Expression and characterization of fungal exocellulases

Sonali Kiran Lagu

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

for The Degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

December, 2014

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Sonali Kiran Lagu

Entitled: Expression and characterization of fungal exocellulases

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Selvadurai Dayanandan

_____ External Examiner
Dr. Vladimir Titorenko

_____ Examiner
Dr. Joanne Turnbull

_____ Examiner
Dr. Justin Powlowski

_____ Supervisor
Dr. Reginald Storms

Approved by _____
Dr. Selvadurai Dayanandan, Graduate Program Director

January 19, 2015 _____
Dr. André Roy, Dean of Faculty

ABSTRACT

Expression and characterization of fungal exocellulases

Sonali Kiran Lagu

Although plant biomass is among the earth's most abundant renewable resources, the high cost of converting it to usable sugars currently prevents it from being used as an effective replacement of petroleum as a feedstock for the production chemicals, power and fuels. One way of reducing costs is to improve the performance of the enzymes used to hydrolyze plant biomass. Bacteria and fungi that degrade plant biomass produce mixtures of enzymes, including endoglucanases, cellobiohydrolases and β-glucosidases, which work synergistically to hydrolyze the polymers in plant biomass to monomeric sugars. Biochemical parameters such as cell wall structure, pH and temperature can critically influence the rate and extent of plant biomass hydrolysis. Cellobiohydrolase (Cbh) types I and II release cellobiose by hydrolyzing cellulose polymers from their reducing and non-reducing ends. Here, I have made an attempt to establish benchmark cellulase hydrolysis assays using commercially available cellulase systems and the model substrate PASC. A library of fungal cellobiohydrolases were cloned and expressed using the filamentous fungus *Aspergillus niger* as the expression host. Activity assays were used to identify cellobiohydrolases that could be expressed as enzymatically active enzymes by *A. niger*. These enzymes were then subjected to biochemical characterization using enhancement studies with commercial cellulase mixtures. In addition, the biochemical properties of the characterized cellobiohydrolases were correlated with their primary sequence information. After initial screening based on levels of expression and enzyme activity, six cellobiohydrolases were selected as promising targets. Enhancement studies of commercial cellulase mixtures using four of these enzymes was done using two experimental designs. Although the properties of the

different cellobiohydrolases were similar in many respects, there were significant differences in expression levels and relative activities, which may make some more attractive as candidates for cellulase system enhancement.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to all the people who have helped and supported me through this long journey. My special gratitude belongs to my supervisor, Dr. Reginald Storms, whose expertise, support, patience and constructive guidance enhanced considerably my graduate experience.

My sincere gratitude for the unconditional support and guidance provided by my supervisory committee members Dr. Justin Powlowki and Dr. Joanne Turnbull.

A big thanks to my lab mates, Dr. Edith Munro, Dr. Sophia Ushinsky, Dr. Simon Bourque, Greg Bukewitsch, Lina Mougharbel, Rebecca Sydenham, Shaghayegh Ostadjoo and Yun Zheng. I thank my colleagues in Dr. Powlowki's lab, Yue-Lei and Racha Cheikh-Ibrahim for sharing their knowledge. Special thanks to Racha for all her unconditional help and understanding. I thank Dr. Enongene Ekwe for his knowledgeable and witty discussions. Appreciation is extended to my friends in Dr. Tsang's lab; Juan for help during lab work and Raji and Amrita for their support throughout the graduate experience. Special thanks to my friend Dr. Paul Riddleberger for all his help.

Finally, a very special thanks to my dearest family; my Mother who was there, taking care of my home and me so I could succeed, my husband Kiran, and son Siddharth for always being there for me. Finally I thank the Almighty God, for making it happen.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A. niger *Aspergillus niger*

BCA bicinchoninic acid

Bgl β-glucosidase

CBH Cellobiohydrolases

CBM Carbohydrate Binding Domain

CM Catalytic Module

CMC Carboxymethyl cellulose

DP Degree of polymerization

DS Degree of synergy

EG Endoglucanases

LPMO Lytic Polysaccharide Monooxygenase

GH Glycoside hydrolases or glycosyl hydrolases

MALDI Matrix-assisted laser desorption/ionization

PASC Phosphoric Acid Swollen Cellulose

PMO Polysaccharide Monooxygenase

PCR Polymerase chain reaction

pNPC 4-Nitrophenyl β-D-cellobioside

RAC Regenerated-amorphous cellulose

SDS- PAGE Sodium dodecyl sulphate- polyacrylamide gel electrophoresis

TCA Trichloroacetic acid

TLC Thin Layer Chromatography

T. reesei *Trichoderma reesei*

Introduction

Ethanol produced from cellulosic biomass can be a sustainable source of energy and a potential substitute for petroleum (Banerjee G et al. 2010); however, costs associated with the enzymatic hydrolysis of cellulosic biomass into the fermentable glucose used for biofuels production remains a major challenge faced by the biofuels industry. The objective of my MSc research was to screen for cellulases that synergistically enhanced the cellulose hydrolyzing efficiency of a commercial cellulase system.

Bioethanol produced using cellulosic feedstocks has several advantages relative to transportation fuels derived from petroleum. These advantages include: the reduction of greenhouse gases emissions by 8% to 30% depending on the type of renewable feedstock and the production method (Borjesson P et al. 2009, Fairley P 2011, Kim S and Dale B 2005, and Curran MA 2007); reduced environmental impact from spills because bioethanol is biodegradable and can easily be diluted to non-toxic concentrations; improved trade balances; and, reduced dependence on foreign oil mitigates national security issues related to reliance on imported oil (Demirba A 2008).

Producing ethanol from sucrose and starch means that the feedstocks used for ethanol production are also used as food for humans and as feed for animals, therefore producing bioethanol from these feedstocks has resulted in the 'Food versus Fuel' controversy (Tenabaum DJ 2008, Mudge SM 2008). An attractive alternative to using cornstarch and sucrose from sugarcane feedstocks is to use a variety of non-food substrates that are economical, locally available and produced in large quantities. Several efforts have been made to use agricultural waste such as corn stover, corn cobs, sugarcane bagasse, forest industry waste such as beetle

infested lodgepole pine wood (Kumar A 2009), and graminaceous plants or grasses such as giant reed (*Arundo donax*), switchgrass (*Panicum virgatum*) and miscanthus (*Miscanthus giganteus)* (Raghu S 2006). A major challenge associated with using these abundant but complex feedstocks is achieving efficient hydrolysis.

