BIOCHEMICAL CHARACTERIZATION OF β-XYLAN ACTING GLYCOSIDE HYDROLASES FROM THE THERMOPHILIC BACTERIUM *Caldicellulosiruptor saccharolyticus*

Thesis
Submitted to
The School of Engineering of the
UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree of
Master of Science in Chemical Engineering

By
Jin Cao
Dayton, Ohio
December, 2012
BIOCHEMICAL CHARACTERIZATION OF β-XYLAN ACTING GLYCOSIDE HYDROLASES FROM THE THERMOPHILIC BACTERIUM *Caldicellulosiruptor saccharolyticus*

Name: Cao, Jin

APPROVED BY:

Donald A. Comfort, Ph.D.  
Advisory Committee Chairman  
Research Advisor & Assistant Professor  
Department of  
Chemical & Materials Engineering

Amy Ciric, Ph.D.  
Committee Member  
Lecturer  
Department of  
Chemical & Materials Engineering

Karolyn Hansen, Ph.D.  
Committee Member  
Assistant Professor  
Department of Biology

John G. Weber, Ph.D.  
Associate Dean  
School of Engineering

Tony E. Saliba, Ph.D.  
Dean, School of Engineering  
& Wilke Distinguished Professor
ABSTRACT

BIOCHEMICAL CHARACTERIZATION OF B-XYLAN ACTING GLYCOSIDE HYDROLASES FROM THE THERMOPHILIC BACTERIUM *Caldicellulosiruptor saccharolyticus*

Name: Cao, Jin
University of Dayton

Research Advisor: Donald A. Comfort

Fossil fuels have been the dominant source for energy around the world since the industrial revolution, however, with the increasing demand for energy and decreasing fossil fuel reserves alternative energy sources must be exploited. Bio-ethanol is a promising prospect for an alternative energy source to petroleum, especially when plant biomass is used as the replacement carbon source. This gives greater benefit for implementation of bioethanol as a viable alternative transport fuel than food stocks such as starch from corn. The implementation of this next generation of bioethanol requires more robust and environmentally friendly methods for degrading cellulose and hemicellulose, the two major components of plant biomass. This work looks to understand the role and mechanism of xylan degradation by glycoside hydrolases from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. The degradation of xylan is achieved through multiple glycoside hydrolase enzymes working synergistically.
Among these genes is the β-xlyosidase from *C. saccharolyticus*, *XynD* (Csac_2409), which was characterized to determine the optimal temperature, pH value, kinetic properties, hydrolysis pattern of oligosaccharides, and effects of inhibition by xylose. This biochemical characterization of the enzyme provides insight into its role in hydrolyzing xylan oligosaccharides by *C. saccharolyticus* and also potential biotechnological applications of the enzyme for upstream processing of hemicellulose for bioenergy application.
ACKNOWLEDGEMENTS

I sincerely thank Dr. Donald Comfort, my advisor, for all the help he has given to me since I started my study in UD. He provided a great stage to let me show my thinking and exchange ideas with him. His patient and open-minding make me never feel tough for my research. I have obtained a great deal of encouragement and approval from him that makes me feel confident to overcome the difficulties and burdens in research.

I would like to thank all faculty and staff from Chemical Engineering Department. Your wonderful classes give me the opportunities to further understand what an engineer truly is. Thanks my defense committee, Dr. Ciric and Dr. Hansen for their direction and invaluable advice along this project.

Thank you to all my friends, without your supports, either from the academic or daily life, I could not stand here today.

This is thesis dedicated to my parents and bother. Your love is my power to face all the difficulties.
TABLE OF CONTENTS

ABSTRACT.......................................................................................................................... iii
ACKNOWLEDGEMENTS......................................................................................................... v
TABLE OF CONTENTS.......................................................................................................... vi
LIST OF FIGURES ................................................................................................................ viii
LIST OF TABLES .................................................................................................................. x
CHAPTER 1 LIGNOCELLULOSIC BIOMASS FOR BIOETHANOL PRODUCTION
................................................................................................................................................ 1
   Introduction .......................................................................................................................... 1
   Bioethanol for Fuel............................................................................................................... 4
   Bioethanol Production....................................................................................................... 7
      Pretreatment and separation of lignin .......................................................................... 8
      Hydrolyzing lignocellulosic biomass .......................................................................... 10
      Caldicellulosiruptor saccharolyticus ......................................................................... 15
   Research Goals................................................................................................................. 16
CHAPTER 2 BIOCHEMICAL CHARACTERIZATION OF β-XYLODIDASE FROM
THE THERMOPHILIC BACTERIUM Caldicellulosiruptor saccharolyticus ............. 17

Introduction ........................................................................................................... 17

Materials and Methods ........................................................................................ 19

PCR, cloning and expression .............................................................................. 19

Purification of recombinant xylan 1,4-β-xylosidase XynD ............................... 21

Determination of the molecular weight of xylan 1,4-β-xylosidase .................... 22

Enzyme assay and kinetics characterization using synthetic substrates .......... 22

Hydrolysis of Natural Substrates and Thin-Layer chromatography (TLC) ......... 23

Results ................................................................................................................. 24

CHAPTER 3 CONCLUSION ..................................................................................... 37

Conclusion ........................................................................................................... 37

Future Work ......................................................................................................... 38

REFERENCES ..................................................................................................... 40

APPENDIX A PRIMER SEQUENCE AND RELEVANT PCR MATERIALS .......... 52

APPENDIX B LIQUID AND SOLID MEDIA ....................................................... 54

APPENDIX C LIGATION ....................................................................................... 57
LIST OF FIGURES

Figure 1: Lignocellulosic Ethanol Pathway ................................................................. 8
Figure 2: Cellulose hydrolysis procedure by three major types of cellulases .......... 12
Figure 3: General structure of xylan ............................................................................. 14
Figure 4: Structure of xylo-oligosaccharide ................................................................. 14
Figure 5: Procedure for gene cloning ......................................................................... 21
Figure 6: SDS-PAGE analysis of purified XynD from C. saccharolyticus ............... 24
Figure 7: a) Lineweaver-Burk plot for XynD incubated with substrate pNP-β-D-xylopyranoside; b) Plot of experimental initial velocity versus substrate concentration.. 26
Figure 8: a) Lineweaver-Burk inhibition plot of xylose on XynD, and b) Michaelis-Menten plot of the same data showing the increase in the corresponding half-maximal. 29
Figure 9: Thin layer chromatography of the products formed from the hydrolysis of xylo-oligosaccharides incubated with XynD ......................................................... 31
Figure 10: Estimated quaternary structure 3D model of the XynD, from C. saccharolyticus. ......................................................................................................................... 33
Figure 11: Amino acid sequence XynD, from *C. saccharolyticus*, aligned with protein XynB, from *T. saccharolyticum*, XynB, from *G. stearothermophilus* and XynB2, from *C. crescentus* ........................................................................................................................................... 34

Figure 12: a) (β/α)8- barrel fold 3D secondary structure of XynD, from *C. saccharolyticus* and the active and b) overlayed 3D structure for Csac_XynD, *T. saccharolyticum*, XynB, *G. stearothermophilus* XynB2, and *C. crescentus* XynB2 ........................................................................................................................................... 36
LIST OF TABLES

Table 1: The composition of potential lignocellulosic materials................................. 6
Table 2: Heating cycle parameters for PCR ........................................................................ 20
Table 3: Normalized activity of XynD for 20 & 90 minutes of incubation at 70°C .......... 25
Table 4: Variety of $K_M$ from different bacterium that had the combined features from GH 39 family and EC 3.2.1.37, with substrate of $p$NPβXyl ...................................................... 27
Table 5: Competitive Inhibitor xylose acted different $K_i$ in variety of enzymes .......... 30

Table A 1: dNTPs recipe................................................................................................. 52
Table A 2: PCR primer sequences .................................................................................. 52
Table A 3: Primer %GC and melting temperature......................................................... 53
Table A 4: Restriction Enzyme cutting position.............................................................. 53

Table B 1: LB media and LB medium plate ..................................................................... 54
Table B 2 SDS-Page Gel Protocol .................................................................................. 55
Table B 3: 5 times TBE recipe ...................................................................................... 56

Table C 1: Ligation mixture .......................................................................................... 58
CHAPTER 1
LIGNOCELLULOSIC BIOMASS FOR BIOETHANOL PRODUCTION

Introduction

Since the start of industrial revolution two centuries ago, the increase in energy consumption has been met by fossil fuels. They are so prominent that currently about 81% of the world’s energy is supplied by fossil fuels (IEA, 2011). The rate of energy consumption has increased dramatically, doubling in only thirty years from 1970 to 2000, and is expected to double again by 2050. Fossil fuels which have supplied this, however, are a nonrenewable source and cannot provide continual and sustained growth of energy demand. Furthermore, the emissions from burning fossil fuels include carbides, sulfides, nitrides, hydrocarbon compounds and suspended particles, among which carbon dioxide (CO$_2$), methane (CH$_4$) and nitrous oxide (N$_2$O) are the first three main greenhouse gases responsible for global warming and climate changing (Cheng, 2010). Of these, CO$_2$ is by far the largest by volume, representing about 80% of the total emissions of these six gases (Ahmad & Wyckoff, 2003). Global energy-related CO$_2$ emissions will rise from 30.2 billion tons in 2008 to 35.2 billion tons in 2020 and 43.2 billion tons in 2035 (IEA, 2011). The concentration is approximately 390 ppm (parts per million) in 2010, which is about
39% higher than in the mid-1800s (280ppm) (IEA, 2011). In order to support continued energy growth, an alternative and renewable source of energy must be found.

