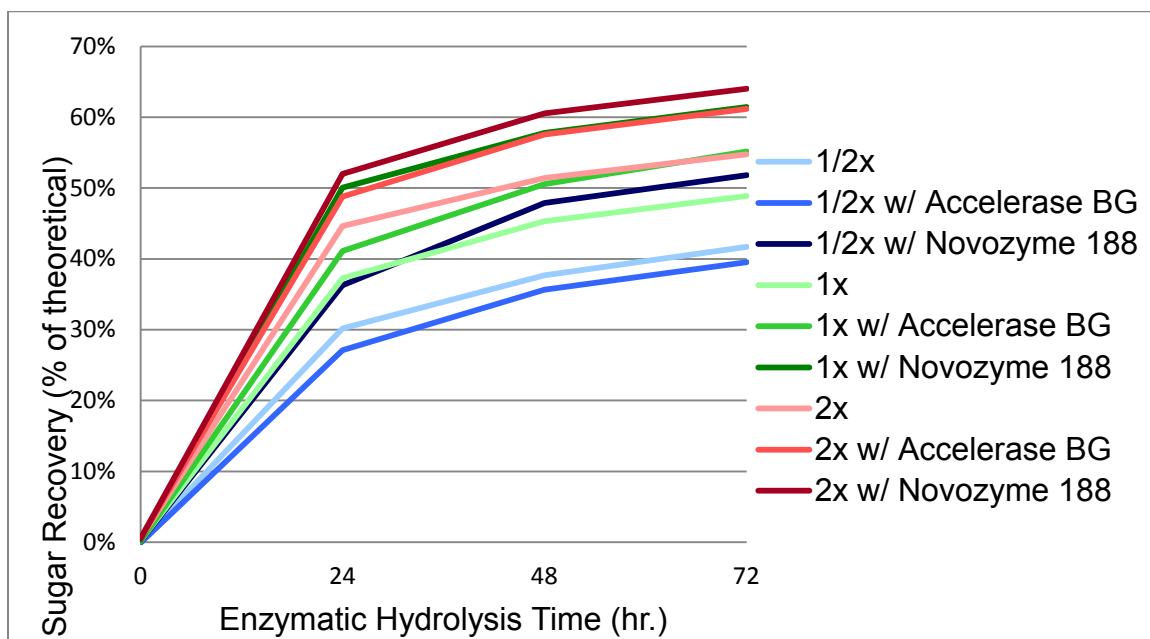

Figure B.22: Enzymatic hydrolysis glucose monomer recoveries for the aspen (2) experiment, performed 12-Dec. 2010.
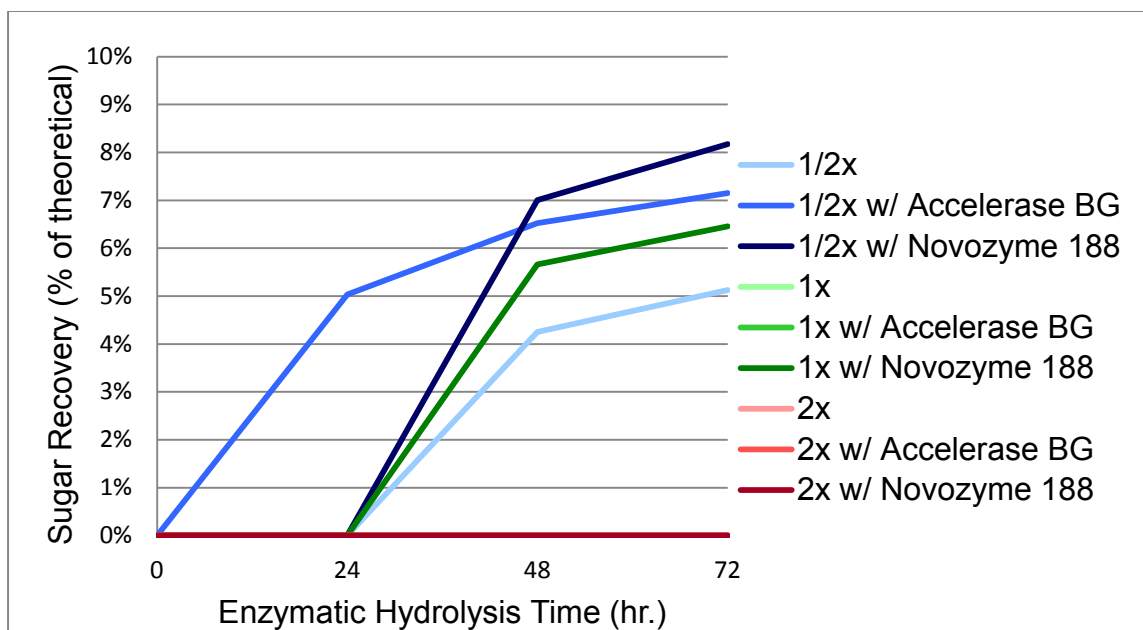
Figure B.23: Enzymatic hydrolysis xylose monomer recoveries for the aspen (2) experiment, performed 12-Dec. 2010.