Biomass Composition:

The major components of lignocellulosic biomass are cellulose (40-50%), hemicellulose (20-30%) and lignin (20-30%) (Kumar R et al. 2008). Minor components are organic extractives such as terpenes and inorganic components such as ash. The structure of the plant cell wall and the relative proportions of its major components vary depending on the plant species, development stage, cell type and the local environmental conditions (Markus P 2008). The general arrangement of the three major components is a hetero-matrix (Figure 1). Cellulose chains, formed of exclusively glucose molecules connected by 1-4 beta-linkages compose the crystalline and amorphous microfibrils. Hemicellulose forms cross-links between cellulose microfibrils and lignin. Hemicellulose and lignin are connected by ester and ether linkages. The cellulose, hemicellulose and lignin are embedded together in a matrix. The main function of the cellulose and hemicellulose polymers are to provide load-bearing strength to the cell wall whereas, lignin provides mechanical support, and contributes to pathogen resistance and vascular function (Figure 1).

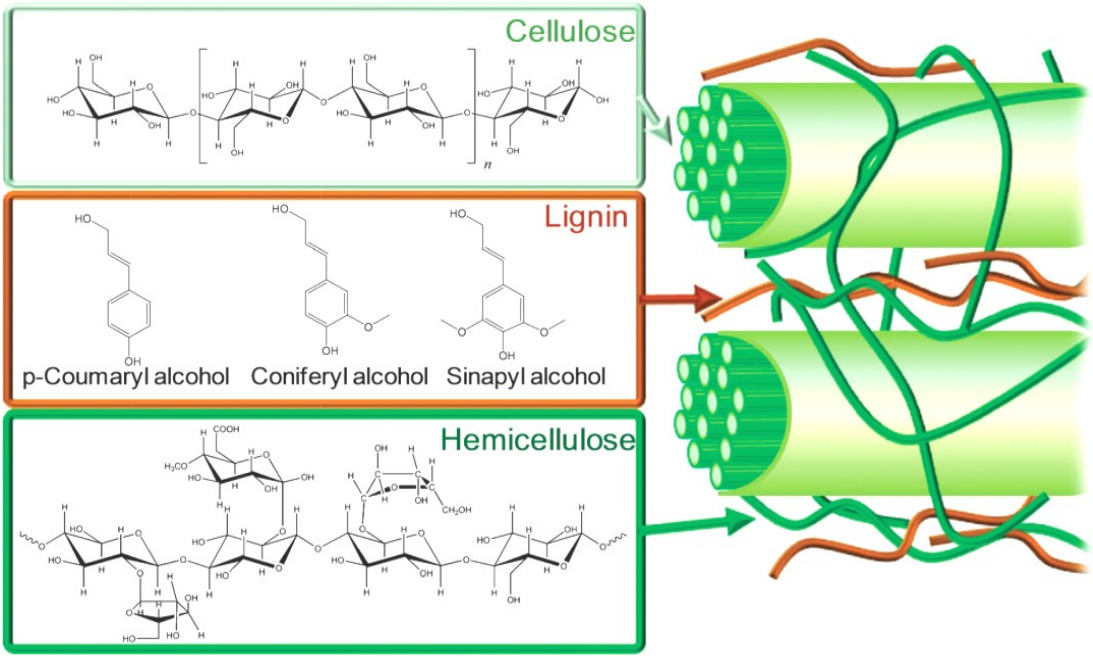


Figure 1. **Structure of lignocellulosic biomass.** The building blocks of cellulose, hemicellulose and lignin, are shown. Cellulose is a linear glucose polymer where glucose residues are linked by β-1,4-glycosidic bonds. Hemicellulose is a branched polymer (five and six carbon sugars), which serves as an interface between lignin and cellulose and is chemically bonded to lignin. Lignin is a polymer of phenylpropane units linked primarily by ether bonds**.** Figure is adapted from "Bimetallic catalysts for upgrading of biomass to fuels and chemicals." (Alonso DM et al. 2012) *Chem Soc Rev 41: 8075-8098* with permission of Royal Society of Chemistry.

Molecular structure of cellulose:

Cellulose is a linear homo polysaccharide consisting of numerous (up to 20,000) ß-1-4 linked glucose units (Zhang YHP 2004). Cellulose is widely distributed in plants, marine animals such as tunicates, and in some algae, fungi, bacteria and protozoans such as *Dictyostelium discoideum* (Habibi Y 2010). The cellulose chains cluster together by van der Waals interactions between the hydrophobic faces of the glucose residues and inter chain hydrogen bonds, giving rise to crystalline microfibrils (Parthasarthi R 2011; Chundawat SPS 2011). Based on X-ray scattering data each vascular plant microfibrils consists of about 36 cellulose chains (Fernandes AN et al. 2011, Somerville C 2004). Each microfibril is composed of about 6 elementary fibrils twisted together, where each elementary fibril is composed of about 6 parallel cellulose chains (Jarvis MC 2013). Microfibrils linked by hemicelluloses form

macrofibrils (bundles of microfibrils). In vascular plants, the elementary fibrils have diameters of about 2 ηm, microfibrils have diameters of about 5 ηm and macrofibrils have diameters up to 1 μm (Figure 2).