Bioethanol is a potential renewable alternative energy that can be used as a transportation fuel in place of gasoline. Bioethanol is produced by the fermentation of simple sugars such as glucose into ethanol and CO₂. The glucose comes from carbohydrate resources such as corn or cane sugar or more complex structural carbohydrates such as lignocellulosic biomass (McGowan, et al., 2009; Taherzadeh & Karimi, 2008). Lignocellulose is the largest source of carbon on the planet and primarily consists of cellulose, hemicellulose, and lignin. Bioethanol is currently being produced from starchy crops, but this method creates competition for food resources driving up prices. Lignocellulosic material, on the other hand, generally has little intrinsic value and represents a largely untapped carbohydrate resource.

Biofuels have the capacity to solve both fossil fuel and energy independence/national security reliance on petroleum. The European Union (EU), for instance, completely relies on imported petroleum to support their transportation sector (Galbe & Zacchi, 2007), but in 2008 produced 8.5 billion liters of biodiesel, a 50% share of the global market (Eisentraut, 2010). Biofuel production is also growing, with 105 billion liters of biofuel was produced around the world in 2010, which was 17% higher than in 2009 (International Energy Agency (IEA)).

One reason that lignocellulosic bioethanol has lagged behind starch sources for conversion into bioethanol is the difficulty of hydrolyzing the material into the sugar monomers. Currently hydrolysis is carried out by acid-based processes, however, this technique causes corrosion and formed inhibitor during acid hydrolysis which leads to
yields that are less than satisfactory (Taherzadeh & Karimi, 2008). A new bioengineering technique has been investigated utilizing enzymes to degrade cellulose and hemicellulose into hexoses and pentoses. This enzymatic hydrolysis method is more selective at degrading polysaccharides, which leads to fewer undesirable side reactions. Moreover, this method is conducted at milder conditions (temperature and pH) compared to the acid hydrolysis (Sun & Cheng, 2002). Given these characteristics, the enzymatic method could give a more economical and effective way to hydrolyze lignocellulosic biomass and produce ethanol.

We are investigating the potential hydrolysis of hemicellulose using glycoside hydrolase enzymes from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*, which has the capacity for metabolizing a large variety of carbohydrates, including hemicellulose. Moreover, the enzymes from *C. saccharolyticus* are thermophilic but stable at temperatures typically up to 70-80°C (Coughlan, 1989). The genome sequence of *C. saccharolyticus* shows a number of enzymes hypothesized to be involved in the breakdown of hemicellulose (van de Werken, et al., 2008). The GH-based genomic loci Csac_2404 to Csac_2411, including XynD (Csac_2409), which appear to be central to the hydrolysis of hemicellulosic substrates, are examined in this research (VanFossen, et al., 2011; van de Werken, et al., 2008). Csac_2408 & Csac_2405 have previously been produced recombinantly and encoded intracellular and extracellular GHs that were active towards xylo-oligosaccharides. Strategic utilization of these enzymes may provide a new platform to produce ethanol utilizing renewable resources of hemicellulose.
**Bioethanol for Fuel**

Ethanol is the most important product in biotechnology in terms of fuel market value (Taherzadeh & Karimi, 2008). It can be used as a motor fuel in its pure form, but usually used as a gasoline additive to reduce vehicle emissions, since it has a high octane rating (113 octane). With a high octane rating, the fuel could have high performance in high-compression engines. Different ratios blended ethanol-gasoline could efficiently raise the compression ratio and increase thermal efficiency. For instance, E10 fuel (E, fuels denote ethanol blended fuels), a mixture of 10% ethanol and 90% gasoline, could raise the fuel's octane rating 2 or 3 that is higher than regular gasoline (www.ethanol.org, 2011). In addition, ethanol is a clean, non-toxic environmental friendly fuel, which could reduce the growth of GHG emissions into the atmosphere from burning fossil fuel. Until 2003, ethanol production around world reached to 3.1 billion liters. 60% of the total production was derived from sugar crops, 33% was derived from other crops and the rest from synthetic means (Hamelinck, et al., 2003). Bioethanol (produced from agricultural feedstocks) has been the primary source for ethanol consumption. However, the cost of bioethanol as an energy source is relatively high compared to fossil fuels. Additionally, starch-based technology leads to competition for agricultural resources.

Biofuels are normally grouped into two different categories based on the type of feedstock they utilize. The first generation fuels utilize feedstocks that are easy to process, including sugar, starches, and vegetable oils. The second generation fuels utilize harder to process feedstocks, most notably lignocellulosic biomass.

The 1st generation biofuels that are derived from starch sugar and vegetable oils or animal fat and some other organic materials. Corn ethanol in the U.S., sugarcane ethanol
in Brazil, oilseed rape biodiesel in European Union countries and other biofuels elsewhere, have been produced to meet the increasing total global demand of energy (IEA, 2009). The production of 1st generation biofuels, mainly from traditional food crops, has increased rapidly over the past few years in response to concerns about energy supply security, rising oil price, climate change, and development of a commercial market as well. However, their potential to be the primary fuels is still limited by various constraints: competition for land, water and other resources in agricultural production; contribution of high food prices; and high costs in the production process. The most important drawback is the net of GHG emission is not reduced distinctly. Considering current situation, a new generation of biofuels, with high performance and conventional production techniques, is required.

The second generation of biofuels utilizes non-food stocks such as lignocellulosic material like straw, forest residues, grass, and organic component of municipal solid wastes (Sims, et al., 2010; Cheng, 2010). Another attractive feature is that lignocellulosic biomass is the largest carbon reserve in the world and has a short harvest life cycle. In fact, abundance, the short harvest life cycle, and no competition for land makes lignocellulosic biomass an ideal source of energy for reliable and sustainable energy (Kim & Dale, 2004; Milbrandt, 2005).

Hemicellulose, cellulose and lignin are the major components of the lignocellulosic biomass. Furthermore, cellulose and hemicellulose are the typical polysaccharides materials which can be converted into monosaccharide through a series of thermochemical or biological processes, and finally converted to ethanol through
fermentation process. Some common lignocellulosic materials and their compositions are listed in Table 1.

| Table 1: The composition of potential lignocellulosic materials  
<table>
<thead>
<tr>
<th>(Jørgensen, et al., 2007; Cheng, 2010)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composition</td>
</tr>
<tr>
<td>Corn stover</td>
</tr>
<tr>
<td>Sugarcane bagasse</td>
</tr>
<tr>
<td>Switchgrass</td>
</tr>
<tr>
<td>Wheat straw</td>
</tr>
<tr>
<td>Leaves</td>
</tr>
<tr>
<td>Hard wood stalk</td>
</tr>
<tr>
<td>Softwood stalk</td>
</tr>
<tr>
<td>Waste papers from chemical pulps</td>
</tr>
<tr>
<td>Hybrid poplar</td>
</tr>
</tbody>
</table>

Cellulose is a linear, homogenous polysaccharide which is comprised of over 10,000 glucose units (Cheng, 2010). It occupies 40 to 60% mass in dry biomass of lignocellulosic materials, so, it is a potentially important source for making ethanol. One of the major challenges with cellulose conversion is the degradation of the polymeric chains. Hydrogen bonds link the cellulose polymers together (Taherzadeh & Karimi, 2008), making them hard to separate and isolate for enzymatic degradation into free sugars. Finding an economic and effective pretreatment method is the key to produce bioethanol in commercial scale.

Hemicellulose constitutes about 20 to 40% of the dry biomass in lignocellulosic material. Hemicellulose differs from cellulose in that it is heterogeneous polymers made up of xylan, xyloglucan, galactoglucomannan, arabinoxylan, arabinogalactan and glucomannan (Collins, et al., 2005). These polysaccharides are constituted by the sugar monomers, like L-arabinose, D-galactose, D-xylose, D-glucose and D-mannose and some sugar derivatives such as D-glucuronic acid, D-4-O-methylglucuronic acid, D-
galacturonic acid, O-acetyl groups or feruloyl-and coumaryl esters linked via L-arabinose residues to the backbone (Aspinall, 1959). These sugars allow the branched hemicellulose to be easily hydrolyzed.

