BIOCHEMICAL CHARACTERIZATION OF THE HIGHLY THERMOSTABLE β-XYLOSIDASE FROM *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

Dilan Karunathilaka Wellalage Don

Dayton, Ohio

August, 2019

BIOCHEMICAL CHARACTERIZATION OF THE HIGHLY THERMOSTABLE

β-XYLOSIDASE FROM *Caldicellulosiruptor saccharolyticus*

Name: Wellalage Don, Dilan Karunathilaka

APPROVED BY:

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

_____
Kevin J. Myers, D.Sc., P.E
Committee Member
Professor
Department of
Chemical & Materials Engineering

_____
Kevin Hinkle, Ph.D.
Committee Member
Assistant Professor
Department of
Chemical & Materials Engineering

_____
Robert J. Wilkens, Ph.D., P.E.
Associate Dean for Research and Innovation
Professor
School of Engineering

_____
Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering

ABSTRACT

BIOCHEMICAL CHARACTERIZATION OF THE HIGHLY THERMOSTABLE β-XYLOSIDASE FROM *Caldicellulosiruptor saccharolyticus*

Name: Wellalage Don, Dilan Karunathilaka
University of Dayton

Advisor: Dr. Donald A Comfort

There has recently been an increasing focus on the conversion of lignocellulosic biomass to biofuel as an alternative to petroleum. The current bottleneck for the process is efficient hydrolysis of lignocellulose into simple sugar molecules for fermentation to ethanol. Xylan represents the major hemicellulose in plants and is identified as the second most abundant polysaccharide on earth. The complete degradation of xylan requires several enzymes working synergistically, including endoxylanases and β-xylosidases. β-xylosidases are capable of hydrolyzing xylo-oligosaccharides to xylose. Thermostable β-xylosidases are more desirable in biofuel production due to their ability to withstand harsh process conditions. This research characterizes glycoside hydrolase enzymes from the extreme thermophilic bacterium *Caldicellulosiruptor saccharolyticus*, which are predicted to possess the ability to degrade xylan into the fermentable sugar xylose. Thermostable β-xylosidase encoded by Csac_2409 of GH39 from *C. saccharolyticus* was recombinantly expressed by GenScript and the protein purified to 75% purity. The protein was then characterized to determine the substrate preference, optimal temperature, pH value, thermal stability, and kinetic constants. Thermostable β-xylosidase showed activity over wide range of pH and temperature with optimal pH of 6.5 and temperature of 80°C. The enzyme indicated high thermal stability at 70°C with half-life close to 3 hours. Michaelis-Menten kinetic parameters, K_M, V_{max}, k_{cat}, and k_{cat}/K_m were determined to be 0.918 mM, 0.251 mM/min, 13.6 s^{-1}, and 14.75 $s^{-1}mM^{-1}$, respectively. Understanding the function and

optimal conditions of the enzyme could help the advancement of the lignocellulosic ethanol process, which would ultimately lead to less fossil fuel usage and more environmentally friendly transportation fuels.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Dr. Donald Comfort for providing this opportunity and for his continuous support of research work. His expertise and guidance have been invaluable throughout this research and in writing the thesis. Without his motivation, patience, and understanding this study would not have been possible.

I would also like to thank the thesis committee: Dr. Kevin Myers and Dr. Kevin Hinkle for their encouragement, time, and guidance.

In addition, I would like to thank Dr. Rochael Swavey for offering me the opportunity to serve as a teaching assistant in the chemistry lab.

Finally, I express my sincere gratitude to my parents for their support and encouragement throughout my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO LIGNOCELLULOSIC BIOFUEL

Objective

There has been a conscious effort to move towards renewable energy sources as a result of both the economic and environmental costs associated with fossil fuels. Unfortunately, these alternative energy sources such as hydropower, nuclear, wind, and solar systems are not directly transferable to transportation energy sources. As a result, liquid fuel sources such as ethanol have been developed over the last few decades. The first-generation biofuels (starch-based ethanol) are currently commercially implemented, however, they utilize food stocks to produce ethanol, whereas lignocellulosic ethanol (second-generation ethanol) avoids this problem. Implementation of cellulosic ethanol has been delayed due to technological challenges associated with low yield and conversions. Despite this, continued research is being invested for biomass energy which is more efficient and sustainable compared to other traditional energy resources.

Several technologies have been developed to produce biofuel from lignocellulosic biomass including biochemical, chemical, and thermochemical processes (NREL, 2006). In fact, alternative methods have been proposed to break down lignocellulosic biomass utilizing cellulolytic enzymes derived from thermophiles to avoid current limitations in biofuel production.

The thermophilic bacterium *C. saccharolyticus* has been considered as a great candidate for converting biomass into small biofuel molecules by producing effective carbohydrate degrading enzymes in harsh bioprocessing conditions. *C. saccharolyticus* grows optimally at 70°C on various carbohydrate substrates, which encodes different Glycoside Hydrolases (GHs). These hydrolytic enzymes function as glucanases, xylanases, and mannanases in addition to side

chain degrading enzymes. In this work, the β-xylosidase from gene Csac_2409 was characterized to evaluate optimal conditions, thermal stability, and enzymatic kinetic parameters.

Background

In recent decades, there has been a significant increase in energy demands as developing countries utilize more energy and developed countries maintain and increase usage. In fact, global energy consumption increased by 2.3% in 2018, and nearly two-fold since 2010 (Global Energy & CO_2 Status Report. GECO, 2019). Transportation fuels, such as petroleum, represent one fuel for which there has been an unsteady, but general increase in demand. Petroleum and other fossil fuels have traditionally been utilized for transportation; however, these resources are limited and require increased investments for extraction. Fossil fuels are nonrenewable sources which account for the vast majority of the global energy supply. Burning fossil fuel releases waste gases contributing to the greenhouse effect, which influences climate change, in addition to the desired energy.

In the United States, the largest source of greenhouse gas emissions is from burning fossil fuel such as coal, hydrocarbon gas liquids, natural gas, and petroleum for electricity, heat and transportation. Carbon dioxide from combustion of fossil fuel, the largest contributor to U.S greenhouse gas emissions has accounted for about 77 percent of GWP-weighted (Global Warming Potential) emissions since 1990. In addition, emissions from fossil fuel combustion has been increased by 6.5 percent (309.4 MMT CO_2 Eq.) from 1990 to 2015. Thus, the changes in anthropogenic emissions from fossil fuel combustion have been the major factor that affects U.S emission trends (Caspeta, L., Buijs, N. A., & Nielsen, J., 2013).

In terms of total greenhouse gas emission in U.S, transportation has accounted for 28.7 percent, which is the largest share of greenhouse gas emissions in 2017. The Intergovernmental Panel on Climate Change (IPCC) has recommended several strategies to reduce greenhouse gas emissions related to transportation such as fuel switching, improving fuel efficiency, improving operating practices, and reducing travel demand. In terms of fuel switching, alternative fuel

sources like biofuels, hydrogen, and electricity from renewable energy sources have been suggested (EPA, 2019).

Bioethanol for Transportation

A mitigation strategy for environmental costs associated with fossil fuel is to use biofuels for transportation. Biofuels are derived from biological material mainly from plants, and they are the most abundant renewable source of energy used in transportation. Biofuels including ethanol, butanol, and long-chain hydrocarbons are generated from starch, cellulose, hemicellulose, and oils. These fuels are usually blended with fossil fuels such as gasoline and diesel additives, however, they can also be used on their own. Using biofuels in addition to petroleum-based fuels helps to reduce the negative impact on the environment caused by burning fossil fuel. Liquid biofuels are categorized based on sources of feedstock, improvements in technology and process involved, and are referred to as first, second, and third generations biofuels.

First Generation Biofuels

The first-generation biofuels include ethanol and biodiesel derived directly from biomass such as cane sugar, starch or vegetable oil. Ethanol is usually generated from C_6 sugars, mainly glucose under the influence of yeast strains such as *Saccharomyces cerevisiae* which processes the sugar for energy extraction via fermentation. The production of first-generation bioethanol uses only a few different feedstocks, mostly sugarcane or corn in addition to whey, barley, and sugar beets (EIA, 2019).

Biodiesel is the other renewable fuel produced from biomass on an industrial scale in the United States. Consumption of biodiesel in the U.S has been increasing since 2001 as a direct substitute for conventional diesel fuel. Transesterification is used to produce biodiesel from vegetable oil or animal fats. During this process, water and other contaminants are removed from feedstock and then mixed with an alcohol and a catalyst to break apart oil molecules (Hood,

2016). As the product of this process, methyl esters (biodiesel) and glycerol are formed, which are then separated from each other. The produced methyl esters can either be mixed with conventional diesel fuel or used in pure form. Biodiesel is often blended with fossil fuel derived diesel at several ratios (EIA, 2019).

The first-generation biofuels provide several advantages including reduction of greenhouse gas emissions and energy balance. However, there are some issues such as high food prices due to competition with food crops, use of fertilizers, limited greenhouse gas emission reduction, deforestation, and negative impact on biodiversity (IEA, 2008).

Second Generation Biofuels

The second-generation biofuels are derived from non-edible feedstock such as wood chips, forestry and agricultural residues, organic solid wastes, and energy crops. Many of the issues related to first-generation biofuels can be addressed with the production of biofuels utilizing lignocellulosic feedstock. In fact, use of these materials provides environmental sustainability and less competition for arable lands (European commission, 2015). Due to their abundance, availability, low cost and high yield of ethanol, lignocellulosic feedstock is considered as a promising option for bio-ethanol production.

The second-generation biofuels are produced according two different pathways: thermocatalytic ("thermo") and biological ("bio") pathways (Figure 1.1). In the thermo pathway biomass is processed in the presence of minimal amounts of oxidizing agents. All processes in the thermo pathway result in three major products: biochar (solid), pyrolytic oil or bio oil, and syngas. When biomass is treated at low temperatures (250-350°C) without oxygen, it undergoes a teorrefaction process and produces biochar. Bio oil can be obtained as the major conversion product when biomass is processed at higher temperatures (550-750°C) without oxygen (pyrolysis process). Furthermore, at greater temperatures (750-1200°C) gasification of biomass

produces syngas as the major product in addition to bio oils and biochar as by-products. The partial or total oxidation of carbon from biomass supplies the required energy to heat biomass. Therefore, thermal processes are self-sufficient up to a certain extent in terms of energy usage (Lee, R A, and Jean-Michel Lavoie., 2013).

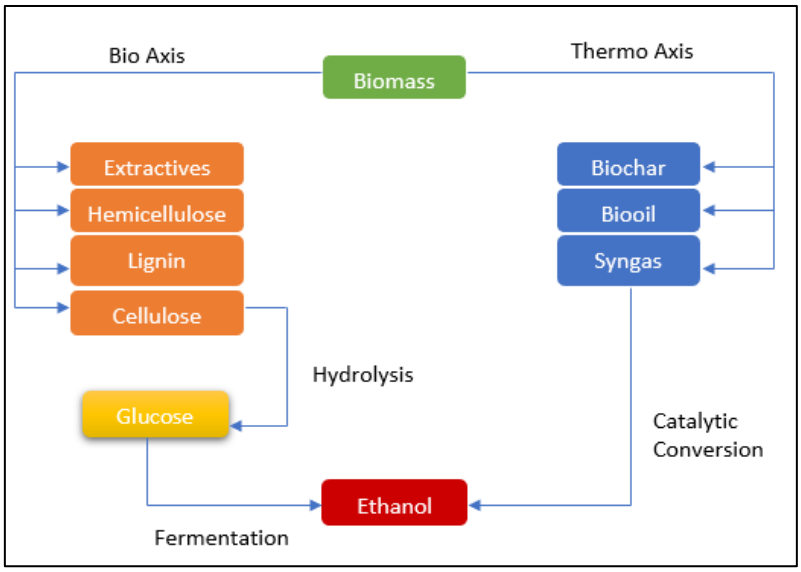


Figure 1.1: Simplified scheme for the "bio" and "thermo" pathways for conversion of lignocellulosic biomass into biofuels. (Lee, R.A., and Jean-Michel Lavoie., 2013).

Biochar is a popular solid biofuel mainly used in pelletization where lignocellulosic biomass can be obtained at low prices (Clarke and Presto, 2011). In addition, pyrolytic oil or syngas is a significant intermediary for transportation fuel. However, in order to use bio oil as a fuel for transportation, further transformations must be made utilizing processes such as hydrodeoxygenation, catalytic cracking, steam reforming and emulsification to use with diesel (Zhang et al., 2006). Syngas from gasification of biomass can be used to produce renewable liquid fuel via the Fischer–Tropsch process. During this process, contaminants of syngas are removed, and catalytic reactor is used to produce liquid fuel from syngas (Hu et al., 2012).

Methanol production from syngas is one of the options to produce synfuel. By using reducing catalysts, direct conversion of carbon monoxide and hydrogen to methanol can be

performed. However, methanol needs to be processed further for use as an additive in fuel. Thus, methanol can be used as a starting material to produce alkanes via methanol-to-gasoline (MTG), and ethanol via carbonylation processes (Lee, R.A, and Jean-Michel Lavoie., 2013). In theory, biomass can be converted into biofuels via thermocatalytic pathways, but there are certain technical and economic challenges which restrict the production.

The bio pathway is similar to pulping process where cellulose is first isolated from biomass. For this process many methods are considered such as classical pulping, steam explosion, and organosolv processes. After the purification, saccharification of cellulose is achieved from enzymatic hydrolysis and/or chemical hydrolysis using acids (Lee, R A., and Jean-Michel Lavoie., 2013). Hemicelluloses in biomass can be hydrolyzed into C_5 sugars to produce ethanol. Lignin from pretreatment of biomass and paper industry is typically disposed as waste. However, lignin could be used to produce high value products such as dispersants, binders, (Cannatelli, M D., and Ragauskas., 2016) value-added chemical compounds like pyrocatechol (Beauchet et al.,2012) and aviation fuel (Shabtai et al.,1999).

Third Generation Biofuels

"The most accepted definition of third-generation biofuels is fuels that derived from algal biomass which has a very distinctive growth yield as compared with classical lignocellulosic biomass" (Brennan, L and Owende., 2010).

The lipid content of the microorganisms is a very important factor in the production of biofuels from algae. Species like *Chlorella* have high lipid content (60-70%; Liang et al., 2009) and productivity (7.4 g/L/d for *Chlorella protothecoides*; Chen et al., 2011). However, the high growth rate of algae presents some challenges, especially industrial scale production of algae-based biofuels needs large amount of water to harvest algae. In addition, high water content of biomass is also an issue when extracting lipids from biomass which requires processes like

centrifugation and filtration to remove water. The extracted lipids are processed via transesterification to produce biodiesel. In fact, hydrogenolysis can be used to produce kerosene rated alkanes.

Bioethanol Production

The biomass used in the process of ethanol production is a major factor. This process includes three major steps: 1) production of fermentable sugars containing solutions, 2) fermentation of sugars and conversion into ethanol, and 3) separation and purification of ethanol (Figure 1.2). Based on the biomass and conversion technology, one or more extra steps are added to these processes. In this case, pretreatment steps are used to extract carbohydrate or to increase accessible area for further extraction (Vohra et al., 2013).

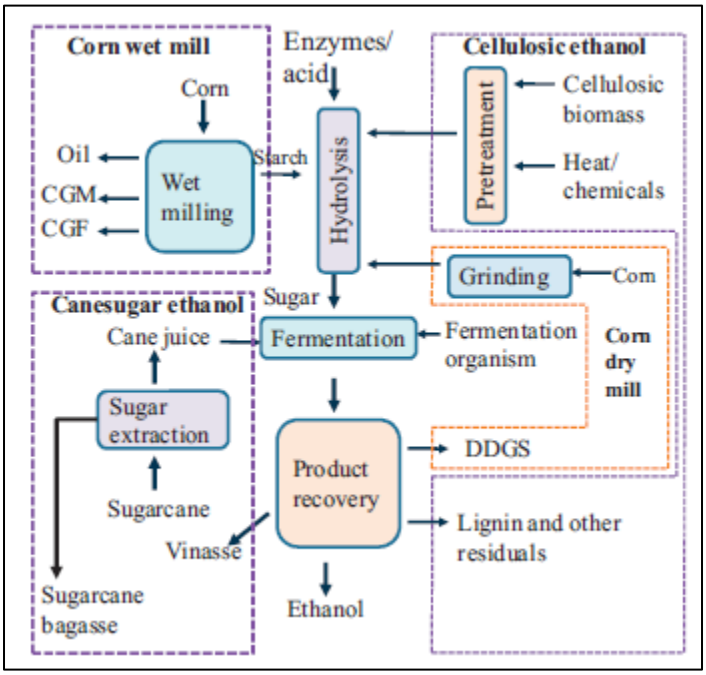


Figure 1.2: Schematic representation of production of ethanol from cane sugar, corn, and cellulosic biomass. (Vohra et al., 2013).

The hydrolysis reactions are used to convert large portions of fibers into simple sugars. At the beginning of the batch reaction, hydrolysate, yeasts, nutrients and other ingredients are

added to fermentation. During these processes, ingredients are continuously added, and the products are removed from the fermentation vessels (Wyman, 2004). The resulting broth typically contains 8-14 vol% ethanol - above this concentration, yeast activity reduces due to inhibition. The distillation process results in an azeotropic mixture of alcohol (95.5%) and water (4.5%) which is then dehydrated to obtain anhydrous ethanol containing up to 99.6% alcohol and 0.4% water (Vohra et al., 2013).