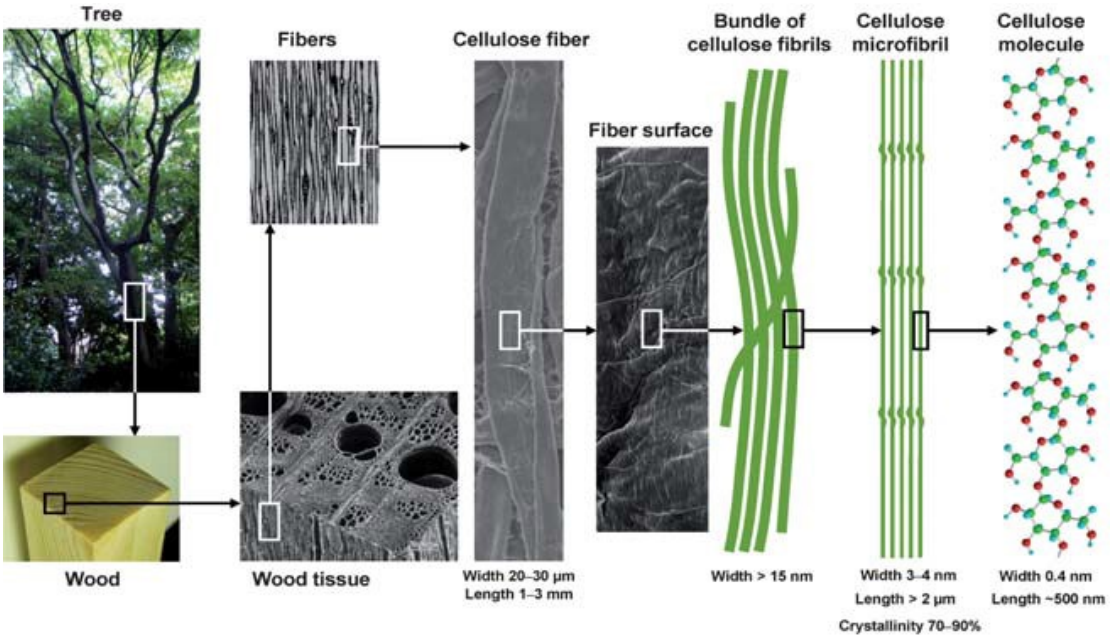


Figure 2. Hierarchical structure of wood biomass and the characteristics of cellulose microfibrils. Cellulose fibers are composed of cellulose macrofibrils. Each cellulose macrofibril is made of many cellulose microfibrils. Each cellulose microfibril consist of about 6 elementary fibrils, and, each elementary fibril (not shown in this figure) is composed of about 36 parallel cellulose chains. Figure is adapted from "TEMPO-oxidized cellulose nanofibers" (A. Isogai et al., 2011) Copyright 2010, Royal Society of Chemistry with permission.

Cellulose chains are synthesized by cellulose synthase, which extends the growing cellulose polymer by adding glucose units to the non-reducing end (Koyama M et al. 1997). The 4C1 chair conformation adopted by D-glucopyranose rings lead to positioning of hydroxyl groups in the equatorial plane, and hydrogen atoms in the axial (vertical) plane (Habibi Y 2010). Each D-glucose residue is added via a β- (1-4) linkage and in crystalline cellulose each glucose unit of a cellulose polymer is rotated by 180 degrees relative to its neighbours. Hence the

disaccharide cellobiose rather than glucose is the repetitive unit of cellulose (Figure 3). Depending upon the source of cellulose the number of glucose units that constitute each polymer varies from 100 to 20,000 (Zhang YHP 2004). The number of glucose units in a cellulose polymer is called the "degree of polymerization" (DP).

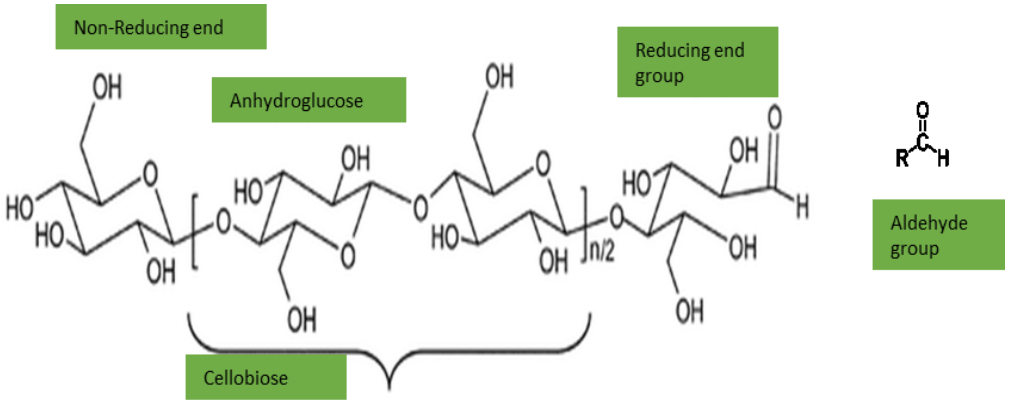


Figure 3. The structure of cellulose. Cellobiose is the structural repetitive unit of cellulose. Adjacent anhydroglucose units are linked by β-1-4 linkages. Each cellulose chain has a stable closed ring structure at one end, the non-reducing end, and an unstable free anomeric carbon hemiacetal at the other end, the reducing end. Figure is adapted from "Cellulose—model films and the fundamental approach" (E. Kontturi et al., 2006) Copyright 2006, Royal Society of Chemistry with permission.

The cellulose chain contains three types of anhydroglucose units: reducing end glucose units with an aliphatic structure that can dissociate to from an open structure with an aldehyde (Figure 3) at the C-1 carbon atom (Figure 4); non-reducing end glucose units with a closed ring structure having a free hydroxyl on the C-4 atom (Figure 4); and, internal anhydroglucose units that are linked together by glycosidic bonds between the C-1 and C-4 carbons of adjacent glucose residues (Figure 3 and 4). The cellulose chains, with a reducing end and a non-reducing end, give each chain a polarity.

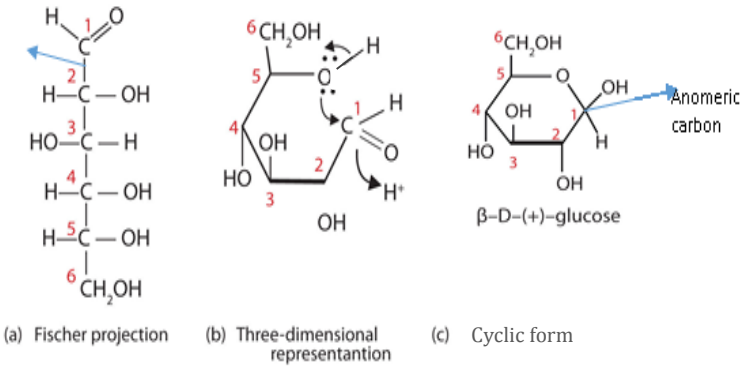


Figure 4. **D-Glucose structure.** D-Glucose structure can be represented with (a) a Fischer projection or (b) three dimensionally (c) β-D-glucose (cyclic monosaccharide) is produced by reacting the OH group on the fifth carbon atom with the aldehyde group. During conversion of carbohydrate molecule (above glucose) from acylic to cyclic form the carbonyl carbon transforms into a new stereocentre or chiral centre referred to as Anomeric carbon atom. Figure was adapted and modified from http://creativecommons.org/licenses/by-nc-sa/3.0/.