The third major component of lignocellulosic biomass is lignin, which comprises 10 to 25% of the biomass. Lignin is a complex polymer that is a key factor for cell walls (Hamelinck, et al., 2003). Cellulose and hemicellulose are strongly linked via lignin by the covalent and hydrogen bonds. This structure drives the lignocellulosic biomass to resist any force from breaking (Limayem & Ricke, 2012). Lignin cannot be hydrolyzed to fuel components due to its phenyl propane and methoxy groups. Moreover, the interpolymer bonding between the cellulosics and hemicellulosics arising from lignin representing a significant reduction in processing efficiency. Although lignin cannot be the source of fermentation, it can be burnt for heat or electricity that is one way to increases the energy efficiency for the entire system (Galbe & Zacchi, 2007).

Besides, lignocellulosic feedstock materials are CO₂ neutral conversion, its conversion is an important derivation of biofuels. 2nd generation biofuel solved two major problems which limited the development of first generation biofuels, feedstock source and carbon neutral fuels.

**Bioethanol Production**

Commercial-scale production of bioethanol has increased dramatically in the past decade, but the approach has not evolved much. The key conversion steps of ethanol production are shown in Figure 1 and include: pretreatment and separation of lignin residues, hydrolysis of hemicellulose and cellulose, fermentation of sugar, and ethanol purification (Galbe & Zacchi, 2007; Hamelinck, et al., 2003; Naik, et al., 2010). Current
challenges in the lignocellulosic processing for ethanol are at the pretreatment step to remove lignin and during enzymatic hydrolysis of cellulose.

**Figure 1**: Lignocellulosic Ethanol Pathway (Galbe & Zacchi, 2007; Naik, et al., 2010)

**Pretreatment and separation of lignin**

Pretreatment of lignocellulosic materials is to disrupt the crystalline structure of macro- and micro-fibrils, and release the polymer chains of cellulose and hemicellulose (Mosier, et al., 2005). This results in the interpolymer spacing between the materials to be expanded allowing the enzymes to penetrate into the fibers. This step is important for practical cellulose and hemicellulose conversion processes (Mosier, et al., 2005), which can further enhance the ethanol production (Galbe & Zacchi, 2007). Technologies of pretreatment based on different properties of each material and are be classified into three categories: physical, physicochemical & chemical, and biological pretreatment.

Physical pretreatment is a method to break the structures of biomass into fine powder, which increases the surface area for faster enzymatic treatment. Chipping, comminution, stream explosion and hot water washing are the common methods of physical pretreatment, however, the low sugar yields of sugar and high costs due to intensive energy requirement make this technology less than ideal for scale-up.
Two dominant chemical pretreatment methods for biomass are alkali and dilute acid treatments (Galbe & Zacchi, 2007). Materials are immersed in dilute sulfuric acid solution (below 4 wt %, temperature above 160°C) and heated from several minutes up to hours (Galbe & Zacchi, 2007; Sun & Cheng, 2005). This method is the most developed for industrial usage since it is low-cost and high yield, however, the maintenance cost on the anti-corrosion and toxic hydrolysate are the major drawbacks for this method. Alkaline treatment method (using lime or NaOH) is more effective on agricultural residues and herbaceous crops comprised of less lignin (Khanal & Surampalli, 2010). The solubilized cellulosic materials are dissolved into the liquid phase, and then available for hydrolysis by enzymes.

Physicochemical methods seek a way to combine the benefits both from physical and chemical pretreatments. Steam pretreatment and ammonia fiber explosion (AFEX) are two of the most widely used physicochemical pretreatment methods. Steam pretreatment has shown good results for hydrolysis of softwood (Galbe & Zacchi, 2007), whereas AFEX has a good performance on agricultural residues, but is not effective on high lignin content materials like wood.

Emerging biological techniques are promising innovations for pretreatment approaches, due to their environmentally friendly nature, less energy and chemicals, and compact reactor design (Khanal & Surampalli, 2010). However, these have not been widely used the reaction rate of biological pretreatment processes is far too low. Even still, this method is continuing to be investigated at the research stage, and has high potential to replace the traditional methods and become a major method in the future.
**Hydrolyzing lignocellulosic biomass**

Hydrolysis is the major method to degrade polysaccharides into sugar monomers. Complete hydrolysis of cellulose results in glucose, whereas hemicellulose results in various pentoses and hexoses (primarily xylose and glucose). Hydrolysis can be catalyzed by concentrated acid, dilute acid, or enzymes.

Concentrated acid generally operates in the acid concentration ranges from 30-70% and at low/medium temperature (around 40°C). This method has been reported to hydrolyze of both hemicelluloses and cellulose and produce high sugar yield. On the other hand, dilution and heating of the concentrated acids would result in extremely corrosive conditions and the acid recovery is mandatory for its economic viability, as well (Arastehnnodeh, 2012). Dilute-acid hydrolysis is the most commonly applied among the chemical hydrolysis methods. Both cellulose and hemicellulose can be hydrolyzed using this process. Dilute-acid hydrolysis is carried out in two stages. In the first stage, which is carried out under relatively mild conditions, hemicellulose is converted to sugar monomers. In the second stage, the residual solids are hydrolyzed under relatively high temperature about 230-240°C (Wyman, 1999), allowing cellulose to be hydrolyzed. However, dilute-acid hydrolysis is not a preferable pretreatment for cellulose hydrolysis since the high temperatures required may lead to formation of a high amount of degradation products (Gírio, et al., 2010). In addition, for both dilute and concentrate acid hydrolysis approaches, the acid have to be removed or neutralized before fermentation, which yields a large amount of waste.

Current researches have proven that a number of thermophiles can hydrolyze lignocellulosic biomass and converse them into bioenergy (Blumer-Schuette, et al., 2008).
Most lignin is removed from the biomass and the crystallinity of the biomass is significantly reduced by the chemical pretreatment. The expanded interspace allows enzymes to enter into the biomass and access the substrates. In general, biomass are catalyzed by cellulolytic and hemicellulolytic enzymes in mild conditions, which results low corrosion problems and low utility consumption. Moreover, the low toxicity of the hydrolysate is also the main advantages of this process.

Complete hydrolysis of cellulose and hemicellulose requires a suite of enzymes to process each of the different carbohydrates and the resulting intermediate hydrolysis products. Endoglucanase, exoglucanase and β-glucosidase are the three major groups of cellulose enzymes, which accomplish hydrolysis by synergistic performance (Wyman, 1996). The combined hydrolysis by them results in the conversion of cellulose into glucose (Figure 2).
Endoglucanases internally cut the cellulose chains, thereby producing shorter cello-oligomers, also called cello-dextrins. Endoglucanases attacks regions of low crystallinity of the cellulose fiber, creating free chain-ends (Sun & Cheng, 2002). Exoglucanases, which include celllobiosyl hydrolases and β-glucosidases, cleave oligosaccharides from the outside at the non-reducing ends of cellulose chains and cello-oligosaccharides (Khanal & Surampalli, 2010). Celllobiosyl hydrolases produce
disaccharide cellobiose products and β-glucosidases produce monosaccharide glucoses. (Sun & Cheng, 2002).

Bioconversion of the hemicelluloses into fermentable sugar monomers is essential to reduce the cost of bioethanol in commercial markets (Wyman, 1999). Unlike cellulose, hemicellulose is a heteropolysaccharide, which requires the synergistic reaction of a number of related enzymes. Great variety of hemicelluloses depends on the various backbone-chain compositions of building blocks. Four main categories include xylans, mannans, β-glucans, and xyloglucans.

Xylans are the most abundant type of hemicellulose and are organized by β-1,4-linked β-D-xylopyranosyl units with side-chain groups arabinofuranose, glucuronic acid, or 4-O-methyl-D-glucuronic acid (Ek, et al., 2009) (Figure 3), depends on the different kinds of plant species. Comprised of pentose and hexose, xylans are the primary source to be decomposed and fermented into ethanol.

The complex structure of xylan requires cooperation from a suite of different enzymes to fully hydrolyze hemicellulose. This group of enzymes is called xylanases, which is one kind of glycoside hydrolases (EC 3.2.1-x) (http://www.cazy.org) with endo-hydrolysis to cleave the 1,4-β-D-xylosidic linkages in xylan. α-L-arabinofuranosidases (EC 3.2.1.55), α-D-glucuronidases (EC 3.2.1.139), acetylxylan esterases (EC 3.1.1.72), ferulic acid esterases (EC 3.1.1.73) and p-coumaric acid esterases (EC 3.1.1.--) are used to remove the side groups.
**Figure 3**: General structure of xylan (Sunna & Antranikian, 1997). Figure was prepared with CS ChemBioDraw Ultra version 12.0.