Sugars derived from sugarcanes, molasses, sugar beet and fruits can be directly fermented into ethanol via fermentation utilizing yeast (first-generation bioethanol). This process does not require pretreatment methods, milling, hydrolysis, and detoxication in order to produce ethanol from biomass. However, ethanol production using fermentable sugars derived from starchy materials requires processes like milling, liquefication, and saccharification (Vohra et al., 2013).

Bioethanol Production from Lignocellulosic Feedstock

Plant cell walls are composed mostly of lignocellulose which is the most abundant organic material on Earth. Lignocellulose consists of three polymers - cellulose, lignin, and hemicellulose in addition to small quantities of acetyl groups, mineral, and phenolic substitutes. Different types of lignocellulosic biomass exist in complex non-uniform three dimensional structures which provide recalcitrance and resistance to degradation. In general, lignocellulosic biomass consist of 35-50% cellulose, 20-35% hemicellulose, and 10-15% lignin. Table 1.1 summarizes different types of lignocellulosic biomass and their compositions (Isikgor, F.H., and Becer, 2015) and shows the breadth of compositional differences.

Table 1.1: Types of lignocellulosic biomass and their chemical composition (Isikgor, F. H., and C. Remzi Becer., 2015).

Lignocellulosic biomass		*Cellulose (%)*	*Hemicellulose (%)*	*Lignin (%)*
Hardwood	Poplar	50.8–53.3	26.2–28.7	15.5–16.3
	Oak	40.4	35.9	24.1
	Eucalyptus	54.1	18.4	21.5
Softwood	Pine	42.0-50.0	24.0-27.0	20.0
	Douglas Fir	44.0	11.0	27.0
	Spruce	45.5	22.9	27.9
Agricultural waste	Wheat Straw	35.0-39.0	23.0-30.0	12.0-16.0
	Barley Hull	34.0	36.0	13.8-19.0
	Barley Straw	36.0-43.0	24.0-33.0	6.3-9.8
	Rice Straw	29.2-34.7	23.0-25.9	17.0-19.0
	Rice Husks	28.7-35.6	12.0-29.3	15.4-20.0
	Oat Straw	31.0-35.0	20.0-26.0	10.0-15.0
	Ray Straw	36.2-47.0	19.0-24.5	9.9-24.0
	Corn Cobs	33.7-41.2	31.9-36.0	6.1-15.9
	Corn Stalks	35.0-39.6	16.8-35.0	7.0-18.4
	Sugarcane Bagasse	25.0-45.0	28.0-32.0	15.0-25.0
	Sorghum Straw	32.0-35.0	24.0-27.0	15.0-21.0
Grasses	Grasses	25.0-40.0	25.0-50.0	10.0-30.0
	Switchgrass	35.0-40.0	25.0-30.	15.0-20.0

Cellulose is the major component of lignocellulose and it is a linear polymer consisting of D-anhydroglucopyranose units (AGU) linked by β-1,4- glycosidic bonds. Many hydroxyl groups in the molecule form intermolecular and intramolecular hydrogen bonds which make cellulose highly crystalline, insoluble in water and resistant to enzyme attack. In addition, the ordered crystalline structures provide essential mechanical properties to the cellulose fibers (Börjesson, M., and Westman., 2015).

Hemicellulose, the second most abundant polymer consists of random and amorphous regions. Unlike cellulose, hemicellulose is composed of different polymers including xylan, galactomannan, glucuronoxylan, arabinoxylan, glucomannan and xyloglucan. These polymers are made of different carbon monosaccharide units such as xylose, arabinose, mannose, glucose and

galactose. Furthermore, hemicelluloses form a complex network that provides structural strength in plant cell walls by connecting cellulose fibers with lignin (Isikgor, F.H., and Becer., 2015).

The third component of lignocellulose is lignin and it is a three-dimensional polymer of phenylpropanoid units. The structure of lignin is formed by oxidative coupling of three different phenylpropane units such as *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol. Lignin acts as a glue which holds the lignocellulose matrix together which provides compressive strength to plant tissue, stiffness to the cell wall and resistance to insects and diseases.

There has been a great effort directed towards developing cost-effective methods and advanced technologies for bioethanol production from lignocellulosic feedstock. However, there are remaining challenges for the commercial-scale production of lignocellulosic bioethanol. Since biofuel production is connected to the commodity market, it is essential to reduce production costs, especially the cost of converting lignocellulosic biomass to ethanol to maintain profitability. The major challenges in using lignocellulosic feedstock are at the pretreatment steps currently used to remove lignin and enzymatic hydrolysis of lignocellulose (Lee, R.A., and Jean-Michel Lavoie., 2013).

Biomass Pretreatment

Lignocellulose is a complex matrix composed of cellulose, hemicellulose, and lignin. These biomasses from plants are recalcitrant in order to protect themselves from harmful microorganisms. Crystallinity of cellulose, porosity, lignin protection, fiber strength and cellulose sheathing contribute to recalcitrance of biomass (Agbor et al., 2011). The conversion of biomass into bioethanol at commercial-scale production, however, needs to be completed in hours or days. Therefore, it is required to make cellulose and hemicellulose more accessible to enzymes by breaking hemicellulose-lignin complex linkages within the biomass. To this end, many pretreatment methods have been developed (Balan, 2014).

The major pretreatment methods include physical, chemical, physiochemical, and biological methods. The physical methods are used to reduce particle size and to increase accessible area via mechanical processing and extrusion. Disk milling/grinding reduces particle size and extrusion methods can shorten the fibers and defibrillation to make biomass more accessible to enzymes. However, due to their high energy usage, physical treatment processes including milling are not widely used in biorefineries.

Chemical methods include processing in acidic, neutral, or alkaline conditions. For acidic treatments, mineral (H_2SO_4, HCl, H_3PO_4, and HNO_3) and organic (fumaric, maleic, and acetic) acids are used to remove cellulose and lignin, and solubilize hemicellulose to monomeric xylose (Mosier et al.,2005). Cellulose is then degraded to monomeric sugars via enzymatic hydrolysis. However, degradation carried out using organic acids is prone to make fewer products compared to mineral acids. For neutral conditions, ionic liquids can be used for extraction, which is capable of solubilizing both cellulose and hemicellulose. Organosolv and oznolysis are also pretreatment methods that can be processed in neutral conditions (Balan, 2014).

Alkaline methods offer several advantages during hydrolysis such as increasing internal surface area, decreasing crystallinity, cleaving carbohydrates and solubilizing lignin. In addition, alkaline methods can solubilize small quantities of hemicellulose and cellulose. The major alkaline processes use strong alkali like NaOH, KOH (cleave ester and ether bonds) and weak alkali like NH_3, or $Ca(OH)_2$ (cleaves only ether bonds) as catalyst and they are performed at varying temperatures and processing times (Bensah, E.C., and Mensah.,2013). Ammonia fiber expansion (AFEX) (90-120°C; 250-400 psi) is another pretreatment process which is widely used on lignocellulosic biomass (Hood, 2016). However, this method does not produce lignin fraction which can be used to manufacture other products, this requires addition of oxidants like O_2 and H_2O_2 to remove lignin.

Steam explosion and CO_2 explosion are two major physiochemical pretreatment methods used in biofuel production. According to Zhang et al. (2012) steam explosion with two-steps (1.1Mpa/4min and intermediate separation of fiber cell - 1.2Mpa/4 min) can improve enzymatic hydrolysis, product conversion and reduce inhibitor conversion compared to single step steam explosion (1.2MPa/8min). On the other hand, the CO_2 explosion process is more efficient than steam explosion and results in fewer degradation products. However, both methods use high pressure reactors therefore, processing costs can be high (Balan, 2014).

Biological methods use enzyme and microbes to reduce length of larger polysaccharides in lignocellulosic feedstock. Biological pretreatment methods require low capital cost when compared to other processes since they are carried out in mild conditions. But they are relatively slow processes which require several days to complete degradation (Balan, 2014).

Hydrolysis of Lignocellulosic Feedstock

Hydrolysis can be performed utilizing either enzymes or mineral acids. However, enzymatic hydrolysis is more beneficial since it offers several advantages compared to acidic hydrolysis. Enzymatic hydrolysis does not require corrosion resistant equipment and results in low acidic waste and fewer unwanted by-products. In addition, there is a possibility for a complete degradation of biomass via enzymatic hydrolysis.

Microorganisms secrete enzymes which are capable of degrading biomass to produce monomeric sugars. In nature, microbes use different systems or patterns to degrade polysaccharides which include cellulosomal, free enzymes, and multifunctional enzymes (Moraïs et al., 2012). Free enzymes have cellulose binding domains (CBMs) that contain about 35 amino acids. In cellulosome enzyme system, instead of CBMs, a dockerin domain-cohesion can be recognized (Kuhad et al., 2011). Free enzymes for commercial use are produced at large scale from fungi and bacteria. This mixture of enzymes contains 40-50 different enzymes including

cellulases, hemicellulases, and pectinases. Various types of enzymes are used to cleave bonds during hydrolysis of lignocellulosic biomass and these enzymes are known as 'molecular scissors'. Extensive research work has been performed in the past to identify synergistic activities of enzymes during hydrolysis (Hu et al., 2013; Zhou, S., and L. O. Ingram, 2000).

Enzymatic hydrolysis includes reactions that convert pretreated lignocellulosic biomass into simple sugar molecules. During this process, cellulose and hemicellulose are converted into glucose, pentoses (xylose and arabinose), and hexoses (glucose, galactose and mannose). These conversion reactions are catalyzed by cellulases and hemicellulases, which are substrate specific.

Enzymatic hydrolysis is performed at mild conditions (pH 4.8 and 45-50°C). There are three major groups of cellulases involved in complete hydrolysis of cellulose (Figure 1.3): endoglucanase, exoglucanase or cellobiohydrolase, and β-glucosidase (Escuder-Rodríguez et al., 2018; Zhang, Xiao-Zhou, and Yi-Heng Percival Zhang, 2013). Endoglucanase attacks low crystalline regions of cellulose fiber creating free chain-ends. Exoglucanase degrades the molecule further by removing cellobiose units (repeating unit of cellulose) from free chain-ends. Finally, β-glucosidase hydrolyzes cellobiose units to produce glucose (Vohra et al., 2014).

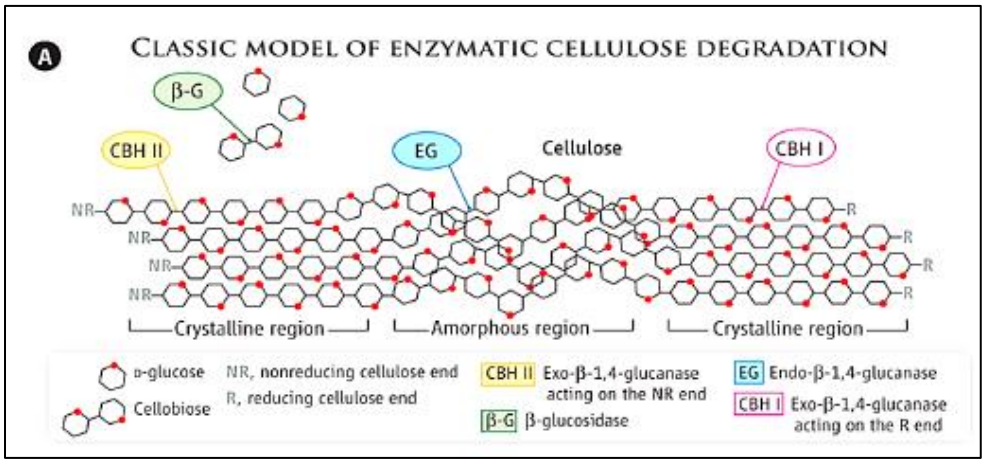


Figure 1.3: Enzymatic hydrolysis of cellulose (Dutta, S., and Wu., 2014)

Hydrolysis of cellulose to produce fermentable sugars in ethanol production is quite expensive process. Therefore, it is necessary to maximize the conversion of lignocellulosic feedstock into other major extractives, mainly hemicellulose to generate monosaccharides for fermentation.

Xylan has been identified as the most abundant hemicellulose in plants and the second largest source of biomass accounting for one-third of all renewable organic carbon on earth (Collins, T., Gerday, C., & Feller, G., 2005). Xylans are composed of pentose sugars, which require metabolic pathways that are different from those used to process hexose sugars from cellulose. In addition, the complex linkages and sidechains in xylan require a suite of hydrolytic enzymes for complete degradation of the polymer. The production of such enzymes could result in significant economic and metabolic costs (Smith et al., 2017).

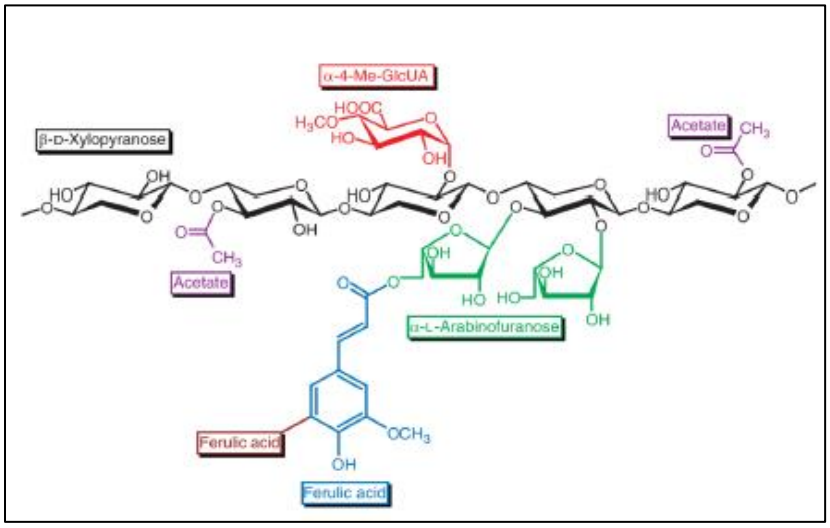


Figure 1. 4: General structure showing the various linkages found in a variety of xylans isolated from plant cell walls (Dodd, Dylan, and Cann, 2009).

Xylan is a branched polysaccharide consisting of a backbone of β-(1,4)-linked xylopyranosyl units which are typically substituted with acetyl (Ac), glucuronosyl, arabinofuranosyl side chains (Figure 1.4). Xylan substituted with different types of side chain groups are specific to the plant, and characteristics of such substitutions depend on the age of

plant and tissue source (Maehara et al., 2017). Xylans can be classified as homoxylans, arabinoxylans, glucuronoxylans, and arabinoglucuronoxylans.

Xylan Degradation

Due to complexity and heterogeneity of xylan, the complete hydrolysis of xylan must be carried out by a variety of cooperatively acting enzymes (Figure 1.5). Thus, xylan backbone is randomly cleaved by endo-1,4-β-D-xylanases while β-D-xylosidases cleave xylose monomers from the non-reducing end of xylo-oligosaccharides and xylobiose. The specific xylanases including α-L-arabinofuranosidases, α-D-glucuronidases, acetylxylan esterases, ferulic acid esterases, and *p*-coumaric acid esterases catalyze the removal of side chains (Collins, T., Gerday., and Feller., 2005). Endoxylanases degrade xylan into xylose monosaccharides and catalyze the endo-hydrolysis of 1,4-β-D-xylosidic linkages in xylans, yielding shorter 1,4-β-D-xylooligosaccharides of different lengths which can be degraded further by other hemicellulases.

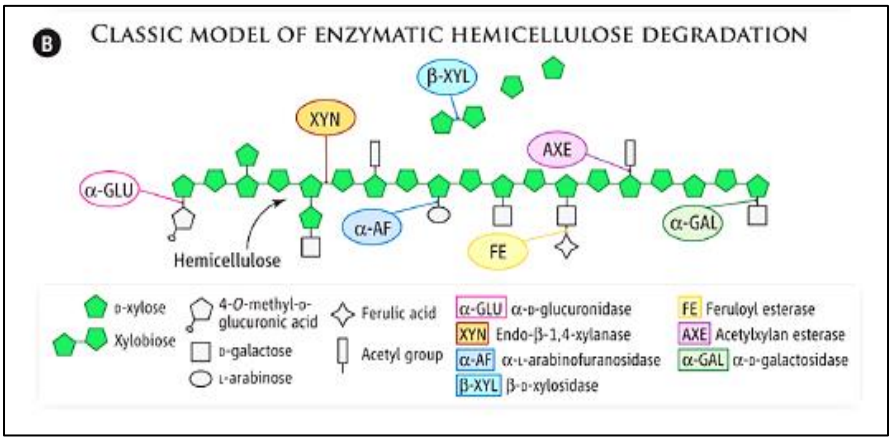


Figure 1.5: Enzymatic hydrolysis of hemicellulose (Dutta, S., and Wu., 2014)

The activity of endo-1,4-*β*-xylanase (xylanase) is critically important for the depolymerization of xylans. These xylanases cleave glycosidic linkage between xylose units in the backbone of xylans. The xylanase enzymes have been categorized into GH families of 5, 7,8,10, 11, and 43. Initially, amino acid sequence similarities are used to classify glycoside

hydrolases and transglucosylases into families. However, structural and mechanical analysis have been used for new classification (Henrissat and Davies, 2000). There are two binding sites (subsites) for xylose residues in xylanases: reducing end and non-reducing end. GH 10 and 11 xylanases are the best studied families and they have different number of subsites - GH 10 possesses four or five subsites and GH 11 has at least seven subsites (Dodd, D., and Cann., 2009).