In vascular plants the cellulose microfibrils contain both crystalline and amorphous regions. The amorphous regions are regions where inter chain dislocations occur in the microfibrils. As a result of internal strain produced by tilting and twisting more disorder occurs towards the surface than in the core of each microfibril (Fernandes AN et al. 2011).

Six different polymorphs of cellulose are found: cellulose I, II, III_i, III_{ii}, IV_i, and IV_{ii}. These polymorphs result from variation in hydrogen bonding pattern and molecular orientation depending on the source of cellulose, method of extraction and treatment. In nature mainly cellulose I (Figure 5) is found. There are two suballomorphs of cellulose I, I-α and I-ß, where the proportion of I-α varies from about 20% in cotton to about 65% in Valonia (Brett CT 2000). Although the cellulose chains adapt a parallel arrangement in both I-α and I-ß polymorphs of cellulose, they differ in hydrogen bonding pattern. The unit cell structure of I-α is triclinic P1 containing only one cellulose chain per unit cell (Figure 6a), whereas I-ß is monoclinic P2 and contains 2 chains per unit cell (Figure 6b). I-α is a metastable form and can be converted to the

I-ß form by twisting in a plane that is at 39° orthogonal to the hydrogen-bonded plane of the cellulose sheet (Sugiyama J 1990; Jarvis M C 2000)

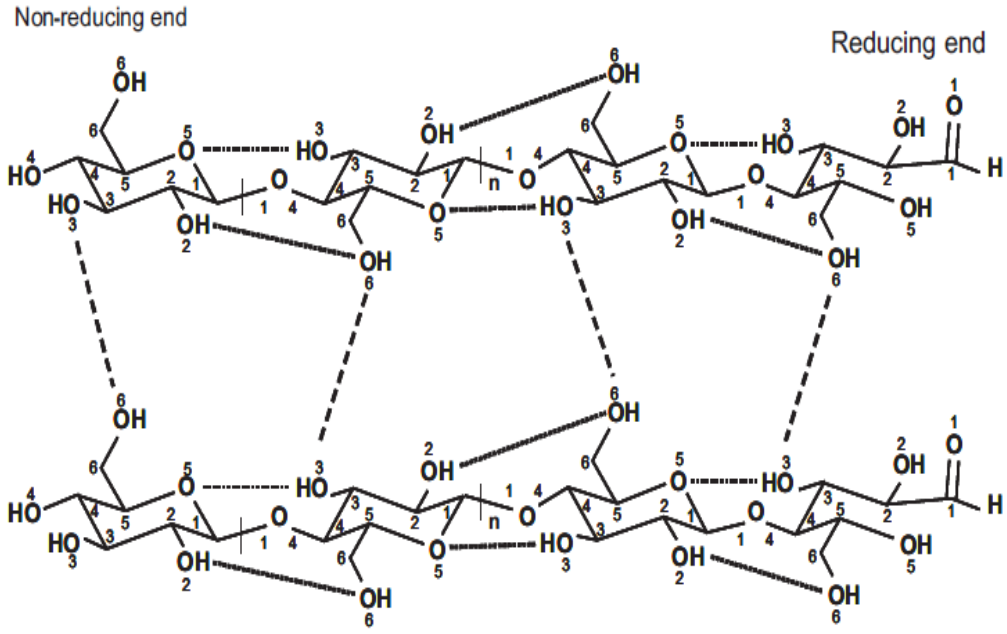


Figure 5. The inter- and intra-chain hydrogen bonding pattern in cellulose I. Dashed lines represent inter-chain hydrogen bonding. Solid and dotted lines represent intra-chain hydrogen bonding. Figure was adapted and modified with permission from "Structure, organization, and functions of cellulose synthase complexes in higher plants" (Festucci-Buselli, R.A., Otoni, W.C., and Joshi, C.P. 2007) *Brazilian Journal of Plant Physiology,* 19(1): 1-13).