Endo-1,4-β-xylanases internally cleave the glycosidic bonds in the xylan backbones, bringing a reduction of degree of polymerization (Figure 4.). The bonds selected for hydrolysis depend on the nature of substrate. The main products are β-D-xylopyranosyl oligomers. But at the later stage, β-D-xylopyranosyl oligomers also can be hydrolyzed to produce smaller molecules, such as mono-, di- and trisaccharides of β-D-xylopyranosyl (Polizeli, et al., 2005). Then, β-xylosidases hydrolyses smaller xylose and xylobiose from the non-reducing terminus. Usually purified β-xylosidases do not directly hydrolyze xylan.

**Figure 4**: Structure of xylo-oligosaccharide. The hydrolysis, xylose, could be the source for bioethanol production. (Cheng, 2010; Collins, et al., 2005). Figure was prepared with CS ChemBioDraw Ultra version 12.0.

Endo-1,4-β-xylanases and β-xylosidases can be produced by fungi, bacteria, yeast, marine algae, insect, seeds and etc. (Sunna & Antranikian, 1997). Several investigations indicate that xylanases are usually highly expressed when the organism is grown on xylan
compounds (Beg, et al., 2001) and different carbon sources have been analyzed to determine their role in effecting the enzymatic hydrolysis.

Another main constituent of hemicellulose is galactoglucomannan, a mixture of glucose and mannose. Endo-1,4-β-mannanases and β-mannosidases act as hydrolytic enzymes to cleave the polymer backbone, in order to breakdown mannose into simple sugars. These mannanases can be found in various species, such as bacteria, yeasts and fungi (Gírio, et al., 2010). Although a number of mannanase producing bacterial sources are available, only a few can be commercially used, such as Bacillus sp. and Streptomyces sp. (Gírio, et al., 2010). The studies on enzymes that hydrolyze mannan have been largely neglected even through glucomannans is an abundant hemicellulose. Therefore, the application of mannanases is as important as xylanases (Gírio, et al., 2010).

Taken together a suite of enzymes that could hydrolyze the bonds of lignocellulosic feedstocks to generate glucose, xylose, and mannose and also allow for separation of lignin would greatly improve the sustainability and economics of lignocellulosic bioethanol.

*Caldicellulosiruptor saccharolyticus*

The enzymes used in these studies come from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *C. saccharolyticus* is a thermophilic, anaerobic bacterium, with an optimal growth temperature of 70°C. This bacterium has been valued for its ability to produce hydrogen from biomass-derived polymers, such as cellulose and hemicellulose (VanFossen, et al., 2008). Recent research shows that *C. saccharolyticus* hydrolyzes cellulose, hemicellulose, starch, and pectin, and ferments corresponding sugar monomers, like glucose, hexoses or pentoses (van de Werken, et al., 2008). To aid with
studies, the *C. saccharolyticus* genome has been sequenced and resulted in identification of the glycoside hydrolase enzymes. As a result of the genome sequencing, over 61 glycoside hydrolase genes have been identified, including those for degrading cellulose and hemicellulose. This makes *C. saccharolyticus* an attractive candidate for digestion of biomass-derived substrates to biofuels.

**Research Goals**

This research emphasizes the purification and characterization of glycoside hydrolase enzymes from the thermophile *C. saccharolyticus* with a focus on those enzymes that degrade xylan to generate xylose. Understanding these enzymes could lead to a new suite of enzymes that effectively generate monosaccharides which may reduce the operational costing and energy consumption of second generation lignocellulosic bioethanol, making it economically competitive. In these studies, the gene for β-xlyosidase was cloned, recombinantly expressed, purified, and characterized. The Michaelis-Menten characterization is completed β-xylosidase using synthetic and natural disaccharide substrates as well as determination of the hydrolysis patterns for larger xylooligosaccharides. Finally, enzyme inhibition by the hydrolysis product xylose is investigated to understand whether this enzyme will be efficient at process xylooligosaccharides to high conversion.
CHAPTER 2

BIOCHEMICAL CHARACTERIZATION OF β-XYLODIDASE FROM THE THERMOPHILIC BACTERIUM *Caldicellulosiruptor saccharolyticus*

Introduction

Hemicellulose is the second most abundant renewable resource in the world, behind only cellulose. This heterogeneous polymer is comprised of four main categories of saccharides: xylans, mannans, mixed linkage β-glucans and xyloglucans (Ebringerová, et al., 2005). Among these, xylans are the most abundant type of hemicellulosic polysaccharide in cell walls, representing up to 35% of the total dry weight (Knob, et al., 2010). Xylan are comprised of β-1,4-linked D-xylopyranosyl units with short carbohydrate side-groups (Ek, et al., 2009). Hemicellulose has potential to play a key role as a carbon source for renewable energy, but this requires the complete degradation of xylan by the synergistic action of enzymes to make this a viable energy source (Collins, et al., 2005). Endo-β-1,4-xylanase (EC 3.2.1.8) acts on xylans and xylooligosaccharides, primarily producing a mixture of xylo-oligosaccharides; β -xylosidase (EC 3.2.1.37) further hydrolyzes the xylo-oligosaccharides into xylose (Grange, et al., 2001), and several auxiliary enzymes, such as α-L arabinofuranosidase (EC 3.2.1.55), β -glucuronidase (EC 3.2.1.139), acetylxylan esterase (EC 3.1.1.72),
ferulic acid esterase (EC 3.1.1.73) and \( p \)-coumaric acid esterase (EC 3.1.1-), cleave the side-chain (Collins, et al., 2005; Subramaniyan & Prema, 2001; Saha, 2003) (Figure 3).

\( \beta \)-xylosidases acts at the non-reducing ends of xylo-oligosaccharides to release xylose through the catalytic hydrolysis of xylobiose and other xylooligosaccharides (Sørensen, et al., 2003). In general, due to the length of xylan, \( \beta \)-xylosidases are not capable of directly degrading xylan (Lee, et al., 1993), \( \beta \)-xylosidases plays an important role in xylan degradation by releasing the end product that limits the hydrolysis activity of xylanase (Knob, et al., 2010). For these reasons, it showed high specificity for xylo-oligosaccharides and attracted the attention of the pulp and paper industry (Yang, et al., 2004).

\( \beta \)-xylosidases have been characterized from a variety of fungi and bacteria, much of which also have been studied to understand the capability for processing and degrading hemicellulosic material. For instance, the thermophilic Thermomonospor fusca (Bachmann & McCARTHY, 1989), the fungi Penicillium capsulatum and Talaromyces emersonii (Filho, et al., 1991), the hyperthermophilic bacterium Caldicellulosiruptor saccharolyticus (VanFossen, et al., 2011) have had \( \beta \)-xylosidase enzymes characterized from them. \textit{C. saccharolyticus} is a thermophilic bacterium that grows optimally at 70°C and was isolated from a thermal spring in New Zealand. The genome has been sequenced and showed a large number of carbohydrate active enzymes, including 53 glycoside hydrolases (VanFossen, et al., 2009). A number of these glycoside hydrolases have been characterized from \textit{C. saccharolyticus}, including Csac_0678 (Ozdemir, et al., 2012), CelA (Csac_1076, GH 48 &9) (Te'o, et al., 1995), CelB (Csac_1078, GH 10 & 5) (VanFossen, et al., 2011), and BgLA (Csac_1089, GH 1) (Hong, et al., 2009) to name a
few. On advantage to using thermophilic proteins for processing carbohydrates is the greater solubility of carbohydrates at elevated temperatures and reduced risk of contamination. This makes thermophilic enzymes an ideal match for carbohydrate processing and led to further investigation of the enzymes from \textit{C. saccharolyticus}.

This work investigates carbohydrate active enzymes from \textit{C. saccharolyticus} identify from the genome sequence, with a focus for those enzymes capable of degrading hemicellulose. A number of genes were successfully cloned (Csac\_2405, Csac\_2408, Csac\_2409, and Csac\_2410), however, only Csac\_2409 (\textit{XynD}, xylan 1,4-\beta-D-xylosidases) was successfully expressed in \textit{E. coli}. In this study, recombinant \textit{XynD} was purified, biochemically characterized for substrate specificity, pH and temperature optima, inhibition effects, and hydrolysis patterns of xylooligosaccharides determined to understand how this enzyme functions in the organism and how it might be used for biotechnological applications such as hemicellulose degradation.

\textbf{Materials and Methods}

\textit{PCR, cloning and expression}

Genomic DNA was extracted from \textit{C. saccharolyticus} using a Qiagen DNeasy Blood and Tissue Kit (Valencia, CA) using the recommended protocol by the vendor. The DNA segment encoding the \textit{XynD} (Csac\_2409) gene was amplified by PCR using the extracted genomic DNA under the conditions shown in Table 2. Reverse and forward primers (shown below, respectively) contained the unique restriction enzyme cleavage site of \textit{NcoI} and \textit{XhoI} (bold).