Xylan-1,4-*β*-xylosidase (*β*-xylosidase) release xylose monomers from the non-reducing end of xylo-oligosaccharides. Xylanases break down xylan into shorter fragments increasing the number of non-reducing ends available for hydrolysis by β-xylosidases. These β-xylosidases are grouped into different GH families such as GH 3, 39, 43, 52, and 54. The most abundant β-xylosidase enzymes belong to GH 3 and 43, and they have been extensively characterized (Dodd, D., and Cann., 2009). To improve efficiency of xylan hydrolysis and yield higher levels of monosaccharides, β-xylosidase should have high tolerance for glucose and xylose. It is essential to increase activity of β-xylosidase to reduce production cost and the amount of enzyme used in biofuel industry. However, most enzyme cocktails used in current commercial scale production are ineffective (Ye, Y.L., and Zhao, 2017).

Xylanases have a wide variety of industrial applications, including paper production, and have been utilized in production of xylose, xylobiose and xylo-oligomers. A majority of these industrial applications of xylanase are performed at elevated temperatures to minimize potential contamination and decrease viscosity of substrates and products. In some cases, it also allows faster rates of reaction by increasing the solubility (Egorova and Antranikian, 2005). Therefore, identifying thermostable xylanase enzymes has received great attention.

Thermophiles for the Production of Biofuels

Thermophilic microorganisms (optimal growth temperature 45 to 80℃) are a preferred source of thermostable catalysts used in industrial applications. They can produce a variety of enzymes including amylases, phosphatases, cellulases, hemicellulases, and a few other enzymes. The enzymes derived from thermophiles have greater resistance to denaturing agents and tolerance for high pressure compared to mesophilic enzymes. Thermostable enzymes are beneficial for industrial applications for the production of value-added products and biofuels. These enzymes have greater thermostability and are less prone to contamination (Bala, A., and Singh, 2018). Due to the importance of thermophiles and their enzymes, extensive research has been conducted to identify their properties and possible applications (Table 1.2).

Table 1.2: Different thermophilic enzymes. (Atalah et al., 2019).

Type of Enzyme	*Organism*	*Enzyme Expression*	*Optimal temp*	*Optimal pH*	*Thermostability (Half-life)*
Lipase (EC 3.1.1.3)	*Pyrococcus furiosus*	Recombinant	80°C	7	6 h at 75 °C
	Bacillus sp. HT19	Recombinant	70°C	9	2 h at 70 °C
	Ureibacillus thermosphaericus	Recombinant	80°C	8	> 6 h at 70 °C
	Stenotrophomonas maltophilia Psi1	Recombinant	70°C	8	N/A
	Archaeoglobus fulgidus	Recombinant	75°C	10	N/A
Xylanase (EC 3.2.1.8)	*Dictyoglomus thermophilum*	Recombinant	75 °C	6.5	> 10 h at 80 °C
	Nonomuraea flexuosa	Recombinant	80°C	8	4.5 h at 80 °C
	Caldicellulosiruptor owensensis	Recombinant	90°C	7	1 h at 80 °C
	Thermotoga neapolitana	Recombinant	102°C	5.5-6	2 h at 100 °C
	Pyrodictium abyssi	Native	110°C	5.5	1.2 h at 105 °C
Nitrilase (EC 3.5.5.1)	*Bacillus pallidus*	Native	65°C	7.6	2.5 h a 60 °C
	Thermotoga maritima MSB8	Recombinant	45°C	7.5	1 h at 75 °C
	Geobacillus pallidus	Recombinant	50°C	7	5 h at 60 °C
	Pyrococcus abyssi	Recombinant	80°C	7.4	6 h at 90 °C
	Pyrococcus sp. M24D13	Native	90°C	7	8 h at 90 °C

Transaminase (EC 2.6.1.X)	*Thermomicrobium roseum*	Recombinant	80 ℃	7.5	5 h at 70 ℃
	Geobacillus thermodenitrificans	Recombinant	65℃	9	N/A
	Alvidobulum sp. SLM16	Recombinant	65℃	9.5	22 days at 50 ℃
	Vulcanisaeta moutnovskia	Recombinant	90℃	8	> 5 h at 70 ℃
	Cloned from a Metagenome	Recombinant	80℃	9	> 7 days at 80 ℃
GDH (EC 1.4.1.2)	*Bacillus sp.*	Native	65 ℃	8	> 8 h at 65 ℃
GDH (EC 1.4.1.3)	*Pyrococcus furiosus*	Native	85 ℃	N/A	10.5 h at 100 ℃
	Thermotoga maritima	Recombinant	75 ℃	N/A	1.8 h at 85 ℃
GDH (EC 1.4.1.4)	*Thermococcus litoralis*	Native	95 ℃	8	2 h at 100 ℃
	Aeropyrum pernix K1	Native	95 ℃	7	> 5 h at 100 ℃
Laccase (EC 1.3.10.2)	*Thermus thermophile*	Recombinant	92 ℃	5.5	14 h at 85 ℃
	Aquifex aeolicus	Recombinant	75 ℃	7	1 h at 80 ℃
	Thermobacullum terrenum	Recombinant	60 ℃	7	8 h at 80 ℃
	Bacillus sp. PC-3	Native	60 ℃	7	3.75 h at 60 ℃
	Bacillus sp. FNT	Native	70 ℃	7	3 h at 70 ℃

Thermophilic Bacterium *Caldicellulosiruptor saccharolyticus*

There has been a great interest in thermophilic species of the *Caldicellulosiruptor* genus because of their ability to convert lignocellulosic biomass into fermentable sugar molecules by producing effective carbohydrate degrading enzymes. *Caldicellulosiruptor* species are distributed around the world and have been isolated from New Zealand hot springs, Iceland, Russia, and North America (Rainey et al., 1994). The optimal growth temperature of these species typically ranges between 65-70°C, with the enzymes from these organisms active near these temperatures as well.

Thermophilic species of the *Caldicellulosiruptor* genus tend to produce more than one type of carbohydrate-degrading enzymes. The genome of *Caldicellulosiruptor saccharolyticus* has been sequenced and it was the first species within this genus that was transcriptionally analyzed to investigate how entire genome responds to various carbon sources (Blumer-Schuette et al., 2013). In addition to producing a wide range of glycoside hydrolase (GH) enzymes, *C. saccharolyticus* is capable of generating hydrogen grown on monosaccharides (van Niel et al., 2002; Kádár et al., 2003; de Vrije et al., 2007) and wastepaper. In fact, *Caldicellulosiruptor* genus has been shown to co-ferment C_5 (xylose) and C_6 sugars (glucose) (Kádár et al., 2003).

To identify potential Carbohydrate Active enZYmes (CAZYs) and their functionality, one can investigate sequenced genomes of thermophiles and other species. The genome and transcriptional studies indicate that *C. saccharolyticus* is capable of encoding a wide variety of GHs including xylanases, (VanFossen et al., 2011) glucanases, mannanases (Morris et al., 1995), in addition to side chain degrading enzymes (Lim et al., 2010) belonging to various GH families (Table 1.3). These enzymes tend to form complex structures by combining with each other, which makes it difficult to isolate native enzymes homogeneously. In this case, recombinant DNA

technology has been a useful technique allowing isolation of pure enzymes expressed in *E. coli* bacterial hosts (Zhengqiang et al., 2001).

Table 1.3: Glycoside hydrolases encoded in the *C. saccharolyticus* genome (VanFossen et al., 2011)

Locus tag	*Glycoside Hydrolase*	*GH family*	*Molecular Weight kDa*
Glucanases			
Csac_1076	1,4-β-Glucanase (CelA)	GH 48 & 9	193.7
Csac_1077[*]	1,4- β -Glucanase (ManA)	GH5 & 44	146.9
Csac_1078[Δ]	1,4-β-Glucanase (CelB)	GH10 & GH5	117.6
Csac_1079	1,4-β-Glucanase (CelC)	GH9	123.8
Csac_1085	Endoglucanase	GH74	101.7
Csac_1089	β -Glucosidase (BglA)	GH1	53.2
Csac_2548	Endoglucanase	GH16	125.9
Xylanases			
Csac_0204	Endo-1,4-β -xylanase	GH10	46.9
Csac_0678	Endo-1,4-β -xylanase	GH5	84.7
Csac_0696	Endo-1,4-β -xylanase	GH10	78.8
Csac_2404	β -Xylosidase (XynC)	GH39	56.4
Csac_2405	Endo-1,4-β-xylanase (ORF4)	GH10	36.5
Csac_2408	Endo-1,4-β-xylanase (XynA)	GH10	40.5
Csac_2409	Xylan 1,4-β -xylosidase (XynD)	GH39	58.5
Csac_2410	Endo-1,4-β-xylanase (XynE)	GH10	80.9
Csac_2732	β-Xylosidase	GH39	35.7
Mannanases			
Csac_0129	β -Mannosidase	GH2	95.7
Csac_0137	β -1,4-Mannanase	GH5	39.0
Csac_0663	β -1,4-Mannanase	GH26	52.9
Csac_1080	β-1,4-Mannanase (ManB)	GH5	27.5
Csac_2528	β -1,4-D-Mannanase	GH5	71.1
Sidechain degrading			
Csac_0359	β-Xylosidase/α-N-arabinofuranosidase	GH43	60.6
Csac_0362	β-Galactosidase	GH2	118.9
Csac_1018	β-Galactosidase	GH42	79.2
Csac_1118	α-Galactosidase (clan GH-D)	GH36	84.5
Csac_1561	α-N-arabinofuranosidase	GH51	57.9
Csac_2411	α-L-Arabinofuranosidase (XynF)	GH43 & 43	152.0
Pectin Acting			
Csac_0360	Unsaturated rhamnogalacturonyl hydrolase	GH105 or 88	43.6
Csac_0361	Galacturan 1,4-α-galacturonidase	GH28	50.0
Csac_0664	Galacturan 1,4-α-galacturonidase	GH28	50.8

Csac_1105	α-L-rhamnosidase N-terminal domain	GH78	58.0
Csac_1107	α -L-Rhamnosidase C-terminal domain	GH78	47.7
Csac_1560	Endo-1,5-α-L-arabinanase	GH43	55.4
Csac_2527	Endo-1,5-α-L-arabinanase	GH43	38.4
Csac_2689	α-Glucuronidase	GH67	80.6
Csac_2730	Unsaturated rhamnogalacturonyl hydrolase	GH105 or 88	44.6

*Csac_1077 also acts as mannanase.

ΔCsac_1078 also acts as xylanase

Research Goals

This research project focuses on isolation and characterization of the glycoside hydrolase enzyme Csac_2409, putative xylan 1,4-β-xylosidase (XynD), from thermophilic bacterium *Caldicellulosiruptor saccharolyticus* to assess potential improvements for hemicellulose pretreatment method for biofuel production. This research aims to evaluate the ability of β-xylosidase from *C. saccharolyticus* to degrade xylan, a major component of hemicellulose into fermentable sugars. Understanding the functions and optimal conditions for enzyme pretreatment could help to reduce the operational costs and high energy consumption of lignocellulosic bioethanol production.

In this project, I made an effort to clone and express the proteins putative β- xylosidase for purification and characterization. The biochemical and biophysical characterization of purified protein was performed to determine enzyme substrate preference, pH optimum, temperature optimum, and thermal stability. Finally, enzyme kinetics were determined including Michaelis-Menten characterization.

References

Agbor, Valery B., Nazim Cicek, Richard Sparling, Alex Berlin, and David B. Levin. 2011.Biomass Pretreatment: Fundamentals toward Application. Biotechnology Advances 29, no. 6: 675-85. doi: 10.1016/j.biotechadv.2011.05.005.

Atalah, Joaquín, Paulina Cáceres-Moreno, Giannina Espina, and Jenny M. Blamey. 2019.Thermophiles and the Applications of Their Enzymes as New Biocatalysts. Bioresource Technology 280: 478-88. doi: 10.1016/j.biortech.2019.02.008.

Bala, Anju, and Bijender Singh. 2019.Cellulolytic and Xylanolytic Enzymes of Thermophiles for the Production of Renewable Biofuels. Renewable Energy 136: 1231-244. doi: 10.1016/j.renene.2018.09.100.

Balan, Venkatesh. 2014. Current Challenges in Commercially Producing Biofuels from Lignocellulosic Biomass.ISRN Biotechnology 2014: 1-31. doi:10.1155/2014/463074

Beauchet, R., F. Monteil-Rivera, and J.m. Lavoie. 2012.Conversion of Lignin to Aromatic-based Chemicals (L-chems) and Biofuels (L-fuels). Bioresource Technology 121: 328-34. doi: 10.1016/j.biortech.2012.06.061.

Bensah, Edem Cudjoe, and Moses Mensah.2013. Chemical Pretreatment Methods for the Production of Cellulosic Ethanol: Technologies and Innovations.International Journal of Chemical Engineering, vol. 2013, Article ID 719607

Blumer-Schuette, S. E., Brown, S. D., Sander, K. B., Bayer, E. A., Kataeva, I., Zurawski, J. V., Kelly, R. M. 2014. Thermophilic lignocellulose deconstruction. FEMS Microbiol Rev, 38: 393–448. doi: 10.1111/1574-6976.12044

Börjesson, Mikaela, and Gunnar Westman.2015. Crystalline Nanocellulose — Preparation, Modification, and Properties. Cellulose - Fundamental Aspects and Current Trends. doi:10.5772/61899

Brennan, L. and Owende, P.2010.Biofuels from microalgae: A review of technologies for production, processing, and extractions of biofuels and coproducts. Renewable & Sustainable Energy Reviews, 14, 557-577. doi:10.1016/j.rser.2009.10.00

Cannatelli, Mark D., and Arthur J. Ragauskas. 2016. Conversion of Lignin into Value-added Materials and Chemicals via Laccase-assisted Copolymerization. Applied Microbiology and Biotechnology 100, no. 20: 8685-691. doi:10.1007/s00253-016-7820-1.

Caspeta, L., Buijs, N. A., & Nielsen, J.2013. The role of biofuels in the future energy supply. *Energy & Environmental Science,6*(4), 1077. doi:10.1039/c3ee24403

Cheng, P., Wang, Y., Liu, T., & Liu, D.2017. Biofilm Attached Cultivation of *Chlorella pyrenoidosa* Is a Developed System for Swine Wastewater Treatment and Lipid Production. *Frontiers in plant science*, *8*, 1594. doi:10.3389/fpls.2017.01594

Clarke, Steve, and Fernando Preto. 2011.Biomass Densification for Energy Production. Accessed May 23, 2019. http://www.omafra.gov.on.ca/english/engineer/facts/11-035.htm.

Collins, T., Gerday, C., & Feller, G. 2005. Xylanases, xylanase families and extremophilic xylanases. FEMS Microbiol Rev, 29(1): 3-23. doi: 10.1016/j.femsre.2004.06.005

De Vrije T, Mars AE, Budde MA, Lai MH, Dijkema C, de Waard P, et al.2007. Glycolytic pathway and hydro-gen yield studies of the extreme thermophileCaldicellulosiruptor saccharolyticus. Appl Microbiol Bio-technol; 74: 1358–1367.

Dodd, Dylan, and Isaac K. O. Cann.2009. Enzymatic Deconstruction of Xylan for Biofuel Production. GCB Bioenergy1, no. 1: 2-17. doi:10.1111/j.1757-1707.2009.01004. x.

Dutta, Saikat, and Kevin C.-W. Wu. 2014.Enzymatic Breakdown of Biomass: Enzyme Active Sites, Immobilization, and Biofuel Production." *ChemInform* 45, no. 51. doi:10.1002/chin.201451258.

Egorova K & Antranikian G.2005. Industrial relevance of thermophilic Archaea. Curr Opin Microbiol 8: 649–655.

Energy Information Administration.2019. Ethanol and the Environment-Energy Explained, Your Guide to Understanding Energy. Accessed May 27, 2019. https://www.eia.gov/energyexplained/index.php?page=biofuel_ethanol_environment.

Energy Protection Agency.2019. Inventory of U.S Greenhouse Gas Emissions and Sinks.

Escuder-Rodríguez, Juan-José, María-Eugenia DeCastro, María-Esperanza Cerdán, Esther Rodríguez-Belmonte, Manuel Becerra, and María-Isabel González-Siso.2018. Cellulases from Thermophiles Found by Metagenomics.*MPDI* 6, no. 66. doi:10.3390/microorganisms6030066.

European Commission, Joint Research Centre, Institute for Energy and Transport, Sustainable Transport Unit.2015. The impact of biofuels on transport and the environment, and their connection with agricultural development in Europe.

Global Energy & CO_2 Status Report. GECO, 2019

Henrissat, Bernard, and Gideon J. Davies.2000. Glycoside Hydrolases and Glycosyltransferases. Families, Modules, and Implications for Genomics.Plant Physiology 124, no. 4 1515-519. doi:10.1104/pp.124.4.1515.

Hood, Elizabeth E.2016. Plant-based Biofuels.F1000Research5: 185. doi:10.12688/f1000research.7418.1.

Hu, Jin, Fei Yu, and Yongwu Lu. 2012.Application of Fischer–Tropsch Synthesis in Biomass to Liquid Conversion. Catalysts 2, no. 2: 303-26. doi:10.3390/catal2020303.

Hu, Jinguang, Valdeir Arantes, Amadeus Pribowo, and Jack N. Saddler. 2013.The Synergistic Action of Accessory Enzymes Enhances the Hydrolytic Potential of a “cellulase Mixture” but Is Highly Substrate Specific. Biotechnology for Biofuels 6, no. 1: 112. doi:10.1186/1754-6834-6-112.

International Energy Agency.2008. From 1st to 2nd Generation Biofuel Technologies Extended Executive Summary.

Isikgor, Furkan H., and C. Remzi Becer. 2015. Lignocellulosic Biomass: A Sustainable Platform for the Production of Bio-based Chemicals and Polymers. Polymer Chemistry 6, no. 25: 4497-559. doi:10.1039/c5py00263j.