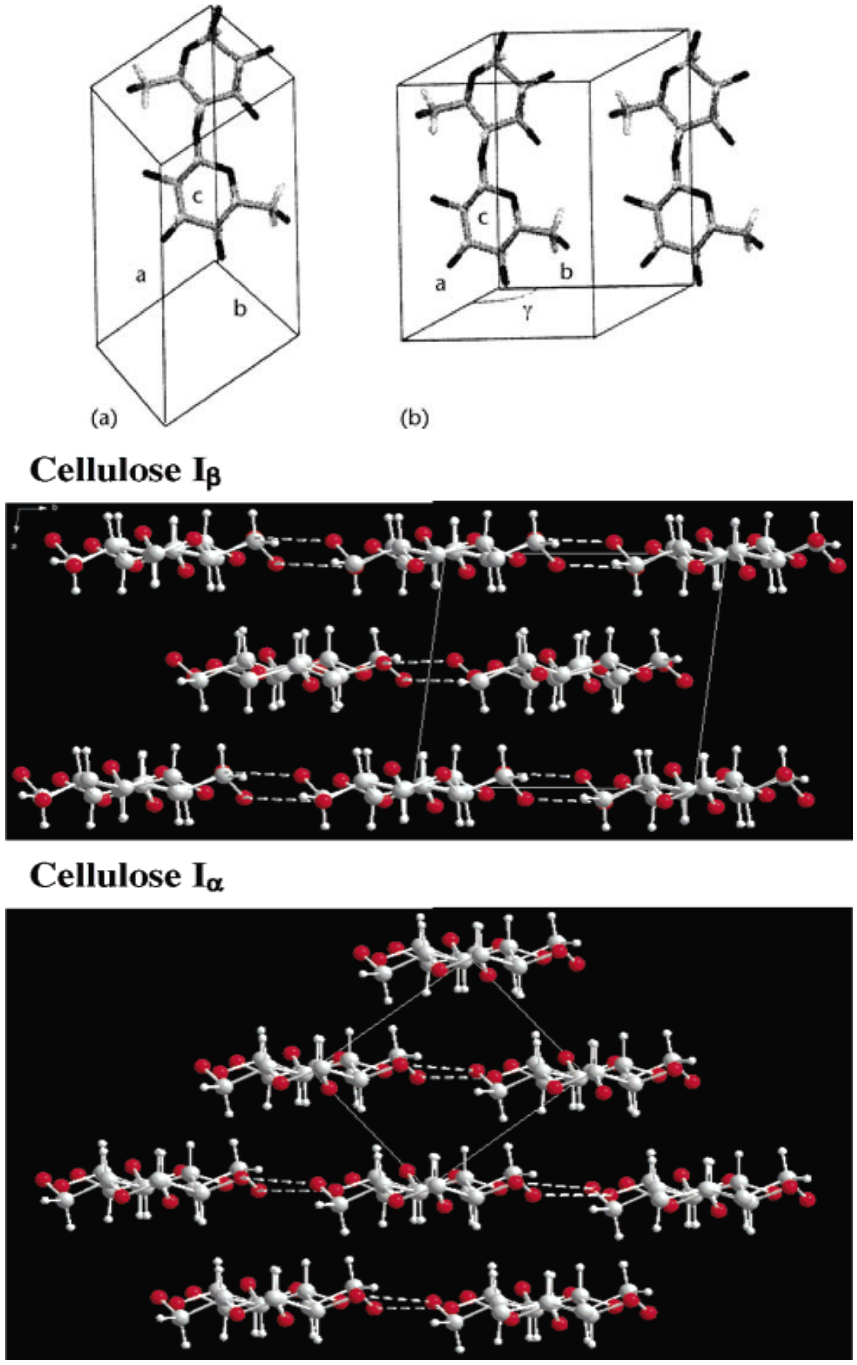


Figure 6. Structures of cellulose Iα and cellulose Iβ. Panel a, the triclinic P1 structure of crystalline cellulose Iα. Panel b, the monoclinic P2 structure of crystalline cellulose Iβ. Panel cellulose Iβ, the covalent and hydrogen bonding of crystalline cellulose 1β. Panel Cellulose Iα, the covalent and hydrogen bonding of crystalline cellulose Iα, C, O, and H atoms are represented as gray, red, and white balls respectively. Covalent and hydrogen bonds are represented as full and dashed sticks.. Adapted with permission from "Cellulose IIII crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction." (Wada M et al. 2004) Macromolecules 37: 8548-8555 Copyright (2012) American Chemical Society.

Cellulose II is the most crystalline and thermodynamically stable form of cellulose. Cellulose II can be obtained by (i) chemical regeneration by dissolving cellulose in a chemical solvent such cupric hydroxide or aqueous ammonia and then reprecipitating in water, or (ii) mercerization by swelling cellulose I using swelling agents like concentrated sodium hydroxide or 65% nitric acid and then removing the swelling agent (Chedin 1959). Cellulose II has an anti-parallel arrangement of chains in the cellulose fibres. This arrangement is more stable than cellulose I, which has a parallel arrangement of chains. Cellulose I has mainly O6-H---O3 inter-chain bonding whereas in cellulose II O6-H---O2 inter chain bonding is dominant. The rigid and linear structure of both the cellulose I and II polymorphs is attributed to intra-chain bonding (Langan P et al. 2001; Nishiyama Y et al. 2002, 2003), with O3-H---O5 intra chain bonding being present in both polymorphs and O2-H---O2 also present in cellulose I

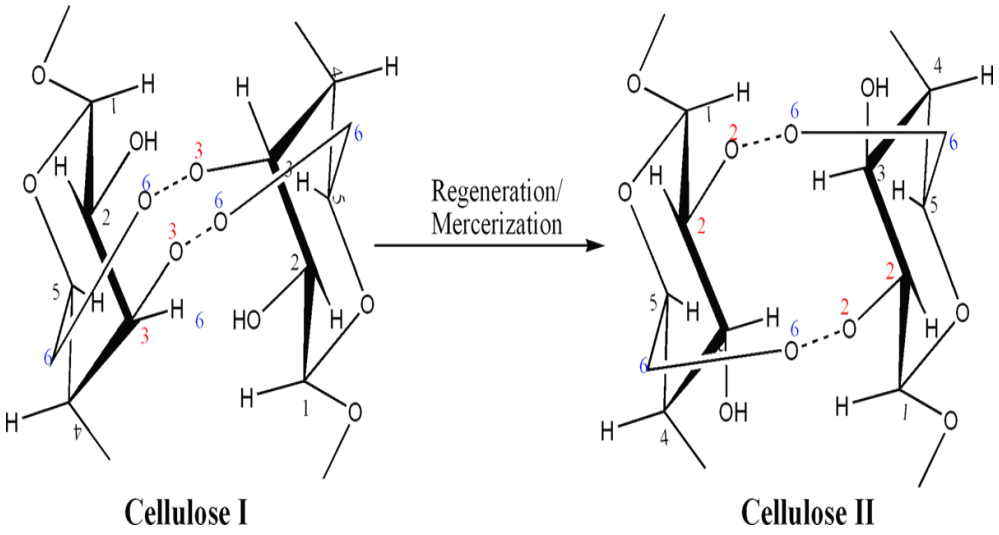


Figure 7. The structures of cellulose I and cellulose II. The major supramolecular distinction between cellulose I and cellulose II is the presence of O6-H---O3 inter-chain hydrogen bonding in cellulose I versus O6-H---O2 hydrogen bonding in cellulose II. Adapted with permission from "Modification of cellulosic fibers by UV-irradiation. Part II: After treatments effects" (Zuber M et al. 2012) Copyright (2012) Elsevier.

The cellulose III_i and cellulose III_{ii} polymorphs of cellulose can be obtained by treating cellulose I and II with swelling agents such as ammonia or other amines and then removing

them. Depending upon the starting form of cellulose either of the polymorphs can be obtained i.e. cellulose I would give cellulose III_i and cellulose II would give cellulose III_{ii}.

Cellulose IV_i and IV_{ii} can be obtained by heating cellulose III_i or cellulose III_{ii} to 260°C in glycerol (Gardiner ES and Sarko A 1985). Cellulose III_i or cellulose III_{ii} can be reverted to native cellulose I and cellulose II by evaporating excess ammonia (- NH_3 in Figure 8) (Chanzy H 1979; Helbert W 1997).