Appendix C: Elsevier Authors' Rights and Responsibilities

This copyright agreement applies to the material in Chapter 3, which was previously published in the journal *Bioresource Technology* (2012) 116:320-326.

Copyright
Intellectual property, in particular copyright (rights in editorial content), trademarks (rights in brands for services or journals), and database rights (rights in compilations of information), form the foundation of Elsevier's publishing services and communications businesses. We in Elsevier embrace the opportunities the digital environment offers for communication and access, while at the same time we recognize the new risks that this environment poses, that being the ease with which unauthorized copies can be made and distributed worldwide. ⇨ Download your practical guide to Elsevier's copyright policy.

Our objective
We aim to manage digital rights and brands amidst the structural changes that the "information society" represents, while at the same time recognizing the shared goals we have with our customers and authors. These include providing the widest possible distribution of scientific and medical content and services in a financially sustainable business model.

Elsevier wants to ensure a proper balance between the scholarly rights which authors retain (or are granted/transferred back in some cases) and the rights granted to Elsevier that are necessary to support our mix of business models. We routinely analyse and modify our policies to ensure we are responding to authors' needs and concerns, and to the concerns in general of the research and scholarly communities.

What rights do I retain as a journal author*?
- the right to make copies (print or electronic) of the journal article for your own personal use, including for your own classroom teaching use;
- the right to make copies and distribute copies of the journal article (including via e-mail) to research colleagues, for personal use by such colleagues for scholarly purposes*;
- the right to post a pre-print version of the journal article on Internet websites including electronic pre-print servers, and to retain indefinitely such version on such servers or sites for scholarly purposes* (with some exceptions such as The Lancet and Cell Press. See also our information on ⇨electronic preprints for a more detailed discussion on these points)*;
- the right to post a revised personal version of the text of the final journal article (to reflect changes made in the peer review process) on your personal or institutional website or server for scholarly purposes*, incorporating the complete citation and with a link to the Digital Object Identifier (DOI) of the article (but not in subject-oriented or centralized

repositories or institutional repositories with mandates for systematic postings unless there is a specific agreement with the publisher. ⤷Click here for further information);

- the right to present the journal article at a meeting or conference and to distribute copies of such paper or article to the delegates attending the meeting;
- for your employer, if the journal article is a 'work for hire', made within the scope of the author's employment, the right to use all or part of the information in (any version of) the journal article for other intra-company use (e.g. training);
- patent and trademark rights and rights to any process or procedure described in the journal article;
- the right to include the journal article, in full or in part, in a thesis or dissertation;
- the right to use the journal article or any part thereof in a printed compilation of your works, such as collected writings or lecture notes (subsequent to publication of the article in the journal); and
- the right to prepare other derivative works, to extend the journal article into book-length form, or to otherwise re-use portions or excerpts in other works, with full acknowledgement of its original publication in the journal.

*Commercial purposes and systematic distribution
Authors of Elsevier-published articles may use them only for scholarly purposes as set out above and may not use or post them for commercial purposes or under policies or other mechanisms designed to aggregate and openly disseminate manuscripts or articles or to substitute for journal-provided services. This includes the use or posting of articles for commercial gain or to substitute for the services provided directly by the journal including the posting by companies of their employee-authored works for use by customers of such companies (e.g. pharmaceutical companies and physician-prescribers); commercial exploitation such as directly associating advertising with such postings; the charging of fees for document delivery or access; the systematic distribution to others via e-mail lists or list servers (to parties other than known colleagues), whether for a fee or for free; the posting of links to sponsored articles by commercial third parties including pharmaceutical companies; institutional, funding body or government manuscript posting policies or mandates that aim to aggregate and openly distribute the accepted, peer reviewed manuscripts or published journal articles authored by its researchers or funded researchers; and subject repositories that aim to aggregate and openly distribute accepted peer reviewed manuscripts or published journal articles authored by researchers in specific subject areas.

For a more detailed discussion of our article posting policies and the different stages of a journal article development that are relevant from a policy perspective, please see the ⤷ Article Posting Policies information page.