Kádár, Z., Vrije, T. D., Budde, M. A., Szengyel, Z., Réczey, K., & Claassen, P. A. 2003. Hydrogen Production from Paper Sludge Hydrolysate. Biotechnol. Fuels Chemicals, 557-566. doi:10.1007/978-1-4612-0057-4_46

Kuhad, Ramesh Chander, Rishi Gupta, and Ajay Singh. 2011.Microbial Cellulases and Their Industrial Applications. Enzyme Research 2011: 1-10. doi:10.4061/2011/280696.

Lee, R A, and Jean-Michel Lavoie.2013. From First- to Third-Generation Biofuels: Challenges of Producing a Commodity from a Biomass of Increasing Complexity.*Animal Frontiers*. doi:10.2527/af.2013-0010.

Liang, Y., Sarkany, N., Cui, Y.2009. Biomass and lipid productivities of <i>Chlorella vulgaris</i> under autotrophic, heterotrophic and mixotrophic growth conditions. Biotechnology Letters, 31(7), 1043-1049.

Lim, Y., Yoon, R., Seo, E., Kim, Y., Park, C., & Oh, D. 2010. Hydrolytic properties of a thermostable α-l-arabinofuranosidase from Caldicellulosiruptor saccharolyticus. Journal of Applied Microbiology, *109*(4), 1188-1197. doi:10.1111/j.1365-2672.2010. 04744.x

Maehara, T., Yagi, H., Sato, T., Ohnishi-Kameyama, M., Fujimoto, Z., Kamino, K., Kaneko, S. 2017. GH30 Glucuronoxylan-Specific Xylanase from Streptomyces turgidiscabies C56. Appl. Environ. Microbiol. 84(4). doi:10.1128/aem.01850-17

Moraïs, Sarah, Yoav Barak, Raphael Lamed, David B. Wilson, Qi Xu, Michael E. Himmel, and Edward A. Bayer. 2012.Paradigmatic Status of an Endo- and Exoglucanase and Its Effect on Crystalline Cellulose Degradation. Biotechnology for Biofuels 5, no. 1: 78. doi:10.1186/1754-6834-5-78.

Morris, D. D., R.A. Reeves, M.D. Gibbs, D.J. Saul, and P. Bergquist.1995. Correction of the the β-mannanase domain of the celC pseudogene from Caldocellulosiruptor saccharolyticus and activity of the gene product on kraft pulp. Applied and Environmental Microbiology 61: 2262–2269.

Mosier, N, C. Wyman, B. Dale et al.2005. Features of promising technologies for pretreatment of lignocellulosic biomass,*Bioresource Technology*, vol. 96, no. 6, pp. 673–686 .

National Renewable Energy Laboratory (NREL).2006. Research Review

Rainey, F., Donnison, A., Janssen, P., Saul, D., Rodrigo, A., Bergquist, P., Morgan, H. 1994. Description of Caldicellulosiruptor saccharolyticus gen. nov., sp. nov: an obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol Lett, 120(3): 263–266.

Sammond, Deanne W., Noah Kastelowitz, Michael E. Himmel, Hang Yin, Michael F. Crowley, and Yannick J. Bomble. 2016. Comparing Residue Clusters from Thermophilic and Mesophilic Enzymes Reveals Adaptive Mechanisms. Plos One 11, no. 1. doi: 10.1371/journal.pone.0145848.

Shabtai, J., W. Zmierczak, and E. Chornet. 1999. Process for conversion of lignin to reformulated hydrocarbon gasoline: The University of Utah Research Foundation.

Smith, Peter J., Hsin-Tzu Wang, William S. York, Maria J. Peña, and Breeanna R. Urbanowicz. 2017. Designer Biomass for Next-generation Biorefineries: Leveraging Recent Insights into Xylan Structure and Biosynthesis. Biotechnology for Biofuels 10, no. 1. doi:10.1186/s13068-017-0973-z.

van Niel E.W.J., Budde M.A.W., Haas G.G. de., van der Wal F.J., Claassen P.A.M., Stams A.J.M. 2002. Distinctive properties of high hydrogen producing extreme thermophiles, *Caldicellulosiruptor saccharolyticus* and Thermotoga elfii. Int J Hydrogen Energy 27: 1391–1398.

Vanfossen, A. L., Ozdemir, I., Zelin, S. L., & Kelly, R. M.2011. Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus. Biotechnol Bioengin, 108(7): 1559-1569. doi:10.1002/bit.23093

Vohra, Mustafa, Jagdish Manwar, Rahul Manmode, Satish Padgilwar, and Sanjay Patil. 2014.Bioethanol Production: Feedstock and Current Technologies. *Journal of Environmental Chemical Engineering*2, no. 1: 573-84. doi: 10.1016/j.jece.2013.10.013.

Wyman C. E. 2004.Ethanol fuel. In: Encyclopaedia of energy. New York: Elsevier Inc.

Ye, Yanxin, Xuezhi Li, and Jian Zhao. 2017. Production and Characteristics of a Novel Xylose- and Alkali-tolerant GH 43 β-xylosidase from Penicillium Oxalicum for Promoting Hemicellulose Degradation. isi*Scientific Reports*7, no. 1. doi:10.1038/s41598-017-11573-7.

Zhang, Qi, Jie Chang, Tiejun Wang, and Ying Xu.2007.Review of Biomass Pyrolysis Oil Properties and Upgrading Research. Energy Conversion and Management 48, no. 1:87-92. doi: 10.1016/j.enconman.2006.05.010

Zhang, Xiao-Zhou, and Yi-Heng Percival Zhang.2013.Cellulases: Characteristics, Sources, Production, and Applications. In Bioprocessing Technologies in Biorefinery for Sustainable Production of Fuels, Chemicals, and Polymers. First ed. John Wiley & Sons.

Zhang, Yuzhen, Xiaoguo Fu, and Hongzhang Chen.2012. Pretreatment Based on Two-step Steam Explosion Combined with an Intermediate Separation of Fiber Cells-Optimization of Fermentation of Corn Straw Hydrolysates. Bioresource Technology 121: 100-04. doi: 10.1016/j.biortech.2012.07.006.

Zhengqiang, J., Kobayashi, A., Ahsan, M. M., Lite, L., Kitaoka, M., & Hayashi, K.2001. Characterization of a Thermostable Family 10 Endo-Xylanase (XynB) from Thermotoga maritima That Cleaves p-Nitrophenyl-. BETA.-D-Xyloside. J Biosci Bioengin, 92(5): 423-428. doi:10.1263/jbb.92.423

Zhou, S., and L. O. Ingram.2000.Synergistic Hydrolysis of Carboxymethyl Cellulose and Acid-Swollen Cellulose by Two Endoglucanases (CelZ and CelY) from Erwinia Chrysanthemi. Journal of Bacteriology 182, no. 20: 5676-682. doi:10.1128/jb.182.20.5676-5682.2000

CHAPTER 2

CHARACTERIZATION OF β-XYLOSIDASE FROM THE THERMOPHILIC BATERIUM *Caldicellulosiruptor saccharolyticus*

Introduction

Production of biofuel from biomass is considered a promising way to reduce dependence of fossil fuels used for transportation. However, current fermentation techniques rely upon carbon sources that also serve as food, mainly starch based materials. There is a push to move to lignocellulosic based ethanol. The main challenge at present is an efficient and environmentally friendly method to break lignocellulosic components into simple, fermentable sugars.

Xylan is the most abundant hemicellulose, consisting of a repeating *β*-1,4-linked xylose backbone substituted with acetyl, arabinofuranosyl, and 4-*O*-methyl glucuronyl groups. Since xylan is mainly composed of pentose sugars like xylose and arabinose, hemicellulose plays a significant role in biofuel production. However, this requires complete degradation of xylan carried out by a variety of cooperatively acting enzymes (Dodd, D., and Cann, 2009)

The depolymerization of xylan is carried out by endo-1,4-β-xylanase (xylanase) which cleaves the β-1,4 glycosidic linkage between xylose residues in the backbone of xylan. Xylan-1,4-β-xylosidase (β-xylosidase) releases xylose monomers from the non-reducing end of xylo-oligosaccharides. Therefore, xylanases are required to break down xylan polymers into shorter fragments, increasing the number of non-reducing ends available for hydrolysis by β-xylosidase (Dodd, Dylan, and Cann, 2009). β-xylosidase enzymes which can be utilized in biofuel production have been isolated and characterized from a variety of microorganisms (Cintra et al., 2017; Li et al., 2018; Shin et al., 2014).

Thermostable xylanases possess advantages for industrial application of xylan degrading enzymes. Therefore, thermophilic enzymes are a preferred option for bioconversion of lignocellulosic biomass in harsh processing conditions.

Of particular interest in these studies is *Caldicellulosiruptor saccharolyticus* which can utilize a range of simple sugars in addition to hemicelluloses as sole carbon sources including xylan, xylose, pectin, and arabinose. This gram-positive extremely thermophilic bacterium grows optimally at 70°C producing effective carbohydrate degrading enzymes in harsh bioprocessing conditions. The genome of *C. saccharolyticus* encodes different glycoside hydrolases (Van de Werken et al., 2008) which are primarily hydrolytic enzymes and have been categorized into GH families.

This research aims to investigate glycosidic hydrolases derived from extremely thermophilic *C. saccharolyticus*. In this work, isolation of endo-xylanase enzymes encoded by Csac_2405 and Csac_2408 of GH10 and the β-xylosidase encoded by Csac_2409 of GH39 were investigated with biochemical characterization of Csac_2409. In particular, *optimum temperature and pH, thermal stability, substrate specificity, and enzyme kinetic characterization of the β-xylosidase* were determined to assess improved pretreatment methods for biofuel production. To study enzyme activity, synthetic substrate *p*NPX (4-Nitrophenyl β-D-xylopyranoside) was used as a substitution for natural substrates due to its high optical sensitivity.

Materials and Methods

Cloning, Expression, and Purification

Overall strategy for protein expression in bacterial system consists of introducing the gene of interest to competent *E. coli* cells and allowing them to transcribe and translate the protein. The targeted protein can then be obtained from lysed cell pellets by protein purification method. The standard SDS-PAGE and Bradford methods can be used to determine purity and

concentration of the purified protein. Gene cloning and protein expression using standard techniques and multiple kits were attempted, but proved unsuccessful, with efforts described in Chapter 3. Given these challenges in obtaining successful expression of the protein, a synthetic gene was synthesized and cloned by biology research company GenScript, USA and the protein expressed and purified to about 75% purity.

Gene Synthesis, Cloning, and Expression by GenScript

The approach by GenScript (Figure 2.1) involves several steps to synthesize DNA strand chemically base-by-base and create a complementary strand. Once the DNA sequence was selected, subsequent codon optimization is essential to improve heterologous protein expression level. This process involves modifying DNA sequence while keeping the amino acid sequence intact. In this case, *GenScript's OptimumGene*™ PSO-driven algorithm has been used to optimize sequences efficiently. After completing the required sequence, it has been analyzed to determine several factors including oligo length, GC content, sequence repeats, and hairpin formations. In order to fabricate DNA, nucleotide monomers have been added stepwise through phosphoramidite chemistry.

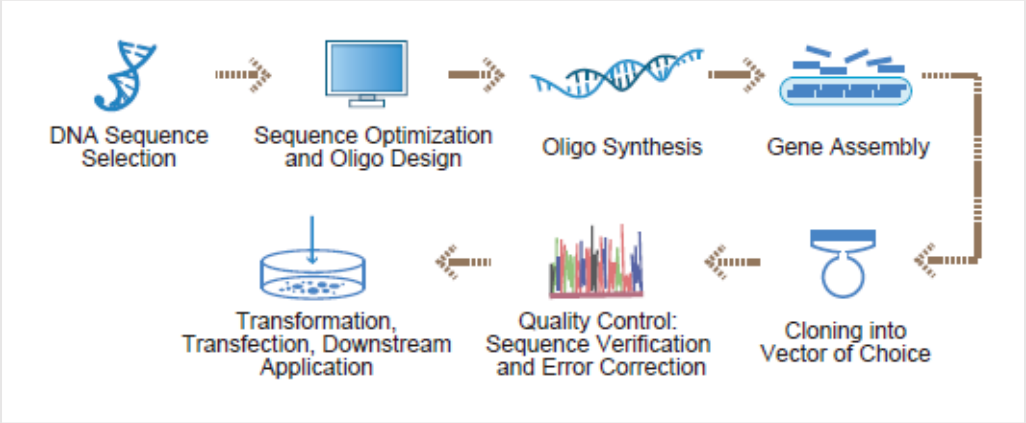


Figure 2.1: Steps in Gene Synthesis (Gene Synthesis handbook 3rd Edition, GenScript). The genetic sequence was optimized for heterologous protein, cloned into a pET expression vector, sequenced, and transferred to *E. coli* for downstream processing.

Once the short oligonucleotides were formed, they were assembled into the complete gene. The synthesized sequence was cloned into the pET-30a (+) vector with NdeI and HindIII restriction sites encoding His_6-tag (Appendix A) for protein expression in *E. coli*. To evaluate expression, the recombinant plasmid was transformed to *E. coli* strain BL21 (DE3) and an individual colony was inoculated into Terrific Broth (TB) media (37°C at 200 rpm) with appropriate antibiotic, kanamycin. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce protein expression.

For scale up expression, recombinant BL21(DE3) has been inoculated into TB medium with antibiotics and incubated at 37°C. Once the optical density (OD_{600}) of bacterial cultures reached about 4, cultures were induced with IPTG at 15°C for 16 hours. These cultures with expressed protein were centrifuged to isolate the cells, the cells were resuspended with lysis buffer, followed by sonication to lyse the cells releasing the soluble, intracellular protein. Nickel-Nitrilotriacetic Acid (Ni-NTA) was used as the primary step in protein purification. Bradford protein assay with BSA as standard was used to determine protein concentration and the purity was evaluated using SDS-PAGE and Western blotting.

Protein Characterization

Standard curve of p-nitrophenol

The synthetic substrate *p*NPX (*p*-nitrophenol-β-D-xylopyranoside) is used for enzymatic reaction, during which a chromophore (*p*NP) is released due to cleavage by the enzyme. This *p*NP absorbs light optimally at 405 nm, which can be used to measure product concentration by spectrophometry. The unknown product concentration is determined using a calibration curve of *p*-nitrophenol from 0 to 1 mM pNP. For this, a set of known standards are measured in series using Multiskan FC microplate photometer (Eppendorf A.G, 2011). Plates were filled to 150µL for both standard and unknown samples.

Substrate Specificity and Enzyme Activity

Substrate specificity of enzyme was determined using different *p*NP substrates: *p*NP-β-D xylopyranoside, *p*NP-α-D xylopyranoside, *p*NP-β-glucopyranoside, *p*NP-α-glucopyranoside, and *p*NP-β-mannopyranoside (Sigma-Aldrich). Conditions for substrate characterization were 70°C, pH 6.5 in sodium acetate buffer. The preferred substrate was then used in subsequent characterization studies for pH and temperature optima and full kinetic characterization and determination of the Michealis-Menten kinetic parameters (k_{cat} and K_M) using 0 mM -1 mM of *p*NPX as substrate. For this, nonlinear regression curve fitting of the data to the following Michaelis-Menten equation (Equation 1) was applied.

$$V_0 = \frac{V_{max}[S]}{K_m + [S]} \quad (1)$$

Determination of pH and Thermal Profiles

The optimum pH of the enzymes was determined by measuring enzyme activities at 70°C using the preferred synthetic substrate. Eight buffer solutions were prepared ranging from pH 4.0 to 7.5 with an interval of 0.5 pH units between each and used to assay the activity of the enzyme. For this, 50 mM sodium citrate (pH 4-6), sodium acetate (pH 5-6), and phosphate (pH 6-7.5) buffers were used. Once the optimum pH was determined, optimum temperature of the enzyme was evaluated using the standard assay at optimum pH level with selected temperatures between 70 and 90°C.

To evaluate the thermal stability, the enzyme was heat-treated for times ranging between 0 and 180 minutes at 70°C and 80°C. The standard assay method was utilized to determine the enzyme activity. The reactions were performed in triplicate.

Enzyme Assay

β-xylosidase activity was assessed using *p*NPX as substrate in 50 mM sodium citrate buffer at pH 5.5. The total volume of 0.5 mL reaction mixture consisted of *p*NPX (1mM) and 10 µL of enzyme. The substrate *p*NPX, citrate buffer and the enzyme were pre-heated at 70°C for 5 minutes in water bath. The reaction was initiated by adding enzyme followed by incubation for 5 minutes. The reaction was terminated by incubating in an ice bath immediately after 5 minutes. The liberated *p*-nitrophenol was measured using spectrophometry at 405 nm.

Results and Discussion

The purity and molecular weight were determined using SDS-PAGE and western blotting (Figure 2.2). The purity of xylosidase protein was estimated to be about 75%. The 1551 bp gene showed a band at 59 kDa which is consistent with reported mass of Xylan 1,4-β-xylosidase (https://www.uniprot.org/uniprot/A4XM51). The concentration of the purified protein determined by Bradford protein assay with BSA as standard is 0.91 mg/ml (Figure 2.3).