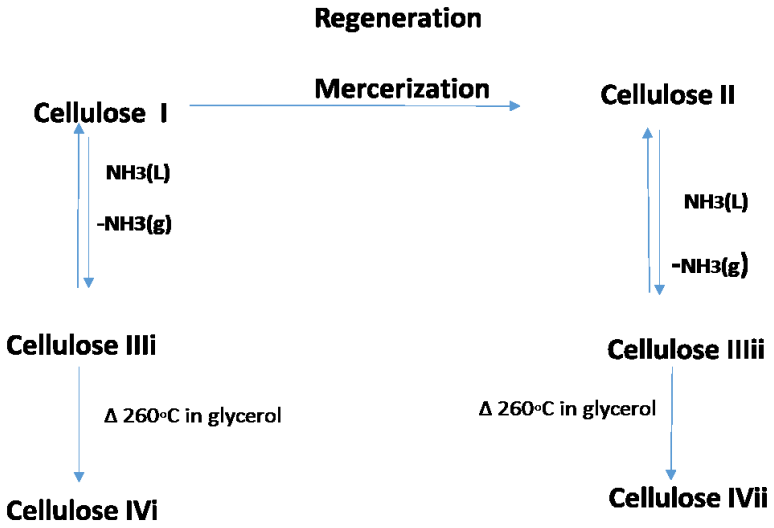


Figure 8. Interconversion of the polymorphs of cellulose. Cellulose IIIi or cellulose IIIii can be obtained from cellulose I and cellulose II, respectively by NH_3 treatment. Cellulose IIIi and cellulose IIIii can be further converted by heating at 260ºC in glycerol to cellulose IVi and cellulose IVii, respectively. NH_3 (L) stands for the liquid ammonia used for treatment and -NH_3 (g) for evaporated ammonia during reversible reaction. Reprinted from "Cellulose: the structure slowly unravels" (Helbert W 1997), Copyright (1997) with permission from Springer.

Hemicellulose:

Hemicellulose is a term used for non-cellulosic polysaccharides. Hemicelluloses vary greatly from each other in their structural and physiochemical properties. Unlike cellulose they are highly branched and are composed of polymers with 500-3,000 sugar units. They are heterogeneous polysaccharides that can have both 5- and 6- carbon sugars. Xyloglucans are abundant in the primary walls of non-graminaceous plants, galactomannans are abundant in the cell walls of solanaceous plants, and mixed-linkage β-glucans are found in the cell walls of graminaceous plants. Xylans are abundant in grasses and woody plants, whereas glucomannans are abundant in gymnosperms (softwoods like spruce and pine). The hemicellulose heteropolymers consist of pentoses such as xylose and arabinose, and hexoses such as mannose, glucose and galactose. Hemicelluloses are easier to degrade than cellulose; however, due to complex branching and acetylation patterns their hydrolysis requires enzyme cocktails that include xylanases and accessory enzymes that can remove branching residues and acetylation from the xylan backbone. Further, hemicellulose hydrolysis yields a mixture of different sugars including pentoses, which are difficult to ferment (Varnai A 2011).

Lignin:

Lignin is complex polymer of aromatic alcohols known as monolignols. It is a relatively hydrophobic aromatic branched polyphenolic macromolecule assembled by mainly beta-O4-aryl linkages. Lignin contains subunits derived from the methoxylated monolignols p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol. The ratio of monolignols varies depending upon the source of the lignin (Chen F 2007). Lignin is cross-linked with the cell wall hemicelluloses and cellulose by ester, ether and glycosidic linkages (Takahashi N and Koshijima T 1988; Sun X et al. 2005). Depolymerization of lignin by microbes in nature is well documented

(Martinez A et al. 2005, Fuchs G et al. 2011). But industrial-scale lignin depolymerization poses a major challenge, because most of the enzymes that act on lignin are co-factor dependent oxidoreductases that are expensive. Present lignin depolymerization efforts are therefore directed at developing thermochemical methods such as pyrolysis, gasification, hydrogenolysis, chemical oxidation and hydrolysis under supercritical conditions (Pandey MP 2011).

Enzymatic Depolymerization of Cellulose:

Enzyme technology seems to be more suitable than thermochemical methods for the depolymerization of lignocellulosic substrates, because enzymatic hydrolysis preserves the carbohydrate structure in the form of monomeric sugar units and does not produce by-products that inhibit the growth of ethanol fermenting yeast (Horn SJ et al. 2012). Recent nano-meter scale atomic microscopy studies (Ding S et al. 2012) showed that the preferred binding sites for microbial cellulases are the hydrophobic planar faces of cellulose molecules. Although these sites are exposed in the primary cell wall, they are covered by lignin in the secondary cell wall, which makes it difficult for enzymes to access the cellulose in secondary cell walls.

The enzymatic degradation of cellulose requires enzymes that can break its ß-1-4-linkages; however, since most vascular plant cellulose is buried within a matrix of lignin, pectin and hemicellulose, a battery of enzymes consisting of ligninases, pectinases and hemicellulases is required to expose the cellulose. The exposed cellulose can then be hydrolysed by cellulases to yield sugars.

A cellulase enzyme system is the collection of cellulase enzymes produced by a cellulolytic strain of bacteria or fungus (Zhang YHP et al. 2006). In general there are two different types of cellulase enzyme systems, namely “anaerobic cellulase enzyme systems” and “aerobic cellulase

enzyme systems" (Zhang YHP et al. 2006). Anaerobic cellulase systems are high-molecular weight, multi-enzyme complexes called cellulosomes; whereas, aerobic cellulase systems are composed of several enzymes that do not physically interact. Recent microscopic studies of delignifiied cell walls showed that cellulosomes hydrolyze cellulose at the cell wall surface and around the plamodesmata, whereas aerobic cellulases systems act by penetrating deep into the cellulose microfibrillar network (Ding S et al. 2012). The collection of cellulase enzymes that can be present in cellulase systems can be divided into two major groups based on the type of reaction catalyzed, the glycosyl hydrolase cellulases (GHs) and the lytic polysaccharide mono oxygenase cellulases (LPMOs).

Glycosylhydrolase cellulases:

The cellulose active GHs include endoglucanases, exoglucanases/cellobiohydrolases, and cellobiases/ß-glucosidases. The endoglucanases (Egs) (EC 3.2.1.4) hydrolyze internal β-1, 4-glycosidic bonds in amorphous cellulose and expose new cellulose polysaccharide chain ends thereby creating additional reducing and non-reducing ends (Lin Z et al. 2011). The exoglucanases (EC 3.2.1.91) (Cbh1s and Cbh2s) hydrolyze terminal glucosidic bonds and make crystalline cellulose amorphous thereby providing an entry point for endoglucanases and water molecules. Cbh1s and Cbh2s cleave the exposed reducing and non-reducing ends respectively of the polysaccharide chains produced by endoglucanase action, thereby releasing cellobiose units (Reese T 1975). The ß-glucosidases (EC 3.2.1.21) convert cellobiose and soluble cellulodextrins to glucose monomers (Figure 9) (Reese T 1975).