Reserve primer: \texttt{5’} – CCC CCA \textbf{CTG GAG} ATA ACC TGG TAT TTT ACT ATC \texttt{3’} \textit{NcoI}

Forward primer: \texttt{5’} – TGT GTG \textbf{GCT AGC} ATA AAG ATT GAG AAA GGC \texttt{3’} \textit{XhoI}
Table 2: Heating cycle parameters for PCR

<table>
<thead>
<tr>
<th>Number of Cycles</th>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial Denaturation</td>
<td>95°C</td>
<td>2min</td>
</tr>
<tr>
<td>2</td>
<td>Denaturation</td>
<td>95°C</td>
<td>45sec</td>
</tr>
<tr>
<td>3</td>
<td>Annealing</td>
<td>44°C</td>
<td>45sec</td>
</tr>
<tr>
<td>4</td>
<td>Extension</td>
<td>72°C</td>
<td>135sec</td>
</tr>
<tr>
<td>5</td>
<td>Repeat to step 2 for 34 or more times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Final Extension</td>
<td>72°C</td>
<td>5min</td>
</tr>
<tr>
<td>7</td>
<td>Soak</td>
<td>10°C</td>
<td>Indefinite</td>
</tr>
</tbody>
</table>

The PCR products were purified using the Wizard SV Gel and PCR clean-up system (Promega, Madison, WI), then, imaged following gel electrophoresis to check purity and determine length of the DNA. Restriction enzyme, NcoI and XhoI, digested double stranded DNA fragment and pET21b vector (Novagen, San Diego, CA). Digestion was at 37°C for 2 hours, followed by purification by Wizard SV PCR kit, and then the digested gene was ligated into the digested pET21b. The constructed expression vector (Figure 5) was transformed into NovaBlue competent cells (San Diego, CA) which were grown at 37°C in LB medium with 100 µg/µL of Ampilillin. The sub-cloned construct was extract by DNA mini-prep and transformed into Rosetta competent cells for protein expression. During protein expression, Isopropyl-β-D-galactopyranoside (IPTG) was added to a final working concentration was 0.4 mM, after when the OD600 reached 0.6-0.8 to induce the lac operon. The cells were then further incubated at 30°C overnight.
Figure 5: Procedure for gene cloning (Wolfe, 1993). Production of recombinant DNA for cloning by restriction enzymes digestion

*Purification of recombinant xylan 1,4-β-xylosidase XynD*

Cells were harvested by centrifugation at 5,000x g at 4°C for 15 min and the resuspended in 20 mM Tris-HCl buffer (pH 8.0). The cells were lysed by sonication and the supernatant heat treated for 30 minutes at 60°C to denature heat-labile *E. coli* proteins. The heat-treated enzyme extract was applied to a 5 mL HiTrap IMAC HP column (GE Healthcare) and gradually eluted by 500 mM imidazole in 20 mM sodium phosphate, 0.5 M NaCl, pH 7.4 buffer.
**Determination of the molecular weight of xylan 1,4-β-xylosidase**

The purified of the xylan 1,4-β-xylosidase XynD was assayed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing conditions and the bands compared to the EZ-Run Rec pre-stained protein ladder (Fisher Scientific, Fair Lawn, New Jersey). All proteins bands were stained with Coomassie blue for visualization. The protein concentration was assayed using BioRad protein concentration assay (Biorad, Hercules, CA) with bovine serum albumin (NEB, Ipswich, MA) used as protein concentration standard.

**Enzyme assay and kinetics characterization using synthetic substrates**

XynD activity was measured using a number of synthetic substrates contain p-nitrophenol (pNP) including: pNP-β-xylopyranoside (pNPβXyl), pNP-β-glucopyranoside (pNPβGlc), pNP-α-glucopyranoside (pNPαGlc), and pNP-β-mannopyranoside (pNP-β-man). The extent of these reactions can be determined by measuring the absorbance of released pNP at 405 nm and comparing against a concentration standard curve for pNP. Optimal pH and temperature conditions were determined using pNPβXyl. Subsequent reactions were carried out at pH 6.5 and 80°C to determine the enzyme kinetic parameters $k_{cat}$, $V_{max}$, and $K_M$, and the enzyme reactions were performed with different substrate concentration range from 50 µM to 5,000 µM. These enzyme kinetic parameters, values were determined by fitting the data to Lineweaver-Burk Plot and the specific activity was defined as the amount of pNP as a product per milligram of enzyme amount per reaction time (µmol · min⁻¹ · µg$_{enzyme}$⁻¹).
Hydrolysis of Natural Substrates and Thin-Layer chromatography (TLC)

The degradation patterns of XynD were analyzed by thin-layer chromatography on silica gel plates, Silica Gel 60 F \textsubscript{254} TLC plate (Merck, Germany). 10 mM xylobiose, xylotriose, xylotetraose, and xylopentaose (Megazyme, Wicklow, Ireland) or 10 mM cellobiose were digested with XynD in 50 mM sodium phosphate buffer (pH 6.5) at 64°C. The incubation time lasted up to 30 hours. Digested samples were spotted onto plates, which were developed twice by n-butanol-ethanol-water (5:3:2 by volume) and sprayed with freshly prepared ethanol-concentrated H\textsubscript{2}SO\textsubscript{4} (9:1 by volume) mixture (Lee, et al., 2006). Dried at room temperature and heated the plates at 200°C for couple minutes until the sugar spots visualized.
Results

This 1521bp gene, as reported by GenBank, encoding XynD that was cloned and expressed in *E. coli*. Protein was expressed using IPTG to induce the lac operon and the resulting enzyme was extracted from the cells. The enzyme was heat treated to denature heat-labile *E. coli* proteins and the heat treated extract purified using a Hi-Trap metal-affinity chromatography with nickel. Purification of XynD was confirmed using SDS-PAGE to check both the purity of the enzyme and affirm that there was a single band with molecular mass about 59kDa (Figure 6), which matched the predicted mass of 58940.3 Da. The final concentration of the purified protein was determined to be 0.088 mg/mL as determined by Biorad Protein Assay.

![SDS-PAGE analysis](image)

**Figure 6**: SDS-PAGE analysis of purified XynD from *C. saccharolyticus*. Lane 1, prestained EZ-Run Rec protein ladder; Lane 2 crude extract enzyme before heat treatment; Lane 3 crude extract enzyme after heat treatment; Lane 4, purified XynD.

β-xylosidase was investigated and found to be capable of removing successive D-xylose residues from the non-reducing termini (Selig, et al., 2008). Substrate specificity
was determined using synthetic pNP-saccharide compounds, comprised of a monosaccharide and a p-nitrophenyl functional group that were connected via a glycosidic bond. β-Xylosidase hydrolyzed the β-1,4-linked glycosidic bond, as indicated by a bright yellow color resulting from the p-nitrophenol. XynD exhibited the greatest activity toward pNP-β-D-xylopyranoside and significant activity on pNP-β-D-glucopyranoside, but had essentially no detectable activity on pNP-α-D-glucopyranoside or pNP-β-D-mannopyranoside, the activity of these pNP substrates are shown in Table 3. This phenomena indicated that XynD had strong specific selectivity to catalyze the degradation of the glycosidic bonds with a terminal β-xylopyranosides from oligosaccharides.

| Table 3: Normalized activity of XynD for 20 & 90 minutes of incubation at 70°C. |
|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
|                               | pNPβXyl | pNPβGlc | pNPαGlc | pNPβman |
| 20 min                        | 100%    | 1.01%   | 0.12%   | 0.15%   |
| 90 min                        | 100%    | 1.90%   | 0.094%  | 0.11%   |

Michaelis-Menten kinetic parameters were investigated utilizing pNP-β-D-xylopyranoside in pH 6.5 sodium phosphate buffer, at 70°C. Substrate concentration varied from 50 µM to 5,000 µM (enzyme concentration: 880 ng per 1000µL). Michaelis-Menten kinetics model was used to determine enzyme catalyzing kinetics. The equation $V = \frac{V_{max}S}{K_M + S}$ shows the relationship between the substrate concentration and reaction velocity. Taking the reciprocal gives: $\frac{1}{V} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}}$. The Michaelis-Menten kinetic parameters, $V_{max}$ and $K_M$ could be determined from Lineweaver-Burk Plot.
Figure 7: a) Lineweaver-Burk plot for XynD incubated with substrate pNP-β-D-xylopyranoside; b) Plot of experimental initial velocity versus substrate concentration.