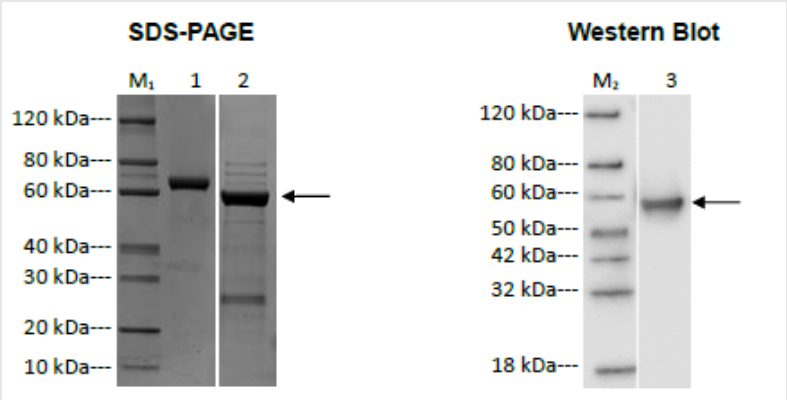


Figure 2.2: SDS-PAGE and Western Blot analysis. M1: Protein Marker, (GenScript, Cat. No. M00516) M2: Protein Marker, (GenScript, Cat. No. M00521), Lane 1: BSA (2.00 µg), Lane 2: Xylosidase (Reducing condition, 2.00 µg), Lane 3: Xylosidase (Reducing condition) Primary antibody: Mouse-anti-His mAb

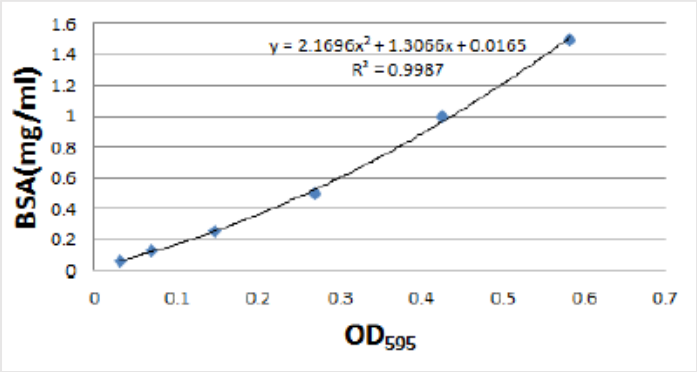


Figure 2.3: Standard curve for Bradford assay

Standard curve of *p*-nitrophenol (Figure 2.4) was created to determine the concentration of *p*NP released during the enzymatic hydrolysis. The data points were fitted with line and the resulting equation provides the concentration of *p*NP per unit absorbance of 0.54 mM.

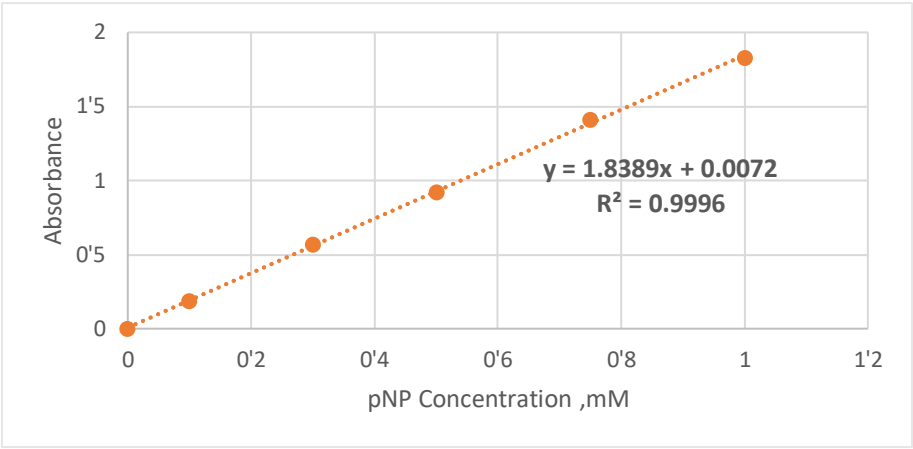


Figure 2.4: Standard curve of *p*-nitrophenol

Substrate specificity of β-xylosidase was determined using synthetic *p*NP linked substrates. *p*-nitrophenyl group is linked to monosaccharide group via a glyosidic bond which cleaves during enzymatic hydrolysis releasing *p*-nitrophenyl. The liberation of this compound can be detected due to visual color change of the solution to yellow and measured at 405 nm. Control

samples were run for all reactions and the results have been normalized to the activity of *p*NP-β-D xylopyranoside (Figure 2.5).

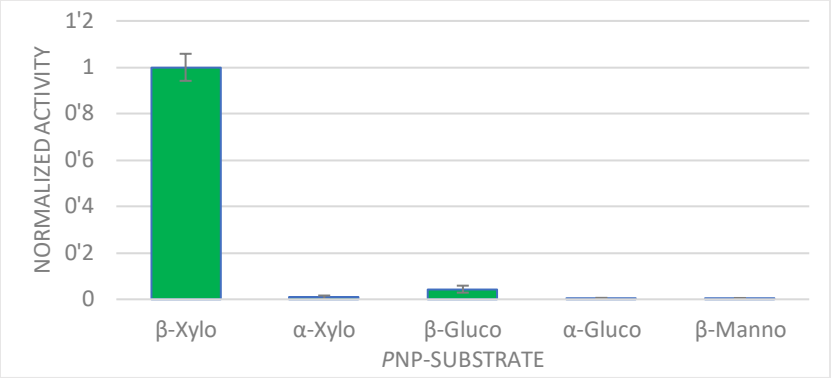
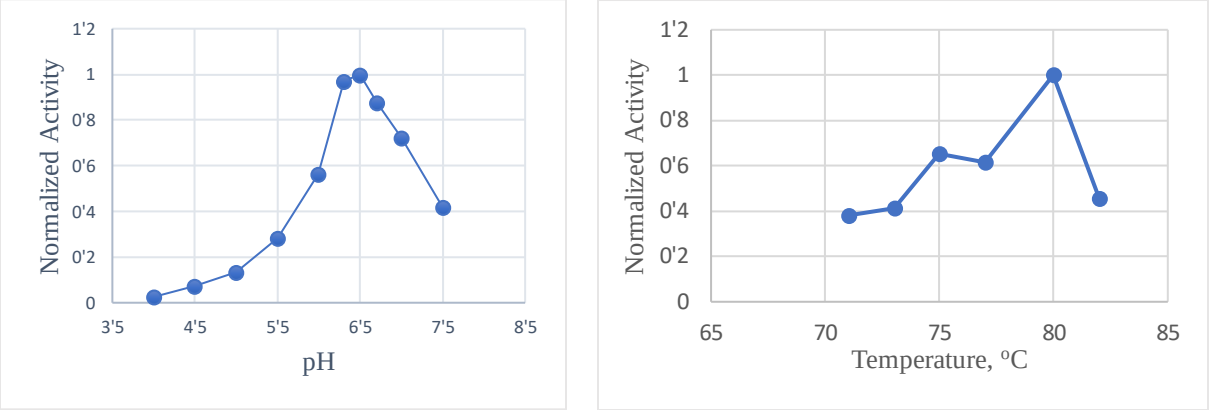


Figure 2.5: Substrate Specificity of β-xylosidase on different synthetic *p*NP substrates. Enzyme activities are normalized to the maximum activity and each bar indicates the averages of triplicates.

The enzyme exhibited the greatest activity on *p*NP-β-D xylopyranoside and lower activity toward *p*NP-β-D glucopyranoside. However, there were no detectable activities observed on *p*NP-α-D xylopyranoside, *p*NP-α-D glucopyranoside, and *p*NP-β-D mannopyranoside. The activity towards *p*NP-β-D xylopyranoside indicates the high selectivity of the enzyme to degrade glycosidic bond between β-xylopyranoside and *p*NP and confirms the putative characterization of the gene as a β-xylosidase. The structural and conformational similarities between *p*NP-β-D glucopyranoside and *p*NP-β-D xylopyranoside justify the significant activity on *p*NP-β-D glucopyranoside and has been observed for β-xylosidases from different organisms (Li et al., 2018; Kirikyali et al., 2014).

β-xylosidase exhibited activity in a range of pH (5 – 7.5) with the optimum around 6.5 (Figure 2.6 A). The reported pH value was similar to that of β-xylosidase of GH39 from thermophilic bacterium *Geobacillus* sp (Bhalla et al., 2014). β-xylosidase was active between 70℃ and 85℃ with highest activity at 80℃ (Figure 2.6 B). The activity increased steadily until 80℃ and decreased with further temperature increase. At the optimal temperature, the greatest activity of enzyme can be observed (Enzymes, rsc.org). Above this temperature, however, protein

denaturation occurs due to the loss of three-dimensional structure, which causes loss of protein functionality. Most of the proteins can be denatured by increased temperature which affects the weak interactions, such as hydrogen bonds. When the temperature is increased gradually the protein structure remains intact until it starts to break down (Nelson and Cox, 2000). The optimum temperature of β-xylosidase reported was higher than that of β-xylosidase from *Geobacillus pallidus* and *Geobacillus sp*. strain WSUCF1 (Quintero et al., 2007; Bhalla et al., 2014).



A B

Figure 2.6: pH and temperature optima of β-xylosidase (XynD). **A)** pH optima of XynD on pNPX **B)** Temperature optima of XynD on pNPX. Enzyme activities are normalized to maximum activity.

Thermal stability of β-xylosidase was determined at 70 and 80°C (Figure 2.7). The reactions were performed in triplicate and results were normalized to the reaction sample at 0 minutes. The data was analyzed with least squares non-linear exponential decay regression to evaluate half-lives. In this work, β-xylosidase indicated half-lives close to 3 hours and 1.8 hours at 70 and 80°C, respectively. Bhalla et al. (2014) reported β-xylosidase from *Geobacillus sp*. strain WSUCF1 had half-life of 9 days at 70°C and endo-1,4-β-xylanase encoded Csac_0678 indicated half-lives of 19 hours at 77°C, less than 3 hours at 80°C, and less than 30 minutes at 85°C (Ozdemir et al., 2011).

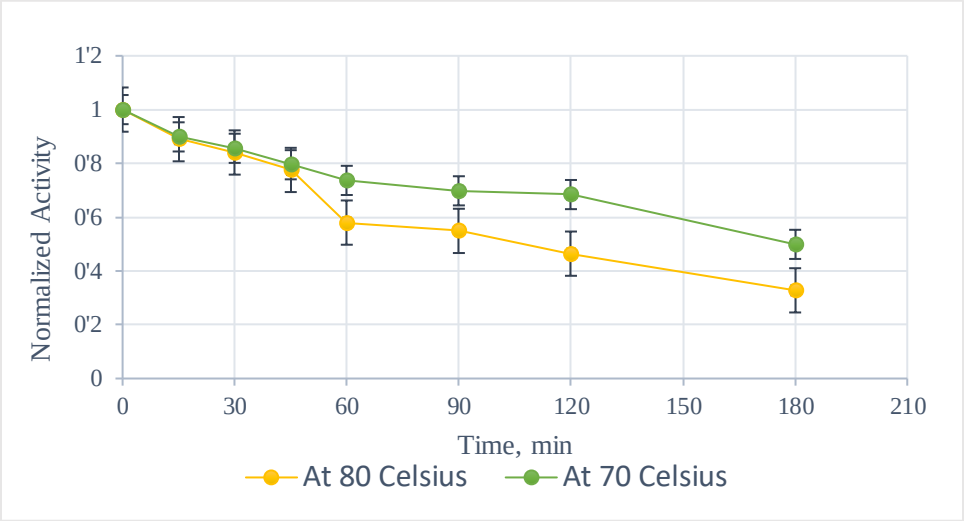


Figure 2.7: Thermal stability of β-xylosidase at 70°C and 80°C. Enzyme activities are normalized to the initial activity. Each point indicates the averages of triplicates and error bars show the standard error.

The kinetic parameters of β-xylosidase were calculated with *p*NPX. *p*NPX substrate concentrations were varied between 0 mM and 1 mM and least squares nonlinear regression was used to determine kinetic parameters (Figure 2.8). The calculated values of K_M and V_{max}, k_{cat} and k_{cat}/ K_m were 0.918 mM, 0.0042 mM/s, 13.6 s^{-1} and 14.75 $s^{-1}mM^{-1}$, respectively.

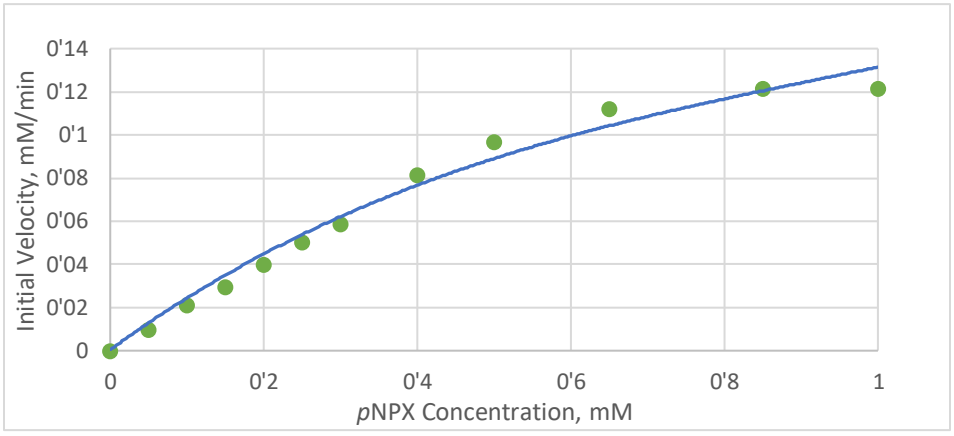


Figure 2.8: Nonlinear regression fitting to the Michealis-Menten equation. ● Experimental data. ─ Regression curve.

Michaelis-Menten equation (Equation 1) mathematically represents the relationship between substrate concentration and reaction rate under isothermal conditions for many enzyme catalyzed reactions (Johnson and Goody, 2011). Where V_0, S and V_{max} denote the initial rate, substrate concentration, and the maximum initial velocity respectively. The maximum velocity V_{max} is theoretically attained when the enzyme has been "saturated" by an infinite concentration of a substrate, and K_m is the Michaelis constant, representing a measure of affinity of the enzyme-substrate interaction. K_m is defined as the substrate concentration at half maximum initial velocity (Marasović et al., 2017).

The difference between calculated and experimental data can be observed (Figure 2.8) when using nonlinear regression to calculate enzymatic kinetic parameters. Typically, the graphical plotting (linear regression) is used to obtain the parameters of Michaelis-Menten equation (Matyska and Kovář, 1985). To this end, there are three common methods to determine the parameters such as Lineweaver-Burk plot (double reciprocal plot), Eadie-Hofstee plot, and Hane-Woolf plot.

The Michaelis-Menten equation (Equation 1) can be transformed to by taking reciprocal of both sides to represent following Lineweaver-Burk equation (Equation 2). The slope and intercept of the straight line of Lineweaver-Burk plot can then be used to calculate K_m and V_{max} (Marasović et al., 2017). However, the linear regression approach is valid only for simple linearizable functions. Furthermore, in some cases, this method does not provide accurate results (Matyska and Kovář, 1985).

$$\frac{1}{V_0} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}} \quad (2)$$

Due to the errors of linearization methods, it has been proposed to use least squares nonlinear regression to estimate parameters with high accuracy (Wilkinson, 1961). It permits determination of values directly from untransformed data points. This procedure begins with

initial estimates and then converges on an estimate in small steps to give the best fit of the fundamental model to the actual data points (Marasović et al., 2017). However, there are some deficiencies of nonlinear regression dealing with experimental data. In this case, initial assumptions made for nonlinear regression do not take into the account deviations originated due to biological errors or experimental errors including variations in the enzyme preparations, variations in measured volume of substrates and enzymes, instrumentation errors. The ordinary least-squares methods can present misleading values for the nonlinear regression because of the outliers in the data (Lim et al., 2013).

Each GH family contains enzymes which are related by their sequence, and by corollary fold and this is useful to make predictions (Cazypedia). GH family 39 is quite small group with less than 400 sequences and includes 12 mammalian α-L-iduronidases (EC 3.2.1.76) and bacterial xylosidases which have been characterized (Lagaert et al., 2014). GH39 β-xylosidases from thermophilic bacteria *T. saccharolyticum* and *G. stearothermophilus* have been analyzed for their crystal structures (Yang et al., 2004; Czjzek et al., 2005). The studies revealed that these enzymes share three domains: catalytic $(\beta/\alpha)_8$ barrel domain, β-sandwich domain, and small α-helical domain (Lagaert et al., 2014). Even though xylosidases in GH family 39 share common sequence of amino acids, the following comparison of enzymes (Table 1.1) shows that their properties vary widely including optimum conditions, thermal stability, and enzyme kinetics.

Table 2.1: Comparison of Michaelis-Menten kinetic parameters of β-xylosidases in GH family 39 for hydrolysis of *p*NPX

Enzyme	Organism	Optimal conditions	K_m	V_{max}	k_{cat} (s^{-1})	k_{cat}/K_m	Reference
XynD	*C. saccharolyticus*	pH 6.5, 80°C	0.918 mM	251 mM/min	13.6	14.75 $s^{-1}mM^{-1}$	This work
XylBH39	*Bacillus halodurans*	pH 7.5, 55°C	8.61±0.48 mM	N/A	5.26±0.78	0.6±0.12 $s^{-1}M^{-1}$	Smaali et al.,2006
β-xylosidase	*Geobacillus sp*	pH 6.5, 70°C	2.38 mM	147 U/mg	N/A	N/A	Bhalla et al.,2014
xln-DT	*Dictyoglomus thermophilum*	pH 6.0, 75°C	1.66 mM	78.46 U/mg	N/A	N/A	Li et al.,2018
xynB2	*Caulobacter crescentus*	pH 6, 55°C	9.3±0.45 mM	402±19 µmol/min	N/A	N/A	Corrêa et al.,2012
BH1068 (For pNP-XylP)	*Bacillus halodurans*	pH 6.5, 47°C	3.27±0.907µM	N/A	1.73±0.286	0.53±0.17 $s^{-1}mM^{-1}$	Wagschal et al.,2007

N/A – Not Available

K_m ranges widely from 0.0301 to 9.3 mM (Table 2.1) which can also vary from enzyme to enzyme. In some cases, the same enzyme will have varying K_m values for different substrates. The term, K_m is often used to measure the affinity of an enzyme for the substrate. However, the actual value of K_m depends on some characteristics of reaction mechanism including number and relative rates of the individual steps of the reaction. K_m cannot be used simply to measure substrate affinity and even when the reaction has several steps after formation of ES (substrate-enzyme) which leads K_m to become a very complex function of many rate constants (Nelson and Cox, 2000). However, low K_m indicates that the enzyme requires a small amount of substrate to become saturated. Thus, the maximum velocity can be reached at relatively low substrate concentration (Robinson, 2015). In fact, K_m value of this work is around 1 mM which is lower than K_m of most β-xylosidase enzymes in GH 39.