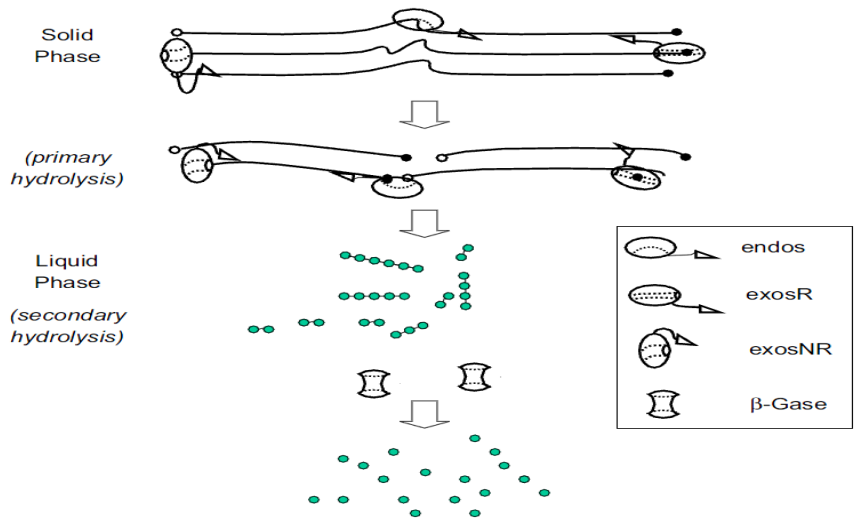


Figure 9: The role of three major cellulases in the hydrolysis of cellulose. The endos represent endoglucanases that act at internal amorphous sites, the exosR and exosNR represent exoglucanases that act at the reducing and non-reducing ends of cellulose polymers, and the β-Gases represent β-glucosidases that act on soluble cellodextrins, mainly cellobiose. Figure was adapted and modified with permission from "Outlook for cellulase improvement: Screening and selection strategies." Zhang Y-P, Himmel ME, Mielenz JR (2006) Biotechnol Adv 24: 452-481

Currently, enzymes from *Trichoderma reesei* and *Aspergillus niger* are used for biomass deconstruction as well as other industrial uses (Lin Z et al. 2011). Enzymes from *Humicola insolens* and *Bacillus sp*. are used for food processing (Lin Z et al. 2011). As a result of multiple rounds of strain improvement, cellulase system production by *T. reesei* has reached to about 100 grams per litre. Although cellulase systems from *A. niger* are used for some commercial applications, due to high levels of cellulase system production by *T. reesei*, it has been the preferred organism for the production of commercial cellulase systems (Zhang YHP and Lynd LR 2004).

The *T. reesei* genome codes for 15 characterized cellulolytic GHs (Glycosyl Hydrolases), two exoglucanases (Cel7A/Cbh1 and Cel6A/Cbh2) (EC 3.2.1.91, CBH), six endoglucanases (Cel7B/Eg1, Cel5A/Eg2, Cel12A/EG3, Cel45A/EG5, Cel74A and Cel5B) (EC 3.2.1.4, EG), and

seven β-glucosidases (EC3.2.1.21, Bgl) (Ouyang et al. 2006). Cbh1, Cbh2, Eg1 and Eg2 are the predominant cellulases in the cellulase system produced by *T. reesei* (Nidetzky B et al. 1994).

The Cbhs progressively cleave the cellulose polymers from the reducing and the non-reducing ends (Divne C et al. 1998; Jeoh et al. 2007), while the Egs cleave the ß-glycosidic bonds internal to the cellulose chain (Karlsson J et al. 2002). Short, soluble oligosaccharides cellobiose, cellotriose and cellotetrose are generated by the combined action of the Cbhs and Egs on cellulose polymers. The soluble oligosaccharides are then cleaved into glucose monomers by the Bgls (Cel3A, Cel1A, Cel3B, Cel3C, Cel1B, Cel3D and Cel3E) (Takashima S et al. 1999; Foreman et al. 2003). These 15 GHs cooperatively hydrolyse the crystalline cellulose in the plant cell walls (Takashima S et al. 1999; Foreman PK et al. 2003). All *T. reesei* cellulases except Eg3 and the seven Bgls have both a catalytic module (CM) and a cellulose-binding module (CBM) (Lynd LR 2002).

T. reesei secretes lower levels of ß-glucosidase than *A. niger*; however, *A. niger* produces lower levels of endoglucanase activity (Kumar R et al. 2008; Zhang YHP and Lynd L R 2004). Also, the major secreted ß-glucosidases from *T. reesei* are more sensitive to glucose inhibition (product inhibition) (Van Zyl WH et al. 2007) than are the ß-glucosidases produced by *Aspergillus species*. ß-glucosidases, Bgl1 and Bgl2 isolated from *Aspergillus aculeatus* by (Sakamoto et al. 1985) were active on soluble cellooligosaccharides ranging from cellobiose to cellohexose and insoluble cellooligosaccharides as long as 20 glucose units.

A recombinant *T. reesei* strain expressing *A. aculeatus* Bgl1 was constructed by Nakazawa H et al. 2012. The culture supernatant obtained from this recombinant strain grown on a mixture of Avicel and xylan as the substrate was able to hydrolyse NaOH treated rice straw at a

significantly lower enzyme loading than was obtained with the supernatant from the parent strain lacking the *A. aculeatus Bgl1* gene. Using a genome based enzyme discovery approach Tambor JH et al. 2012 screened for novel fungal endoglucanases that in binary combinations with *T. reesei* Cel7A increased hydrolysis rates relative to rates obtained with *T. reesei* Cel7A and Cel5A. They identified three fungal endoglucanases that gave increased hydrolysis rates relative to that obtained with TrCel5A/Eg2. Based on the above findings it should be possible to formulate new cellulase cocktails that are capable of cellulose saccharification efficiencies greater that that obtained with the *T. reesei* cellulase system.