The $K_M$ is the substrate concentration at which the reaction rate is half of $V_{max}$. In this study, $K_M$ value was determined to be 2.68 mM; the initial velocity $V_{max}$ was 58.5 μM/min; the $k_{cat}$ was 0.13 s$^{-1}$, and $k_{cat} / K_M$ is 0.049 s$^{-1}$ mM$^{-1}$. The lower value of $K_M$ means that the substrate has greater affinity for the substrate so a lower concentration of substrate is required to saturate the enzyme, thus reducing the catalytic efficiency $k_{cat} / K_M$. The specific activity for XynD on pNPβXyl at pH 6.5 and 70°C was determined to be 46 μmol min$^{-1}$ mg$_{enzyme}^{-1}$.

Glycoside hydrolases are grouped into families based upon amino acid sequence similarities (Henrissat, 1991). XynD appears in glycoside hydrolase family (GH) 39, along with other β-xylosidases (which also appear in GH 3, 30, 39, 43 52 and 54) and α-L-iduronidases. Enzymes within a family share a common sequence and structural motifs, especially at the active sites, which allows for identification of the catalytically essential amino acids involved in the catalytic reaction through sequence alignment once those residues are identified for any of the proteins in the family. A comparison of the
biochemical properties of some of the other closely related GH 39 β-xylosidases is shown in Table 4, which demonstrates that even with similarities in structure, the properties of the enzyme can vary widely for temperature and pH optima as well as kinetic properties for $K_M$, which ranges from 0.036 to 28 mM on $p$NP-$β$-xylopyranoside.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Bacterium</th>
<th>Conditions</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>XynD</td>
<td><em>Caldicellulosiruptor saccharolyticus</em></td>
<td>pH 6.5, 70°C</td>
<td>2.68</td>
<td>0.13</td>
<td>0.049</td>
<td>This study</td>
</tr>
<tr>
<td>xynBl</td>
<td><em>Geobacillus stearothermophilus</em> T-6</td>
<td>pH 4.0, 40°C</td>
<td>5.2</td>
<td>8.2</td>
<td>1.58</td>
<td>(Bravman, et al., 2001)</td>
</tr>
<tr>
<td>n/a</td>
<td><em>Geobacillus stearothermophilus</em></td>
<td>n/a</td>
<td>1.2</td>
<td>n/a</td>
<td>n/a</td>
<td>(Nanmori, et al., 1990)</td>
</tr>
<tr>
<td>xylC</td>
<td><em>Thermoanaerobacterium saccharolyticum</em></td>
<td>pH 6.0, 65°C</td>
<td>28</td>
<td>n/a</td>
<td>n/a</td>
<td>(Shao, et al., 2011)</td>
</tr>
<tr>
<td>XynB</td>
<td><em>Thermoanaerobacterium saccharolyticum</em></td>
<td>pH 5.5, 37°C</td>
<td>0.036</td>
<td>9.72</td>
<td>270</td>
<td>(Vocadlo, et al., 2002)</td>
</tr>
</tbody>
</table>

The value of Michaelis-Menten constant, $K_M$, measures the affinity for the substrate to bind to the enzyme. Thus, the lower $K_M$ value for an enzyme, the higher affinity with the substrate; in other words, less substrate will cause the reaction to proceed at its half maximal reaction velocity. The $K_M$ value of enzyme depends on specific substrate, temperature, pH value and ionic strength as well (Berg, et al., 2002). Comparing the $K_M$ values for XynD with other enzymes, the affinity for $p$NP-$β$Xyl to XynD was at an intermediate level, with at least an order of magnitude difference in each direction. Therefore, a high substrate concentration would be required to achieve the maximum reaction velocity. The affinity of the substrate to the enzyme plays an important role in the catalyzed reaction and is impacted by the structure of the enzyme, which will be discussed in the following section.
The ratio of $k_{\text{cat}}/K_M$ represents the catalytic efficiency of the enzyme and determines whether there is a physiological limitation in the reaction and is also used to compare the enzymes’ preference for substrates. The greater the value of $k_{\text{cat}}/K_M$, the more rapidly and efficiently the substrates were converted into products. The affinity of $XynD$ with $p$NPβXyl was only in an intermediate level and therefore the catalytic efficiency for this substrate was poor at a value of 0.0049 mM$^{-1}$ s$^{-1}$ (Table 4).

Since β-xylosidases cleave terminal xylose residues from the non-reducing end of xylo-oligosaccharides, xylose is the principle natural end product inhibitor of xylo-oligosaccharides (Rasmussen, et al., 2006). Therefore, inhibition effects of the hydrolysate were investigated at xylose concentrations of 3 and 10 mM using pNPβXyl as substrate. The inhibition dissociation constant, $K_i$, was determined from equation fitting on Lineweaver-Burk plots (Figure 8) using inhibition and found to be 24 mM (Dixon, 1953).
Figure 8: a) Lineweaver-Burk inhibition plot of xylose on XynD, and b) Michaelis-Menten plot of the same data showing the increase in the corresponding half-maximal

When the hydrolysis reaction was run in the presence of xylose, Lineweaver-Burk plot analysis of the resulting data indicated that competitive inhibition was occurring as characterized by the intercept near the y-axis. In competitive inhibition, the inhibitor and substrate share affinity for the same active site and compete to occupy it. When the inhibitor is bound to the active site, it excluded the substrate from binding and therefore from being processed to the product. In the presence of a competitive inhibitor, it takes a higher substrate concentration to achieve the same velocities that were reached in its absence. The apparent $K_M$ value for the substrate would be increased, but no change in the apparent $V_{max}$ should be observed. Because the $V_{max}$ remains constant, it can be approached by using excessive substrate to overwhelm the inhibitor and produce saturated enzyme-substrate complexes.

Xylose acting as a competitive inhibitor to β-D-xylosidases has been reported in many papers (Table 5). The $K_i$ for most these enzymes was less than 10 mM, indicated a potential inhibitory effect at relatively low concentrations of xylose. In the case of XynD,
the $K_i$ value was 24 mM, indicating that minimal inhibition occurs in the presence of xylose, whereas β-TXE was inhibited at a lower concentration of only 1.3 mM. The xylose tolerant of enzyme would be essential to keep the efficiency of a developed process.

**Table 5:** Competitive Inhibitor xylose acted different $K_i$ in variety of enzymes.

<table>
<thead>
<tr>
<th>Enzyme name</th>
<th>Organism</th>
<th>Glycoside hydrolase Family</th>
<th>$K_M$ (mM)</th>
<th>$K_i$ (mM)</th>
<th>Reference</th>
<th>Substrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>XynD</td>
<td><em>Caldicellulosiruptor saccharolyticus</em></td>
<td>GH 39</td>
<td>2.68</td>
<td>24</td>
<td>This work</td>
<td>pNP-β-xylopyranoside</td>
</tr>
<tr>
<td>N/A</td>
<td><em>Fusarium proliferatum</em></td>
<td>GH 43</td>
<td>0.77</td>
<td>5</td>
<td>(Saha, 2003)</td>
<td></td>
</tr>
<tr>
<td>xlnD</td>
<td><em>Aspergillus niger</em></td>
<td>GH 3</td>
<td>0.255</td>
<td>3.35 ±0.78</td>
<td>(Selig, et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>XylA</td>
<td><em>Aspergillus japonicus</em></td>
<td>GH 11</td>
<td>0.33 ±0.04</td>
<td>2.9 ±0.2</td>
<td>(Semenova, et al., 2009)</td>
<td>pNP-β-xylopyranoside</td>
</tr>
<tr>
<td>N/A</td>
<td><em>Trichoderma reesei</em></td>
<td>GH 11</td>
<td>0.51 ±0.04</td>
<td>1.4 ±0.1</td>
<td>(Semenova, et al., 2009)</td>
<td>pNP-β-xylopyranoside</td>
</tr>
<tr>
<td>β-XTE</td>
<td><em>Talaromyces emersonii</em></td>
<td>GH 3</td>
<td>0.06 ±0.01</td>
<td>1.3</td>
<td>(Rasmussen, et al., 2006)</td>
<td>pNPX</td>
</tr>
<tr>
<td>β-XTR</td>
<td><em>Trichoderma reesei</em></td>
<td>GH 3</td>
<td>0.8 ±0.006</td>
<td>2.4</td>
<td>(Rasmussen, et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>XynBH6</td>
<td><em>Thermoanaerobacterium saccharolyticum</em></td>
<td>GH 39</td>
<td>N/A</td>
<td>20</td>
<td>(Vocadlo, et al., 2002)</td>
<td>N-bromoacetyl-α-D-xylopyranosylamine</td>
</tr>
</tbody>
</table>

The degradation of xylo-oligosaccharides with two to five xylose residues was investigated using TLC to understand the hydrolysis pattern of XynD. Oligosaccharides were incubated for up to 30 hours at 64°C and the end products resulting from degradation imaged, as shown in Figure 9. Each of the oligosaccharides was capable of being degraded into xylose, however, hydrolysis of xylobiose appears to have the worst performance of the oligosaccharides.