The number of reaction steps in addition to rate-limiting step vary from enzyme to enzyme. In a situation described below (Equation 3), product (P) release from enzyme-product complex ($EP \rightarrow E + P$) becomes the rate-limiting step.

$$E + S \underset{\mathbf{k_{-1}}}{\overset{\mathbf{k_1}}{\rightleftharpoons}} ES \underset{\mathbf{k_{-2}}}{\overset{\mathbf{k_2}}{\rightleftharpoons}} EP \overset{\mathbf{k_3}}{\rightleftharpoons} E + P \qquad (3)$$

Where E, S, P, ES, and EP denote the enzyme, substrate, product, enzyme-substrate complex, and enzyme-product complex respectively. In here, majority of the enzyme is in the EP complex at saturation, therefore,

$$V_{max} = k_3[E_t] \qquad (4)$$

Where k_3 and $[E_t]$ represent rate constant of rate-limiting reaction and total enzyme concentration.

The rate constant k_{cat} can be used to define the limiting rate of an enzyme-catalyzed reaction at saturation. Therefore, in a multi-step reaction, the rate constant of the rate-limiting

step would be equivalent to k_{cat} and for the reaction of equation 2, $k_{cat} = k_3$. However, when several steps in a reaction become partially rate-limiting, Michaelis-Menten equation (Equation 1) can be written in following form (Equation 5).

$$V_0 = \frac{k_{cat}[E_T][S]}{K_m + [S]} \quad (5)$$

The k_{cat} constant is the first-order rate constant, which is also called turnover number. This represents the number of substrate molecules that can be converted to product by a single enzyme molecule per unit time. k_{cat} and K_m kinetic parameters are useful to compare different enzymes to evaluate the kinetic efficiency of enzymes (Nelson and Cox, 2000; Robinson, 2015). The ratio k_{cat}/K_m is referred to as the "specific constant" which is a useful index to compare relative rates of an enzyme acting on alternative, competing substrates (Eisenthal, 2007). k_{cat}/K_m rate constant is a measure of catalytic efficiency which can be used to compare different substrate preferences of an enzyme (Berg et al., 2002). In this work, β-xylosidase indicated greater catalytic efficiency value of 14.75 $s^{-1}mM^{-1}$ compared to other GH39 β-xylosidase enzymes from different organism (Table 2.1).

References

Berg J. M, J. L Tymoczko and L Stryer. 2002.The Michaelis-Menten Model Accounts for the Kinetic Properties of Many Enzymes Biochemistry. 5th edition. New York: W H Freeman; Section 8.4. Available from: https://www.ncbi.nlm.nih.gov/books/NBK22430/

Bhalla, Aditya, Kenneth M. Bischoff, and Rajesh K. Sani. 2014.Highly Thermostable GH39 β-xylosidase from a Geobacillus Sp. Strain WSUCF1.BMC Biotechnology 14, no. 1. doi:10.1186/s12896-014-0106-8.

Cintra, Lorena Cardoso, Amanda Gregorim Fernandes, Izadora Cristina Moreira De Oliveira, Saulo J. Linhares Siqueira, Ilítia Ganaê Oliveira Costa, Francieli Colussi, Rosália Santos Amorim Jesuíno, Cirano José Ulhoa, and Fabrícia Paula De Faria. 2017.Characterization of a Recombinant Xylose Tolerant β-xylosidase from Humicola Grisea Var. Thermoidea and Its Use in Sugarcane Bagasse Hydrolysis. International Journal of Biological Macromolecules 105: 262-71. doi: 10.1016/j.ijbiomac.2017.07.039

Colquhoun, D.1969.Appl. Stat. 18, 130-140

Corrêa, Juliana Moço, Luciana Graciano, Josielle Abrahão, Eduardo Alexandre Loth, Rinaldo Ferreira Gandra, Marina Kimiko Kadowaki, Caroline Henn, and Rita De Cássia Garcia Simão.2012. Expression and Characterization of a GH39 β-Xylosidase II from Caulobacter Crescentus. Applied Biochemistry and Biotechnology 168, no. 8: 2218-229. doi:10.1007/s12010-012-9931-1.

Czjzek M, David AB, Bravman T, Shoham G, Henrissat B, Shoham Y. 2005. Enzyme–substrate complex structures of a GH39 β-xylosidase from Geobacillus stearothermophilus. J Mol Biol 2005; 353:838–4

Dodd, Dylan, and Isaac K. O. Cann.2009.Enzymatic Deconstruction of Xylan for Biofuel Production." *GCB Bioenergy*1, no. 1: 2-17. doi:10.1111/j.1757-1707.2009. 01004.x

Eisenthal, Robert, Michael J. Danson, and David W. Hough.2007.Catalytic Efficiency and Kcat/KM: A Useful Comparator? Trends in Biotechnology25, no. 6: 247-49. doi: 10.1016/j.tibtech.2007.03.010

Enzymes. (n.d.).Retrievedfromhttps://www.rsc.org/Education/Teachers/Resources/cfb/enzymes. Htm

Gene Synthesis handbook 3rd Edition, GenScript

Hydrogenomics of the Extremely Thermophilic Bacterium Caldicellulosiruptor Saccharolyticus. Applied and Environmental Microbiology 74, no. 21: 6720-729. doi:10.1128/aem.00968-08

Henrissat, Bernard, and Gideon J. Davies.2000. Glycoside Hydrolases and Glycosyltransferases. Families, Modules, and Implications for Genomics. Plant Physiology 124, no. 4: 1515-519. doi:10.1104/pp.124.4.1515.

Johnson, K. A and Goody, R. S.2011. The original Michaelis constant: translation of the 1913 Michaelis-Menten paper.Biochemistry, vol. 50, no. 39, pp. 8264–8269.

Kirikyali, Narin, Jonathan Wood, and Ian F. Connerton.2014. Characterisation of a Recombinant β-xylosidase (xylA) from Aspergillus Oryzae Expressed in Pichia Pastoris.AMB Express 4, no. 1. doi:10.1186/s13568-014-0068-1.

Lagaert, Stijn, Annick Pollet, Christophe M. Courtin, and Guido Volckaert.2014.β-Xylosidases and α-l-arabinofuranosidases: Accessory Enzymes for Arabinoxylan Degradation." *Biotechnology Advances*32, no. 2: 316-32. doi:10.1016/j.biotechadv.2013.11.005.

Li, Qi, Tao Wu, Zhipeng Qi, Linguo Zhao, Jianjun Pei, and Feng Tang.2018. Characterization of a Novel Thermostable and Xylose-tolerant GH 39 β-xylosidase from Dictyoglomus Thermophilum." *BMC Biotechnology*18, no. 1. doi:10.1186/s12896-018-0440-3.

Lim, C., Sen, P. K., and Peddada S. D.2013. Robust nonlinear regression in applications. Journal of the Indian Society of Agricultural Statistics, vol. 67, no. 2, pp. 215–234, 291

Luthi E., Jasmat N.B., Bergquist PL.1990a. Xylanase from the extremely thermophilic bacterium Caldocellum saccharolyticum: Overexpression of the gene in Escherichia coli and characterization of the gene product. Appl Environ Microbiol 56(9): 2677–2683.

Luthi E., Love D.R., McAnulty J., Wallace C., Caughey P.A., Saul D., Bergquist P.L. 1990b. Cloning, sequence analysis, and expression of genes encoding xylan-degrading enzymes from the thermophile Caldocellum saccharolyticum. Appl Environ Microbiol 56(4): 1017–1024

Marasović, Maja, Tea Marasović, and Mladen Miloš.2017. Robust Nonlinear Regression in Enzyme Kinetic Parameters Estimation. Journal of Chemistry 2017: 1-12. doi:10.1155/2017/6560983.

Martin Armbrecht, Eppendorf AG, Hamburg, Germany.2011.Generation of a standard curve for a colorimetric assay in the Eppendorf BioSpectrometer® basic and Eppendorf BioSpectrometer® kinetic. user guide

Matyska, L., and J. Kovář.1985. Comparison of Several Non-linear-regression Methods for Fitting the Michaelis-Menten Equation.Biochemical Journal 231, no. 1: 171-77. doi:10.1042/bj2310171.

Nelson, David L., and Michael M. Cox.2000. Lehninger Principles of Biochemistry. Third ed. New York, NY: Worth Publishers.

Ozdemir, Inci, Sara E. Blumer-Schuette, and Robert M. Kelly.2011. S-Layer Homology Domain Proteins Csac_0678 and Csac_2722 Are Implicated in Plant Polysaccharide Deconstruction by the Extremely Thermophilic Bacterium Caldicellulosiruptor Saccharolyticus. Applied and Environmental Microbiology 78, no. 3: 768-77. doi:10.1128/aem.07031-11.

Park, Joshua I., Michael S. Kent, Supratim Datta, Bradley M. Holmes, Zhaohua Huang, Blake A. Simmons, Kenneth L. Sale, and Rajat Sapra.2011. Enzymatic Hydrolysis of Cellulose by the Cellobiohydrolase Domain of CelB from the Hyperthermophilic Bacterium Caldicellulosiruptor Saccharolyticus. Bioresource Technology 102, no. 10: 5988-994. doi: 10.1016/j.biortech.2011.02.036.

Quintero, Denny, Zoraida Velasco, Estefanía Hurtado-Gómez, José L. Neira, and Lellys M. Contreras.2007. Isolation and Characterization of a Thermostable β-xylosidase in the

Thermophilic Bacterium Geobacillus Pallidus. Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics1774, no. 4: 510-18. doi: 10.1016/j.bbapap.2007.02.002.

Robinson, Peter K.2015. Enzymes: Principles and Biotechnological Applications. Portland Press Limited, Essays Biochem. 59, 1–41: doi: 10.1042/BSE0590001

Saul, D. J., Williams, L. C., Grayling, R. A., Chamley, L. W., Love, D. R., & Bergquist, P. L. 1990. celB, a gene coding for a bifunctional cellulase from the extreme thermophile "Caldocellum saccharolyticum". Applied and environmental microbiology, 56(10), 3117–312

Shin, Kyung-Chul, Min-Ju Seo, and Deok-Kun Oh.2014. Characterization of β-xylosidase from Thermoanaerobacterium Thermosaccharolyticum and Its Application to the Production of Ginsenosides Rg1 and Rh1 from Notoginsenosides R1 and R2. Biotechnology Letters 36, no. 11: 2275-281. doi:10.1007/s10529-014-1604-4.

Smaali, Issam, Caroline Rémond, and Michael J. O'Donohue.2006. Expression in Escherichia Coli and Characterization of β-xylosidases GH39 and GH-43 from Bacillus Halodurans C-125.Applied Microbiology and Biotechnology73, no. 3: 582-90. doi:10.1007/s00253-006-0512-5.

Van De Werken H. J. G., M. R. A. Verhaart, A. L. Vanfossen, K. Willquist, D. L. Lewis, J. D. Nichols, H. P. Goorissen, E. F. Mongodin, K. E. Nelson, E. W. J. Van Niel, A. J. M. Stams, D. E. Ward, W. M. De Vos, J. Van Der Oost, R. M. Kelly, and S. W. M. Kengen. 2008.Hydrogenomic of the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus. Appl Environ Microbiol; doi:10.1128/AEM.00968-08.

Vanfossen, Amy L., Inci Ozdemir, Samantha L. Zelin, and Robert M. Kelly.2011. Glycoside Hydrolase Inventory Drives Plant Polysaccharide Deconstruction by the Extremely Thermophilic Bacterium Caldicellulosiruptor Saccharolyticus. Biotechnology and Bioengineering 108, no. 7: 1559-569. doi:10.1002/bit.23093

Wagschal, Kurt, Diana Franqui-Espiet, Charles C. Lee, George H. Robertson, and Dominic W. S. Wong. 2007. Cloning, Expression and Characterization of a Glycoside Hydrolase Family 39 Xylosidase from Bacillus Halodurans C-125. Applied Biochemistry and Biotechnology 146, no. 1-3: 69-78. doi:10.1007/s12010-007-8055-5.

Wilkinson, G. N.1961. Biochem. J. 80, 324-332

Xu, Hui., Dongmei Han, and Zhaohui Xu.2015. Expression of heterologous cellulases in thermotoga sp. strain rq2. Biomed research international 2015: 1-11. doi:10.1155/2015/304523.

Yang J K, Yoon H J, Ahn H J, Il Lee B, Pedelacq JD, Liong EC, et al. 2004.Crystal structure of β-D-xylosidase from *Thermoanaerobacterium saccharolyticum*, a family 39 glycoside hydrolase. J Mol Biol; 335:155–65

CHAPTER 3

ALTERNATIVE METHODS AND STRATEGIES FOR CLONING AND EXPRESSION OF PUTATIVE XYLANASE GENES FROM *Caldicellulosiruptor saccharolyticus*

Introduction

The major components of lignocellulosic biomass include cellulose and hemicelluloses, cross-linked with lignin and spread throughout plant cell walls. Next to cellulose, hemicellulose is the most abundant material in plants which accounts for 40% of lignocellulose material. In secondary cell walls, xylan is the main hemicellulose component (Ji et al., 2011). Xylan is a branched heteropolysaccharide consisting of a backbone of β-(1,4)-linked xylopyranosyl units which are typically substituted with acetyl (Ac), glucuronosyl, or arabinofuranosyl side chains. In addition to xylose, xylan consists of arabinose, glucuronic acid, acetic, ferulic, and *p*-coumaric acids. The composition of branches and their frequency in xylan depend on the plant (Saha, 2003). Different methods for pretreatment of lignocellulosic biomass include acid and alkali hydrolysis, steam explosion, and ammonia explosion. Among these methods, acid hydrolysis can damage hemicellulose fraction while other methods keep it intact. Therefore, enzymatic hydrolysis can be considered as an alternative to release sugar monomers from hemicellulose (Su et al., 2012).

The total biodegradation of xylan is carried out by several synergetic enzymes. The xylan backbone is randomly cleaved by endo-1,4-β-D-xylanases, while β-D-xylosidases cleave xylose monomers from the non-reducing end of xylo-oligosaccharides and xylobiose. The accessory enzymes including α-L-arabinofuranosidases, α-D-glucuronidases, acetylxylan esterases, ferulic acid esterases and *p*-coumaric acid esterases catalyze the removal of side chains

(Collins, Gerday, and Feller, 2005). Endoxylanase and xylosidase are the crucial enzymes due to their direct involvement in xylan degradation (Andrade and Sant, 2013). Many microbes produce xylan degrading enzymes, however, thermostable enzymes are more desirable in biofuel production due to their ability to withstand harsh process conditions. In this case, thermophilic species of the *Caldicellulosiruptor* genus have been considered as great candidates due to their ability to convert lignocellulosic biomass into fermentable sugar molecules by producing effective carbohydrate degrading enzymes. Thermophilic species of the *Caldicellulosiruptor* genus tend to produce a wide variety of carbohydrate-degrading enzymes.

The genome of *Caldicellulosiruptor saccharolyticus* has been sequenced and it was the first member of *Caldicellulosiruptor* genus that was transcriptionally analyzed to investigate how entire genome responds to various carbon sources (Blumer-Schuette et al., 2013). In addition to producing a wide range of glycoside hydrolase (GH) enzymes, *C. saccharolyticus* is capable of generating hydrogen grown on monosaccharides (van Niel et al., 2002; Kádár et al., 2003; de Vrije et al., 2007; Chowdry et al., 2015) and wastepaper. In fact, species of *Caldicellulosiruptor* are able to co-ferment C_5 (xylose) and C_6 sugars (glucose) from biomass (Cha et al., 2013).
To identify potential Carbohydrate Active enZYmes (CAZYs) and their functionality, genomes of hyperthermophiles and other species can be investigated. The genome and transcriptional studies indicate that *C. saccharolyticus* is capable of encoding wide variety of GHs including Xylanases, (VanFossen et al., 2011) Glucanases, Mannanases (Morris et al., 1995), in addition to side chain degrading enzymes (Lim et al., 2010) belonging to various GH families. These enzymes tend to form complex structures by combining with each other, which makes it difficult to isolate native enzymes homogeneously. In this case, recombinant DNA technology has been a useful technique allowing isolation of pure enzymes expressed in *E. coli* bacterial hosts (Zhengqiang et al., 2001).

The focus of this study was to investigate glycoside hydrolases from the extremely thermophilic bacterium *C. saccharolyticus.* In this work, I made an effort to clone genes, Csac_2405 and Csac_2408 of GH10 in addition to Csac_2409 of GH39 using several cloning methods and express the protein recombinantly. The biochemical characterization of putative β-xylosidase gene Csac_2409 was described in Chapter 2. This chapter discussed efforts to clone and express each of Csac_2405, Csac_2408, and Csac_2409 which collectively could be used for improved pretreatment methods for biofuel production.

Methods

DNA Amplification

The extracted genomic DNA from *C. saccharolyticus* was utilized to amplify the gene of interest encoded Csac_2409 (XynD) by polymerase chain reaction (PCR) under the conditions shown in Table 3.1. In order to identify optimal annealing temperature in PCR cycle, temperature was varied between 40-50°C. A PCR matrix (Appendix B) was used to identify the optimal concentrations of different components in PCR templates. The primers were designed on the basis of Csac_2409 gene sequence using online primer designing tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Following reverse and forward primers (Table 3.2) containing restriction sites (NheI and XhoI) were used.

Table 3.1: PCR cycle

Cycle Number	Steps	Temperature (°C)	Time (Seconds)
1	Initial denaturation	95	120
2	Denaturation	95	45
3	Annealing	Varied 40-50	45
4	Extension	72	135
Repeat steps 2 through 4 for 34 cycles			
5	Final extension	72	300
6	Soak	10	-

Table 3.2 : Forward and reverse primers incoprotated restriction sites for Csac_2409 (underlined)

Forwrd Primer	5' - TGT GTG GCT AGC ATA AAG ATT GAG AAA GGC - 3' NheI
Reverse Primer	5' – CCC CCA CTC GAG ATA ACC TGG TAT TTT ACT ATC - 3' XhoI

Cloning Strategies

PCR products were purified using QIAquick PCR purification kit (QIAGEN) and loaded on agarose gel with loading dye for electrophoresis in TBE buffer (Appendix C). The separated DNA fragments were visualized by UV transilluminator to determine gene length and purity. The purified DNA fragments and pET21b cloning vectors were double digested with restriction enzymes, *NheI* and *XhoI*, followed by incubation at 37ºC for 2 hours. The insert DNA was ligated into digested vector backbone using T4 DNA ligase enzyme (Promega) and incubated at room temperature for 2 hours. The ligation products were transformed into competent cells (NovaBlue) which were grown in LB medium with appropriate antibiotic (Ampicillin, 100 µg/mL) at 37ºC overnight. The standard mini-prep procedure (QIAprep Spin Miniprep Kit, QIAGEN) was used to isolate plasmid DNA and imaged by gel electrophoresis.

Alternative Cloning Methods and Strategies

Dephosphorylation – Thermostable Alkaline Phosphatase (TSAP)

The strategy of dephosphorylation is used in traditional cloning process to prevent re-circularizing of vector during ligation. Thermosensitive Alkaline Phosphatase (TSAP, Promega) is an enzyme which is effective on 3´ overhangs, 5´ overhangs, and blunt ends. In addition, it can be used to prepare DNA for 5' end-labeling. Here, TSAP was used to remove 5' phosphate groups from DNA once the vector has been cut with restriction enzymes (Figure 3.1). This leads to prevent the re-circularization and re-ligation of linearized cloning vector DNA during ligation. During transformation, the background can be reduced by decreasing the frequency

re-circularization without an DNA gene insert. This method does not require cleanup process before ligation because TSAP is inactivated by heating at 74℃.

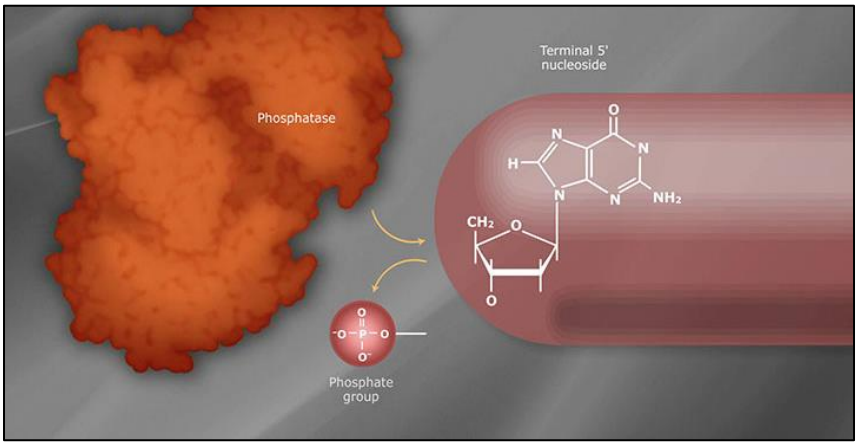


Figure 3.1: Process of dephosphorylation (New England Biolabs)

Protocol

Vector DNA was double digested with required restriction enzymes. 1 µl of TSAP was added to double digested DNA vector reactions and incubated at 37ºC for 15 minutes to dephosphorylate the vector DNA overhangs in MULTI-CORE 10X buffer. TSAP in reaction mixtures was heat inactivated by incubating at 74ºC for 15 minutes. The vector DNA and insert were ligated using T4 DNA ligase and incubated at room temperature for 2 hours. The ligated products were transformed into *E. coli* competent cells and plated on agar containing appropriate antibiotic.

NEB PCR cloning kit

The PCR cloning kit from NEB consists of optimized cloning mixes and a linearized vector. These cloning mixes contain an exclusive ligation enhancer, and the provided vector has the ability to suppress background colonies of re-ligated vector. This method allows simple and quick cloning of PCR amplicon which is independent from proofreading DNA polymerase in amplification reactions. This kit permits direct cloning from amplification reactions without PCR

purification. In fact, this approach performs well even without using the primers in the PCR containing 5'-phosphate groups.

It is difficult for *E. coli* to process genes encoding few numbers of amino acids which are also known as minigenes. These minigenes allow the premature production of small peptides from ribosome. The last amino acid residues of such genes are still bonded to their tRNA therefore, tRNA will not be able to participate in protein synthesis. The protein synthesis will be inhibited if the vector reconnects without the gene of interest and results in no colonies. This method allows to clone into the interrupted minigene and only the vector containing the insert leads to cell growth (Figure 3.2).

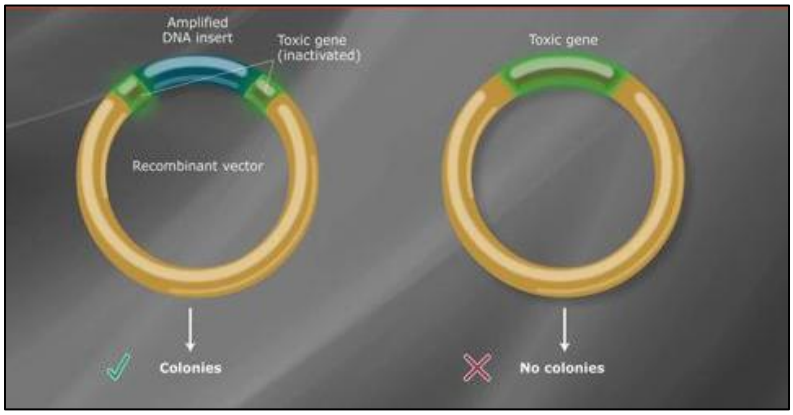


Figure 3.2: **Left**: Recircularization of vector with amplified DNA insert in the middle of interrupted toxic mini gene. **Right**: Recircularization of vector without an insert (NEB® PCR Cloning Kit, New England Biolabs)

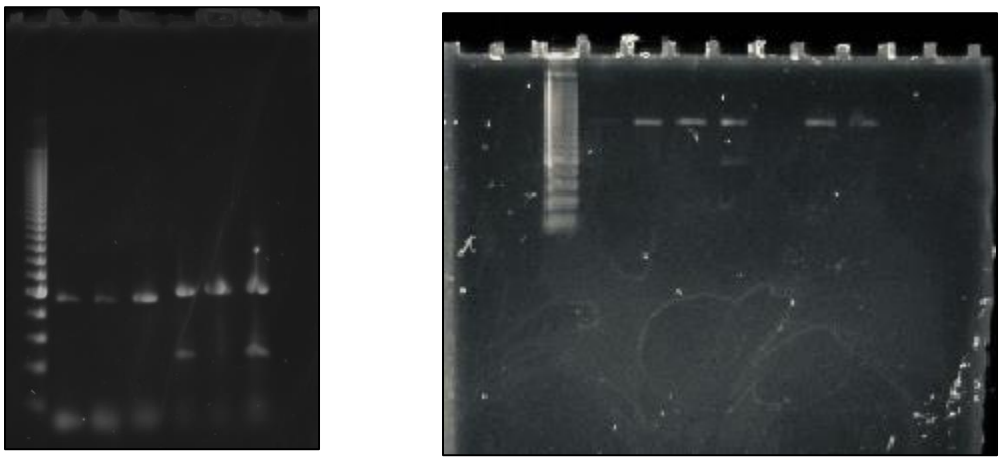
Protocol

The insert DNA was mixed with 1 µL linearized vector, and water was added to bring total volume of the reaction to 5 µL. Then 5 µL cloning mixes were added for a total of 10 µL per ligation reaction. The reaction mixtures were incubated for 15 minutes at room temperature and 2 minutes on ice. Then ligation products transformed to competent cells immediately using heat shock method. The cells were spread on agar plates containing 100 µg/ml ampicillin and

incubated overnight at 37°C. The mini-prep process was used to isolate DNA plasmid from overnight cultures from individual colonies and imaged using gel electrophoresis.

Results/Null Results

The optimal annealing temperature and PCR conditions were determined for genes of interest. The results indicated optimal annealing temperature of 43°C. In addition, from PCR matrix it was determined that high concentrations of forward and reverse primers result in increased PCR products (Figure 3.3A).



A **B**

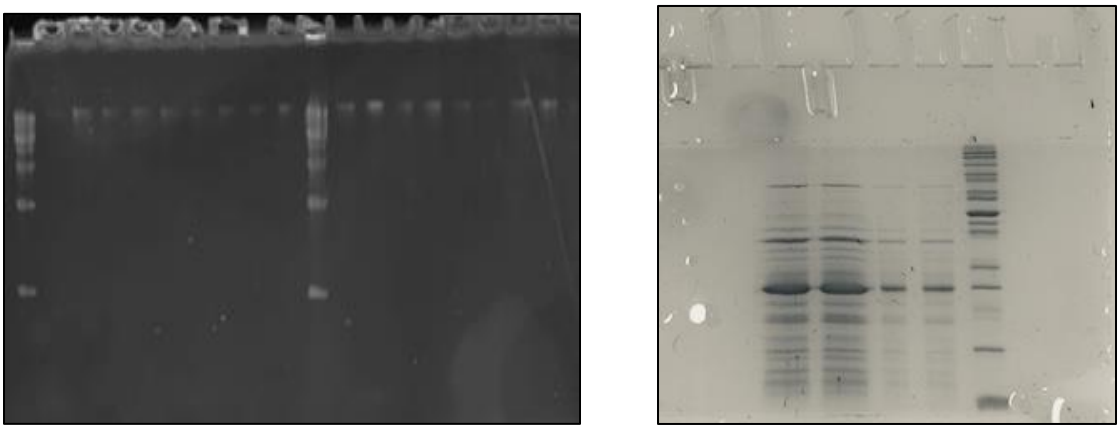
Figure 3.3: Ethidium-Bromide stained gel visualized by UV transilluminator **A)** PCR product of Csac_2405 and Csac_2408 at 43 °C annealing temperature Left to Right: 200bp ladder (lane 1), Csac_2405 (lanes 2-4), Csac_2408 (lanes 5-7) **B)** Plasmid DNA from PCR cloning kit purified with Mini-prep. Left to Right: 1kb ladder (lane 1), Csac_2408 (Lanes 2-7).

The standard mini-prep procedure used to isolate plasmid DNA from traditional cloning method and imaged using gel electrophoresis. Several trials were attempted to clone genes of interest including TSAP and results indicated no success of cloning methods. However, colonies were observed with PCR cloning kit and overnight cultures were made from individual colonies with appropriate antibiotic. Plasmid DNA was isolated via mini-prep process and imaged with gel electrophoresis (Figure 3.3B). The isolated plasmid DNA was sent for DNA sequencing (Sanger method) at CCHMC (Cincinnati, OH). Finch TV 1.4.0 chromatogram viewer was used to view

trace data from Sanger DNA Sequencing. However, results indicated that gene was not inserted properly.

Protein Expression Trials

It was attempted to express previously cloned genes of interest from bacterial cells stored in 40% glycerol solutions at -20°C and -70°C. The overnight cultures were made with appropriate antibiotic and mini-prep procedure was used to isolate plasmid DNA (Figure 3.4A). The results showed the absence of plasmid DNA from previously cloned genes.



A B

Figure 3.4: **A)** Ethidium-Bromide stained gel of plasmid DNA purified with Mini-prep **B)** SDS-PAGE analysis of purified protein from *C. saccharolyticus*. Lane 1, pre-stained EZ-Run Rec protein ladder; Lane 2 crude extract of Csac_2405 enzyme (2 µl protein with 2 µl loading buffer) after heat treatment; Lane 3 crude extract of Csac_2408 (2 µl protein with 2 µl loading buffer) enzyme after heat treatment; Lane 4 crude extract of Csac_2405 enzyme (10 µl protein with 10 µl loading buffer) after heat treatment; Lane 5 crude extract of Csac_2408 (10 µl protein with 10 µl loading buffer) enzyme after heat treatment.

Previously expressed genes of interest from Rosetta bacterial cell pellets (stored at -70°C) were used to prepare 5 ml overnight cultures at 37°C. Overnight cultures were used to inoculate 100 ml LB media with appropriate antibiotics (100 µg/ml ampicillin and 25 µg/ml chloramphenicol). The cultures of 100 ml were centrifuged, and the cell pellets were resuspended in 10 ml 20 mM Tris-HCl (pH 8.0). In order to lyse cells, lysis reagent (FastBreak, Promega) was used followed by centrifugation for 15 minutes. The lysed cells were heat treated at 70°C for 30

minutes to denature most of the *E. coli* proteins. The denatured proteins and cell debris were removed by centrifugation for 20 minutes and the resulting supernatant analyzed by 10% SDS-PAGE (Appendix D) to determine molecular weight (Figure 3.4B).

Learnings and Recommendations

This study aimed to extend the cloning of putative xylanase genes from the extreme thermophilic bacterium *Caldicellulosiruptor saccharolyticus.* The genes of interest have been recognized for their ability to hydrolyze xylan of hemicellulose fraction into monomeric sugars which can be fermented into ethanol. Microorganisms produce a variety of xylan degrading enzymes, which tend to form complex structures with each other. As a result of this, isolation of such enzymes homogeneously has become a difficult process. However, as a useful technique recombinant DNA technology has been considered to isolate pure enzymes expressed in mesophilic hosts (Zhengqiang et al., 2001).

Bacterial cells are often used for production of recombinant proteins. The properties of the protein and its origin are essential in order to select mesophilic host and their expression system. The Gram-negative *E. coli* is widely utilized as bacterial hosts to produce DNA recombinantly since they offer several advantages such as simple molecular process, low-cost media, and rapid protein production in a short period of fermentation. However, challenges like inclusion bodies formation and complex protein folding may be present when recombinant proteins are expressed intracellularly in *E. coli.* (Zhou et al., 2018). Recombinant expression of thermostable protein via thermophilic hosts has been studied. In this case, thermophilic hosts like *Bacillaceae* have been identified as a potential group of organisms which can synthesize, fold and result in high yields compared to *E. coli* (Eivind et al., 2018).

In this work traditional cloning method was attempted to clone genes of interest and express the protein recombinantly. For this process typically, polymerase chain reaction (PCR) is used to amplify the genes of interest and add required restriction sites. These restriction sites are

added to the ends of DNA sequences and also, they are used to cut both the template and the vector. Then the sticky ends of the template and vector are joined together using DNA ligase. The ligated products are finally transformed into competent *E. coli* cells. This traditional cloning method has been widely used since it is versatile which permits easy manipulation of DNA to create desired constructs. However, this process requires frequent optimization, and the reagents used can be expensive. In addition, restriction sites can cause sequence constraints (GenScript; NEB).

Due to unsuccessful cloning attempts with traditional cloning method, several alternative cloning methods and strategies were followed in order to make an effort to achieve successful cloning. For this, dephosphorylation using TSAP, PCR cloning kits in addition to cloning, and expression from previously cloned genes of interest were attempted. However, all the efforts to clone and express the genes of interest were not productive despite the various methods and strategies applied in the process.

It is essential to trace the workflow of cloning process when something impedes the progress even though the protocol was followed properly. In this case, checking each step of the procedure is required to detect issues that may have occurred. Traditional cloning method could be performed within a few weeks by following the protocol. However, troubleshooting of each step involved would extend the time to finish the cloning process. It has been suggested that use of controls in each step would make troubleshooting more convenient. The common problems with traditional cloning method include poor transformation efficiency, ineffective ligation, increased background, fewer colonies, colonies without plasmid, and extra bands in the gel and PCR. In order to address these issues, troubleshooting guides by ThermoFisher Scientific and New England Biolabs are recommended.

There are several possible factors that may have affected the cloning process. One potential source of error could have been the genomic DNA. While DNA is reported to be viable when stored at -20°C, the genomic DNA used in these experiments was extracted in 2009.

Unfortunately, per Dr. Comfort, the culture of *C. saccharolyticus* was not maintained and could not be revived for growth to generate more genomic DNA for use in these experiments. Likewise, the pET vectors used in the experiments were also aged and this may have contributed to the challenges associated with obtaining successful clones by numerous approaches.