β-xylosidases are exo-acting hydrolases that can act on oligosaccharides of various lengths, but always hydrolyze the substrate from the non-reducing end (Kumar & Ramón, 1996). Using thin layer chromatography to characterize the hydrolysis patterns of xylobiose, xylotriose, xylotetraose, and xylopentaose, continued incubation at 64°C
resulted in a greater extent of hydrolysis. \( \text{XynD} \) appears to have greater activity against the longer oligosaccharides than on xylobiose. This phenomenon was also reported by a \( \beta \)-xylosidase, from \textit{Trichoderma reesei} (Herrmann, et al., 1997), whose activity increased slightly with increasing oligosaccharides chain length and liberated xylose from xylan. \( \text{XynD} \) would be atypical for a \( \beta \)-xylosidase, but was not a xylanase that was supported by \( \text{XynD} \) showed no activity with xylan.

![Thin layer chromatography of the products formed from the hydrolysis of xylo-oligosaccharides incubated with \( \text{XynD} \). Lane 1, mixture of xylo-oligosaccharides; Line 2 to 4, xylobiose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 5 to 7, xylotriose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 8 to 10, xylotetraose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 11 to 13 xylopentaose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively.]

\textbf{Figure 9}: Thin layer chromatography of the products formed from the hydrolysis of xylo-oligosaccharides incubated with \( \text{XynD} \). Lane 1, mixture of xylo-oligosaccharides; Line 2 to 4, xylobiose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 5 to 7, xylotriose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 8 to 10, xylotetraose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively; Line 11 to 13 xylopentaose incubated with \( \text{XynD} \) with 16/30 hour and the positive control, respectively.

Gene expression levels of \( \text{XynD} \) were investigated when \textit{C. saccharolyticus} was grown in different kinds of substrates (VanFossen, et al., 2011). In these studies, they tested the utilized DNA microarrays to determine the transcriptional expression levels for genes of \textit{C. saccharolyticus} grown on switchgrass, poplar, glucose, xylose, xylan,
xyloglucan and xylogluco-oligosaccharide, respectively. The highest differential expression of \textit{XynD} occurred in switchgrass and poplar, two xyloglucan carbon sources, which result in xylogluco-oligosaccharides as major hydrolysis components during processing. Our studies support the hydrolysis of the terminal xylose saccharide by \textit{XynD}, however, the activity of \textit{XynD} is very low on \(\beta\)-glucosidic bonds, which indicates that an alternate method for hydrolyzing the glucosidic bonds would be required. This is observed with expression of \(\beta\)-glucosidase (BglA, \textit{Csac\textunderscore 1089}) as well in the transcriptomic studies. The preference for hydrolysis of larger xylo-oligosaccharides by \textit{XynD} supports the transcriptional results through supporting complete hydrolysis of larger xylooligosaccharides arising from the action of xylanases.

The \textit{XynB} (GH 39), from \textit{Thermoanaerobacterium saccharolyticus} B6A-R1 (Lee & Zeikus, 1993), has been reported to have strong sequence similarity to \textit{XynD} and was active towards \(p\)NP\(\beta\)Xyl and xylo-oligosaccharides (n=2 to 5), but no activity to xylan. \textit{T. saccharolyticum} and \textit{C. saccharolyticus} are both members of the order \textit{Thermoanaerobacterales}. Based on the same affiliation with the protein of crystal structure of \textit{XynB} (PDB code: 1uhv), from \textit{T. saccharolyticum}, (Yang, et al., 2004) the amino acid sequence (BLAST) of \textit{XynD} was modeled to get the 3D structure (Figure 10, by \url{http://swissmodel.expasy.org/}). The model of \textit{XynD} 3D structure was therefore estimated by using the known \(\beta\)-xylosidase structure as model. The consensus pattern of GH family 39 active site was determined as W-x-F-E-x-W-N-E-P-[DN], located from amino acids 153 to 162 in these organisms (for \textit{XynD}, located from 159 to 168), and the second glutamic acid (E) was predicted as the active site residue (Bravman, et al., 2001; Yang, et al., 2004).
Figure 10: Estimated quaternary structure 3D model of the \textit{XynD}, from \textit{C. saccharolyticus}, were present in the asymmetric unit (visualized by VMD).
Figure 11: Amino acid sequence XynD, from *C. saccharolyticus*, aligned with protein XynB, from *T. saccharolyticum*, XynB, from *G. stearothermophilus* and XynB2, from *C. crescentus*. Black, the active site loci; grey, same amino acid.
Beside β-xylosidase *XynB*, from *T. saccharolyticum*, two other structures for members of GH39 haven been solved: *XynB*, from *Geobacillus stearothermophilus* (Czjzek, et al., 2005), and *XynB2*, from *Caulobacter crescentus* (Santos, et al., 2012). The primary sequence alignment results showed that the sequence identity for these enzymes with Csac_*XynD* is 72%, 62% and 33%, respectively (Figure 11). This primary sequence similarity also extends to secondary and tertiary structural similarities, as is expected within a GH family. In particular, around the active site, the enzymes from GH 39 family share a similar catalytic domain of (β/α)_8 barrel fold (Czjzek, et al., 2005; Yang, et al., 2004), containing eight parallel β-strand and eight parallel α-helices, which formed a pocketed shape. Two active sites glutamate residues, Glu160 and Glu277 (for Ts. *XynB*), act as the acid/base catalyst and nucleophile to assist the hydrolysis reaction taking place, are locating at the C-terminal ends of β-strands 4 (acid/base) and 7 (nucleophile), respectively (Yang, et al., 2004; Czjzek, et al., 2005). The predicted folded structure for Csac_*XynD* shows common structural aspects, both in the global structural features and also at the active site (Figure 12). Using the information from sequence alignments the active site residues for Csac_*XynD* are predicted to be Glu166 and Glu283 (Yang, et al., 2004).
**Figure 12:** a) $(\beta/\alpha)_n$-barrel fold 3D secondary structure of *XynD*, from *C. saccharolyticus* and the active site Glu166 (acid/base) / Glu283 (nucleophile); b) overlaid 3D structure for Csac$_{xynD}$ (pink), *T. saccharolyticum, XynB* (blue), *G. stearothermophilus XynB2* (green), and *C. crescentus XynB2* (yellow) constructed using PyMol.
CHAPTER 3

CONCLUSION

Conclusion

\(XynD\), a 1,4-\(\beta\)-xylosidase, from \textit{C. saccharolyticus} is a thermostable \(\beta\)-xylosidase of molecular mass 59kDa. It was cloned into pET21b and recombinantly expressed in \textit{E. coli}, which is used for characterization studies. Substrates, \(pNP-\beta\)-xylopyranosides were incubated with \(XynD\) to investigate the hydrolysis pattern at optimal temperature and pH condition. The kinetics parameters \(K_M\) value was 2.68 mM; the initial velocity \(V_{max}\) was 58.5 \(\mu\)M/min, the \(k_{cat}\) was 0.13 s\(^{-1}\) and \(k_{cat}/K_M\) is 0.049 s\(^{-1}\) mM\(^{-1}\), respectively. These series of parameters indicated \(XynD\) was indeed a \(\beta\)-xylosidase (EC 3.2.1.37) from GH 39 family, but with limited catalytic efficiency. This enzyme showed the specific selectivity to catalyze the degradation of the glycosidic bond for terminal \(\beta\)-xylopyranosides from oligosaccharides, which hydrolyzed each xylo-oligosaccharide from \(n=2\) to 5 to xylose or shorter chain. \(XynD\) showed a preference for hydrolyzing longer xylooligosaccharides.

The three-dimension structure of \(XynD\) was stimulated by the confirmed crystal structure of \(\beta\)-D-xylosidase from \textit{T. saccharolyticum}. In addition, enzymes from \textit{G. stearothermophilus} and \textit{C. crescentus} also showed similarities with \(XynD\) both from kinetics characterizations and the three-dimension structure. The specific two active site
residues of \( \textit{XynD} \) were determined to be Glu166 (acid/base) and Glu283 (nucleophile) and active sites pocket of \( \textit{XynD} \) was stimulated and aligned with the other relevant enzymes confirming similarities in the global structure.

Herein, we characterized the kinetic properties and hydrolysis capability of Csac\_\textit{XynD} for processing xylo-oligosaccharides. The high \( K_M \) value and resulting low catalytic efficiency of this enzyme compared to other \( \beta \)-xylosidases, indicate that it does not greatly improve the processing capability of xylo-oligosaccharides. On the other hand, \( \textit{XynD} \) shows high tolerance against inhibiting xylose that would be beneficial in practical application. Enzyme chemical characteristics stem from the structural organization of the enzymes, which can provide insight into the function and specificity of the enzyme. The studies in 3D structure of \( \textit{XynD} \) indicated the classic barrel fold and two active sites. High amino acid similarities of GH 39 family enzymes showed the high alignment in 3D structure as well.