The purity and concentration of genomic DNA are important for cloning process. The issues with starting material can be caused by several reasons: using shared materials, stored in freezer for an extended time period. This requires verification of the sequence before PCR in order to prevent problems which may occur end of the cloning process. In addition, repeated sequences in the gene of interest can be problematic which results in improper annealing in PCR and subsequent ligation. Therefore, high-fidelity polymerase and high-performance thermocyclers have been suggested to help with this issue. When designing primers, GC content is vital to determine temperatures for PCR step since there are three hydrogen bonds between guanine (G) and cytosine (C). High GC content of sequences makes it difficult to amplify the gene and requires high melting temperatures or addition of DMSO/betaine. In addition, amplification of longer gene sequences in full length using PCR and replication of plasmid DNA in bacterial cells could be difficult in cloning process (GenScript).

Despite this lack of success, there are many upcoming technologies that facilitate cloning. Herein, I applied many in the attempts to successfully clone the gene into an expression vector, resulting in a greater understanding of the process and modern molecular biology technologies.

References

Biolabs, New England. "Dephosphorylation." New England Biolabs: Reagents for the Life Sciences Industry. https://www.neb.com/applications/cloning-and-synthetic-biology/dna-end-modification/dephosphorylation.

Blumer-Schuette, S. E., Brown, S. D., Sander, K. B., Bayer, E. A., Kataeva, I., Zurawski, J. V., Kelly, R. M. 2013.Thermophilic lignocellulose deconstruction. FEMS Microbiol Rev, 38: 393–448. doi: 10.1111/1574-6976.12044

Cha, Minseok, Daehwan Chung, James G. Elkins, Adam M. Guss, and Janet Westpheling. 2013.Metabolic Engineering of *Caldicellulosiruptor Bescii* Yields Increased Hydrogen Production from Lignocellulosic Biomass. Biotechnology for Biofuels 6, no. 1: 85. doi:10.1186/1754-6834-6-85.

Chowdhary, Nupoor, Ashok Selvaraj, Lakshmi Krishnakumaar, and Gopal Ramesh Kumar. 2015.Genome Wide Re-Annotation of Caldicellulosiruptor Saccharolyticus with New Insights into Genes Involved in Biomass Degradation and Hydrogen Production. Plos One 10, no. 7 doi: 10.1371/journal.pone.0133183.

Collins, T., Gerday, C., & Feller, G. 2005. Xylanases, xylanase families and extremophilic xylanases. FEMS Microbiol Rev, 29(1): 3-23. doi: 10.1016/j.femsre.2004.06.005
De Vrije T, Mars AE, Budde MA, Lai MH, Dijkema C, de Waard P, et al. 2007.Glycolytic pathway and hydro-gen yield studies of the extreme thermophileCaldicellulosiruptor saccharolyticus. Appl Microbiol Bio-technol.; 74: 1358–1367.

Drejer, Eivind, Sigrid Hakvåg, Marta Irla, and Trygve Brautaset.2018. Genetic Tools and Techniques for Recombinant Expression in Thermophilic Bacillaceae. Microorganisms 6, no. 2: 42. doi:10.3390/microorganisms6020042

Ji, Xiao-Jun., Huang, H., Nie, Zhi-Kui, Qu, L., Xu, Qing & T Tsao, George. 2012. Fuels and Chemicals from Hemicellulose Sugars. Advances in biochemical engineering/biotechnology. 128. 199-224. 10.1007/10_2011_124.

Kádár, Z., Vrije, T. D., Budde, M. A., Szengyel, Z., Réczey, K., & Claassen, P. A. 2003. Hydrogen Production from Paper Sludge Hydrolysate. Biotechnol. Fuels Chemicals, 557-566. doi:10.1007/978-1-4612-0057-4_46

Andrade L., F., C. C. P., and M. H. A. Sant.2013. A Review of Xylanase Production by the Fermentation of Xylan: Classification, Characterization and Applications. Sustainable Degradation of Lignocellulosic Biomass - Techniques, Applications and Commercialization.doi:10.5772/53544

Lim, Y., Yoon, R., Seo, E., Kim, Y., Park, C., & Oh, D.2010. Hydrolytic properties of a thermostable α-l-arabinofuranosidase from Caldicellulosiruptor saccharolyticus. *Journal of Applied Microbiology, 109*(4), 1188-1197. doi:10.1111/j.1365-2672.2010. 04744.x

Morris, D. D., R.A. Reeves, M.D. Gibbs, D.J. Saul, and P. Bergquist.1995. Correction of the the β-mannanase domain of the celC pseudogene from Caldocellulosiruptor saccharolyticus and activity of the gene product on kraft pulp. Applied and Environmental Microbiology 61: 2262–2269.

Neb PCR Cloning Kit Instructional manual.2017. New England Biolabs

Saha, Badal C.2003.Hemicellulose Bioconversion.Journal of Industrial Microbiology and Biotechnology 30, no. 5: 279-91. doi:10.1007/s10295-003-0049-x.

Su, Xiaoyun, Yejun Han, Dylan Dodd, Young Hwan Moon, Shosuke Yoshida, Roderick I. Mackie, and Isaac K. O. Cann. 2012. Reconstitution of a Thermostable Xylan-Degrading Enzyme Mixture from the Bacterium Caldicellulosiruptor Bescii. Applied and Environmental Microbiology 79, no. 5: 1481-490. doi:10.1128/aem.03265-12

Van Niel E.W.J., Budde M.A.W., Haas G.G. de., van der Wal F.J., Claassen P.A.M., Stams A.J.M., 2002. Distinctive properties of high hydrogen producing extreme thermophiles, Caldicellulosiruptor saccharolyticus and Thermotoga elfii. Int J Hydrogen Energy 27: 1391–1398.

Vanfossen, A. L., Ozdemir, I., Zelin, S. L., & Kelly, R. M.2011. Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus. Biotechnol Bioengin, 108(7): 1559-1569. doi:10.1002/bit.23093

Zhengqiang, J., Kobayashi, A., Ahsan, M. M., Lite, L., Kitaoka, M., & Hayashi, K.2001. Characterization of a Thermostable Family 10 Endo-Xylanase (XynB) from Thermotoga maritima That Cleaves p-Nitrophenyl- BETA-D-Xyloside. J Biosci Bioengin, 92(5): 423-428. doi:10.1263/jbb.92.423

Zhou, Y.; Lu, Z.; Wang, X.; Selvaraj, J.N.; Zhang, G.2018. Genetic engineering modification and fermentation optimization for extracellular production of recombinant proteins using Escherichia coli. Appl. Microbiol. Biotechnol.102, 1545–1556

CHAPTER 4

CONCLUSION

Complete degradation of xylan to xylose sugars requires several enzymes, mainly endoxylanases and β-xylosidases. β-xylosidases are essential to hydrolyze xylo-oligosaccharides to monomeric xyloses. Thermostable β-xylosidase has been considered for biofuel production due to their high thermostability which is capable of withstanding harsh bioprocessing conditions. Cloning and recombinant expression of 1,4-β-xylosidase (XynD) derived from extreme thermophile *C. saccharolyticus* was attempted. Due to unsuccessful cloning attempts, the gene encoding Csac_2409 was chemically synthesized and cloned by GenScript, which was then biochemically and biophysically characterized. To investigate enzyme activity, synthetic substrate *p*NPX was utilized as an alternative for natural substrate, xylobiose. β-xylosidase showed activity over wide range of temperature and pH with optimal temperature of 80°C and pH 6.5. Michaelis-Menten kinetic parameters, K_M, V_{max}, k_{cat},and k_{cat}/K_m were determined to be 0.918 mM, 0.0042 mM/s, 13.6 s^{-1} and 14.75 $s^{-1}mM^{-1}$ respectively.

The characterization of the β-xylosidase from *C. saccharolyticus* shows that it is an improvement over other β-xylosidases in GH family 39 (Table 2.1) that have been characterized to date. These benefits are not just catalytic with a greater catalytic efficiency, but also in the higher thermostability, which can be applied to reduce potential bacterial contamination of the hydrolyzed sugar substrates. This provides an option for the finishing step of the xylose degradation, but much work remains in initial degradation of the cellulose, hemicellulose, and lignin. Hemicellulolytic enzymes and cellulases play a significant role in hydrolysis of lignocellulosic biomass. Thus, the interaction between cellulases and hemicellulases is in fact essential to maximize the sugar yield.

To achieve complete hydrolysis of xylan, several xylan degrading enzymes are required including endo-1,4- β-xylanases and β-xylosidases. The synergism of these enzymes appears to

be an effective option for improved conversion. In fact, the evaluation of synergistic cooperation of genes encoding endo-1,4-β-xylanases (Csac_2405, Csac_2408, and Csac_2410) and β-xylosidases derived from *C. saccharolyticus* for efficient hydrolysis is essential to reduce cost associated with second-generation biofuels.

APPENDIX A

Cloning Strategy

Full length protein:

Xylosidase

NdeI--ATG**--Xylosidase**--His tag--Stop codon--HindIII

Protein Length=512 MW=59762.7 Predicted pI=6.43 vector: pET30a

M**GEGIVTNIKIEKGKQIGIFPDKWKFCVGSGRIGLALQKEYIDALLYVKKHIDFKYIRAHGLLHDDVGIY
REDIIDGNEVSFYNFTYIDRIYDSFLELGIRPFVEIGFMPSKLASGTQTVFYWRGNVTPPKDYGKWERLIK
SVVKHFIDRYGEKEVVQWPFEIWNEPNLNVFWKDANQAEYFKLYEVTAKAIKEVNENIKVGGPAICGG
SDYWIDDFLHFCYKNKVPVDFLTRHAYTGKPPQYTPHFVYQDVHPIEYMLNEFKSVREKVRNSPFPDLP
IHITEFNSSYHPLCPIHDTPFNAAYLARVLSEAGDYVDSFSYWTFSDVFEEADVPRAIFHGGFGLVAFNNI
PKPVFHMFTFFNAMGNKILYRDEHILITERENGTIAIVAWNEVMKKEEATDRKYKIEIPVDYNDVFIKQK
LIDEENGNPWRTWIQMGRPRYPTKEQIETLREVAKPKISTFRMMAENGYITLEFTLSKNAVVLFEVSKV
VDESDTYIGLDDSKIPGY**HHHHHH**..

DNA sequence: 1551bp

**CATATGGGAGAAGGGATAGTAACAAATATAAAGATAGAAAAGGGAAAGCAAATCGGTATCTTCCCG
GACAAATGGAAATTCTGCGTGGGCAGCGGTCGTATTGGCCTGGCGCTGCAGAAGGAGTACATCGA
CGCGCTGCTGTATGTGAAGAAACACATCGATTTCAAATACATTCGTGCGCACGGTCTGCTGCACGA
CGATGTGGGCATCTATCGTGAGGACATCATTGATGGCAACGAAGTTAGCTTCTACAACTTTACCTA
TATCGACCGTATTTACGATAGCTTCCTGGAGCTGGGTATCCGTCCGTTCGTTGAAATTGGCTTTAT
GCCGAGCAAGCTGGCGAGCGGTACCCAAACCGTGTTTTACTGGCGTGGCAACGTTACCCCGCCGA
AGGACTATGGTAAATGGGAACGTCTGATTAAGAGCGTGGTTAAACACTTCATCGATCGTTACGGCG
AGAAAGAAGTGGTTCAGTGGCCGTTTGAGATTTGGAACGAACCGAACCTGAACGTGTTCTGGAAG
GACGCGAACCAAGCGGAGTACTTTAAACTGTATGAAGTTACCGCGAAGGCGATCAAAGAGGTGAA
CGAAAACATTAAGGTTGGTGGCCCGGCGATCTGCGGTGGCAGCGACTACTGGATCGACGATTTCC
TGCACTTTTGCTATAAGAACAAAGTGCCGGTTGATTTCCTGACCCGTCATGCGTACACCGGTAAAC
CGCCGCAGTACACCCCGCACTTCGTGTATCAAGATGTTCACCCGATTGAGTATATGCTGAACGAAT
TTAAGAGCGTGCGTGAGAAAGTTCGTAACAGCCCGTTCCCGGACCTGCCGATCCACATTACCGAAT
TTAACAGCAGCTACCACCCGCTGTGCCCGATCCACGATACCCCGTTCAACGCGGCGTATCTGGCGC
GTGTGCTGAGCGAGGCGGGTGACTACGTTGATAGCTTTAGCTATTGGACCTTCAGCGACGTGTTTG
AGGAAGCGGATGTTCCGCGTGCGATTTTCCACGGTGGCTTTGGTCTGGTGGCGTTCAACAACATCC
CGAAGCCGGTTTTCCACATGTTTACCTTCTTTAACGCGATGGGCAACAAAATCCTGTACCGTGACG
AACACATCCTGATTACCGAGCGTGAAAACGGTACCATCGCGATTGTGGCGTGGAACGAGGTTATGA
AGAAAGAGGAAGCGACCGACCGTAAGTACAAAATCGAAATTCCGGTGGACTATAACGATGTTTTCA
TCAAGCAGAAACTGATTGATGAGGAAAACGGCAACCCGTGGCGTACCTGGATTCAGATGGGTCGT
CCGCGTTACCCGACCAAGGAGCAAATTGAAACCCTGCGTGAGGTGGCGAAGCCGAAAATCAGCAC
CTTTCGTATGATGGCGGAGAACGGTTATATCACCCTGGAATTCACCCTGAGCAAGAACGCGGTGGT
TCTGTTTGAGGTTAGCAAAGTGGTTGACGAAAGCGATACCTACATCGGTCTGGACGACAGCAAGAT
TCCGGGCTACCATCACCACCACCATCATTAATGAAAGCTT**

APPENDIX B

PCR Matrixing Template

The optimum conditions for a PCR are dependent on many variables. Matrices like this are used to systematically find the optimum component concentrations.

A. Vary Template	T25 (25 ng templ.)	T100 (100 ng templ)	T200 (200 ng templ)
Template	1*	2 (2 micro)	4
200 μM dNTPs	5	5	5
0.3 μM Fwd Primer	1.5	1.5	1.5
0.3 μM Rev Primer	1.5	1.5	1.5
1.25 U DNA Pol.	1	1	1
10X Buffer	5	5	5
dH_20	35	34	32

*dilute stock

B. Vary dNTP	dNTP 100 (100μM dNTP)	dNTP 200 (200μM dNTP**)**	dNTP 250 (250 μM dNTP)
Template	1	1	1
200 μM dNTPs	2.5	5	5.5
0.3 μM Fwd Primer	1.5	1.5	1.5
0.3 μM Rev Primer	1.5	1.5	1.5
1.25 U DNA Pol.	1	1	1
10X Buffer	5	5	5
dH_20	37.5	35	34.5

C. Vary Primer	P0.1 (0.1 μM Primer)	P0.3 (0.3 μM Primer)	P0.5 (0.5 μM Primer)
Template	1	1	1
200 μM dNTPs	5 (use 1 micro)	5 (1 micro)	5 (1 micro)
0.3 μM Fwd Primer	1*	1.5	2.5
0.3 μM Rev Primer	1*	1.5	2.5
1.25 U DNA Pol.	1	1**	1
10X Buffer	5	5	5
dH_20	36	35	33

* dilute stock (10 μM). (1 μL stock (10 μM) + 1 μL H_2O) → add 1 μL of this

** 0.5 U Pfu Turbo (1μL, 1.25 U/mL + 1.5 μL H_2O) → add 1 μL of this

Mixture Order: H_2O → 10X Buffer → dNTPs → DNA template → primers → Polymerase

APPENDIX C

5X TBE Recipe

5X TBE recipe	
Tris base	54g
Boric acid	27.5g
EDTA	3.72g (20 mL of 0.5M)
DI water	To 1000mL

APPENDIX D

SDS-PAGE Gel Protocol

High Destain
40% Ethanol
10% Glacial Acetic Acid
50% Water

Low Destain
10% Ethanol
10% Glacial Acetic Acid
80% Water

Tris-Glycine Running Buffer (10X) *

Tris Base	29.0 g
Glycine	144.0 g
SDS	10.0 g
DI Water to	1.0 L

(Best if made fresh)

4x Loading Buffer

0.5 M Tris-HCl, pH 6.8	1.0 mL
Glycerol	0.8 mL
10% (w/v) SDS	1.6 mL
2-mercaptoethanol	0.4 mL
0.1%(w/v) bromophenol blue	0.4 mL
DI	3.8 mL

(store at room temp)

Loading Buffer* (2X)

0.5 M Tris-Cl, pH 6.8	2.5 ml
Glycerol	2.0 ml
10% (w/v) SDS	4.0 ml
0.1% Bromophenol blue	0.5 ml
β-mercaptoethanol**	0.5 ml
DI water to	0.5 ml

Loading Buffer (1X)

50 mM Tris-Cl
10% glycerol
2.0% SDS
0.1% Bromophenol blue
100 mM Dithiothreitol**
(Add DTT just before use)

NATIVE PAGE GEL BUFFERS

Running Buffer (10X)

Tris Base	29.0 g
Glycine	144.0 g
DI Water to	1.0 L

Loading Buffer (2X)

1.5 M Tris-Cl, pH 8.8	4.0 ml
Glycerol	2.0 ml
0.1% Bromophenol blue	0.5 ml
DI Water	to 10 m

SDS-PAGE GELS (from Sambrook, Fritsch, and Maniatis)

Resolving Gel (10 mL of 12%)		**Stacking Gel (5 mL of 5%)**	
DI Water	4.9 mL	DI Water	5.5 mL
30% Acrylamide-bis	6.0 mL	30% Acrylamide-bis	1.3 mL
1.5 M Tris (pH 8.8)	3.8 mL	1.0 M Tris (pH 6.8)	1.0 mL
#10% SDS	150 µL	#10% SDS	80 µL
#10% Ammonium persulfate	150 µL (fresh)	#10% Ammonium persulfate	80 µL (fresh)
#TEMED	6 µL	#TEMED	8 µL
		#1% Bromophenol Blue	10 µL