\textbf{Future Work}  

\( \textit{XynD} \) showed a preference for hydrolysis of longer xylo-oligosaccharides, which could be a benefit when combined with the disaccharide processing capability of Ts\_\textit{XynB} (which is also active at elevated temperatures). The synergistic effect of Cs\_\textit{XynD} and Ts\_\textit{XynB} would have the benefit of hydrolyzing both the larger oligosaccharides and the disaccharides. This synergistic action only comes about once the xylan is hydrolyzed by endo-acting xylanases into the xylooligosaccharides. Therefore, complete breakdown of xylan still requires the synergistic action of several hemicellulolytic enzymes, including endo-1,4- \( \beta \)-xylanases and the \( \beta \)-xylosidases. The endo-1,4- \( \beta \)-xylanases, Csac\_2405 and Csac\_2408, are potential candidates for xylan
processing enzymes that also come from *C. saccharolyticus*. Future studies could evaluate the synergistic effect of a β-xylosidase pairing such as these which would have the benefit of hydrolyzing the larger oligosaccharides and the disaccharides. The continued discovery and characterization of enzymes, such as *XynE* (Csac_2410), for processing cellulose, Csac_2405 (putative xylanase), and *XynA* (Csac_2408) for xylan are important to reduce the costs for the second generation biofuels. The verification of the biochemical capabilities of these enzymes is important as a means of complimenting functional genomics and proteomic studies identifying the enzymes
REFERENCES


Available at:
http://www.rsc.org/Education/EiC/issues/2009May/Biofuelsthenextgeneration.asp

[Accessed 6 2012].


Sørensen, H. R., Meyer, A. S. & Pedersen, S., 2003. Enzymatic hydrolysis of water-soluble wheat arabinoxylan. 1. Synergy between α-L-arabinofuranosidases, endo-1,4-β-
xylanases, and β-xylosidase activities. *Biotechnology and Bioengineering*, 81(6), pp. 726-731.


APPENDIX A

PRIMER SEQUENCE AND RELEVANT PCR MATERIALS

<table>
<thead>
<tr>
<th>Table A 1: dNTPs recipe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dATP</strong></td>
</tr>
<tr>
<td><strong>dTTP</strong></td>
</tr>
<tr>
<td><strong>dCTP</strong></td>
</tr>
<tr>
<td><strong>dGTP</strong></td>
</tr>
<tr>
<td>DI water</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Working stock concentration for each of them is 10mM. Aliquot into 2 tubes.

<table>
<thead>
<tr>
<th>Table A 2: PCR primer sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ORF4</strong> (Csac_2405)</td>
</tr>
<tr>
<td><strong>XynA</strong> (Csac_2408)</td>
</tr>
<tr>
<td><strong>XynD</strong> Csac_2409</td>
</tr>
<tr>
<td><strong>XynE</strong> Csac_2410</td>
</tr>
</tbody>
</table>
### Table A 3: Primer %GC and melting temperature


<table>
<thead>
<tr>
<th></th>
<th>2405 Fwd</th>
<th>2405 Rev</th>
<th>2408 Fwd</th>
<th>2408 Rev</th>
<th>2409 Fwd</th>
<th>2409 Rev</th>
<th>2410 Fwd</th>
<th>2410 Rev</th>
</tr>
</thead>
<tbody>
<tr>
<td>% GC</td>
<td>33.33</td>
<td>26.92</td>
<td>26.92</td>
<td>33.33</td>
<td>43.33</td>
<td>45.45</td>
<td>40</td>
<td>30.77</td>
</tr>
<tr>
<td>Basic melting temperature, °C</td>
<td>51</td>
<td>50</td>
<td>50</td>
<td>51</td>
<td>60</td>
<td>63</td>
<td>62</td>
<td>52</td>
</tr>
</tbody>
</table>

### Table A 4: Restriction Enzyme cutting position

<table>
<thead>
<tr>
<th></th>
<th>5'...C A T A T G...3'</th>
<th>3'...G T A T A C...5'</th>
<th>5'...C C A T G G...3'</th>
<th>3'...G G T A C C...5'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>Ndel</td>
<td>NcoI</td>
<td>NcoI</td>
<td>NcoI</td>
</tr>
<tr>
<td></td>
<td>5'...G C T A G C...3'</td>
<td>3'...C G A T C G...5'</td>
<td>5'...C T C G A G...3'</td>
<td>3'...G A G C T C...5'</td>
</tr>
</tbody>
</table>
APPENDIX B

LIQUID AND SOLID MEDIA

<table>
<thead>
<tr>
<th></th>
<th>LB substrate</th>
<th>LB medium plate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>10g</td>
<td>2.5g</td>
</tr>
<tr>
<td>NaCl</td>
<td>10g</td>
<td>2.5g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>5g</td>
<td>1.25g</td>
</tr>
<tr>
<td>DI water</td>
<td>1 liter</td>
<td>250mL (for 15 plates)</td>
</tr>
<tr>
<td>Agar</td>
<td>N/A</td>
<td>15g/L</td>
</tr>
</tbody>
</table>

Autoclave

| Antbiotics     | N/A          | Add the corresponding anti-biotic into liquid when substrate temperature drop to 50-55°C, 1µL anti-biotic / 1mL liquid. |
| Storage        | Keep them at 4°C until use | Keep them at 4°C until use |
### Table B 2: SDS-Page Gel Protocol

<table>
<thead>
<tr>
<th>Resolving Gel</th>
<th></th>
<th>Stacking Gel</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>Amount</td>
<td>Component</td>
<td>Amount</td>
</tr>
<tr>
<td>DI Water</td>
<td>4.9mL</td>
<td>DI Water</td>
<td>5.5mL</td>
</tr>
<tr>
<td>30% Acrylamide-bis</td>
<td>6.0mL</td>
<td>30% Acrylamide-bis</td>
<td>1.3mL</td>
</tr>
<tr>
<td>1.5M Tris (pH 8.8)</td>
<td>3.8mL</td>
<td>1 M Tris (pH 6.8)</td>
<td>1.0mL</td>
</tr>
<tr>
<td>10% SDS Page</td>
<td>150μL</td>
<td>10% SDS Page</td>
<td>80μL</td>
</tr>
<tr>
<td>10% Ammonium persulfate</td>
<td>150μL (fresh)</td>
<td>10% Ammonium persulfate</td>
<td>80μL (fresh)</td>
</tr>
<tr>
<td>TEMED</td>
<td>6μL</td>
<td>TEMED</td>
<td>8μL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1% Bromeophenol Blue</td>
<td>10μL</td>
</tr>
</tbody>
</table>

1% Agarose shell and 5X stock TBE solution recipe

1. 1.9g Agarose dissolve into 200mL DI water.
2. Microwave the solution about 3 min, until bubbles just arise from the bottom. (time always depends on the power of microwave oven)
3. Add 15 sec each time, until the solution is boiling
4. Incubate the bottle at 55°C, water bath, 20 min
5. Pour into tray and put notches
6. Cooling it for 30 min and remove the notches
7. Add 1 time TBE buffer and make the gel immerse into the buffer.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tris base</strong></td>
<td>54g</td>
</tr>
<tr>
<td><strong>Boric acid</strong></td>
<td>27.5g</td>
</tr>
<tr>
<td><strong>EDTA</strong></td>
<td>3.72g (20 mL of 0.5M)</td>
</tr>
<tr>
<td><strong>DI water</strong></td>
<td>To 1000mL</td>
</tr>
</tbody>
</table>
APPENDIX C

LIGATION

SAP Treatment

1. Incubate SAP (1 unit/µg DNA) with restriction-digested vector at 37°C for 15 min in 1 time SAP reaction buffer in a final volume of 30-50µL.

2. Inactivate SAP by heating to 65°C for 15 min.

3. Centrifuge the reaction mixture. Then, do the Ta ligation.

Ligation Mixture

Ligation is the process of introducing a gene into a plasmid. The gene and the plasmid are prepared using the same restriction enzyme. The plasmid is treated with SAP to prevent it from self-ligating. The plasmid and gene are then placed in the same solution with a DNA ligase. The ligase fuses the gene and plasmid together. The following table lists the components used when introducing the target gene into a plasmid, either pET-21b or pET-24d.
<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid</td>
<td>5μL</td>
</tr>
<tr>
<td>Gene</td>
<td>2μL</td>
</tr>
<tr>
<td>10x Ligase Buffer</td>
<td>1μL</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>1μL</td>
</tr>
<tr>
<td>Nuclease Free Water</td>
<td>1μL</td>
</tr>
</tbody>
</table>

**Table C 1: Ligation mixture**