The oxidative cellulases:

Chitin is one of the biosphere's most abundant biopolymers and is an important structural component in a diverse set of organisms, including: the cell walls of fungi; the exoskeletons of crustaceans such as crabs, lobsters and shrimps; the exoskeleton and wings of insects; the radula of mollusks; and the internal shells of cephalopods including squids and octopuses. Chitin $(C_8H_{13}O_5N)_n$ is a long chain of N-acetyl glucosamine (2-(acetylamino)-2-deoxy-D-glucose) residues linked by β-1, 4 linkages similar to those found between glucose resides in cellulose. The amine group allows increased hydrogen bonding between adjacent polymers increasing the strength of the chitin matrix. The glucose residues of chitin are modified by replacing a hydroxyl group with an N-acetyl group (Figure 10).

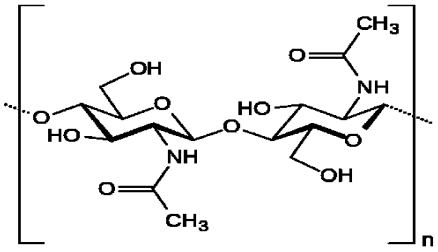


Figure 10. Structure of chitin molecule. Structure of the chitin molecule, showing two β-1, 4 linkaged *N*-acetylglucosamines, the repeat unit in chitin chains. Reprinted from http://en.wikipedia.org/wiki/File:Chitin.svg

Like the cellulase enzyme systems chitinolytic enzymes systems have evolved in bacterial and fungi. The most well characterized chitinolytic enzymes system is that of the enterobacteria *Serratia marcescens*. The *S. marcescens* chitinolytic enzyme system includes at least 5 glycosylhydrolases (4 GH18 chitinases, ChiA, ChiB, ChiC1 and ChiC2, and one GH20 chitobiase, Chb) (Fuchs RL et al. 1986, Kless H et al. 1989). In 2005 Vaaje-Kolstad et al. discovered that *S. marcescens* produces a protein (CBP21) that increases substrate accessibility and promotes chitin hydrolysis. This protein was originally classified as a family 33 carbohydrate-binding module (CBM33). Moser F et al. 2008 showed that *Thermobifida fusca* CBM33 proteins could act as a potentiating factor for chitin hydrolysis by chitinases (Sanchez B et al. 2011) and, possibly, cellulose hydrolysis by cellulases. Although bacteria and viruses have been known to have CBM33 encoding genes they are rare in eukaryotes., Fungal proteins formerly classified as family 61 glycoside hydrolases (GH61) that are structurally similar to CBM33 proteins have shown to enhance cellulose hydrolysis by cellulases (Harris PV et al. 2010).

Until 2010, it was unclear how CBM33s and GH61s worked; however, ((Vaaje-Kolstad et al. (2010), Forsberg Z et al. (2011), Quinlan RJ et al. 2011 and Vaaje-Kolstad et al. (2012)) showed that CBP21 is an enzyme that cleaves glycosidic bonds in chitin in an oxidative manner,

generating a normal non-reducing chain end and a chain end comprising a C1-oxidized sugar called aldonic acid. These studies also showed that the activity of CBP21 is boosted by adding electron donors such as ascorbic acid and that enzyme activity depends on the presence of polysaccharide monooxygenases (LPMOs) and the CAZy database has reclassified CBM33 LPMO proteins (Agger JW et al. 2014), as family 10 auxiliary activity enzymes or AA10 enzymes (Levasseur A et al. 2013). Similar results have been obtained for the cellulose active GH61 proteins, which have now been reclassified as lytic polysaccharide monooxygenases and assigned by CAZy to auxiliary activity enzyme family 9 (AA9 enzymes). AA9 enzymes like AA10 members require bivalent copper (Cu^{+2}) co-ordinated in the active site (Li X et al. 2012). AA9 proteins also need an electron donor such as ascorbic acid, reduced glutathione, or gallic acid to boost their activity (Quinlan RJ et al. 2011).

Oxidative cleavage occurs when a C-C bond is broken to form a C-O bond. CBP21 cleaves glycosidic bonds in chitin by oxidation (molecular O_2 is involved) generating a non-reducing end and a C1-oxidized aldonic acid end (Figure 12). CBP21 activity depends on the presence of bivalent copper (Cu^{+2}) coordinated in the active site; however, addition of electron donors like ascorbic acid further boosts its activity (Vaaje-Kolstad G et al. 2010) (Figure 11).

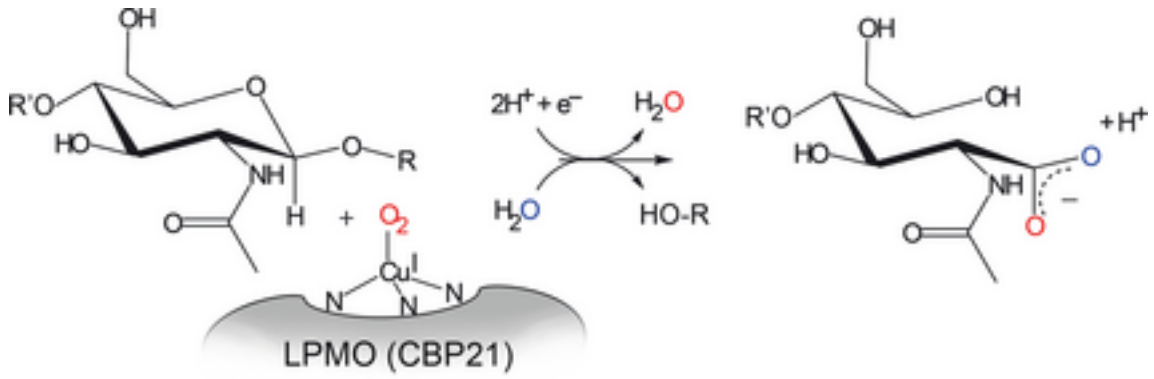


Figure 11. Schematic overview of the reaction catalyzed by CBP21 a LPMO. The enzyme carries a reduced copper ion that transfers an electron to molecular oxygen, which results in forming a superoxo intermediate that initiates cleavage of the substrate. A non-oxidized product ('HO-R') with a newly generated non-reducing end is produced and a product where the downstream end is oxidized to yield an aldonic acid. The aldonic acid contains one oxygen coming from O_2 (red) and one as a result of hydrolysis by bulk water (blue). Note that GH61-type LPMOs seem to vary with respect to which side of the scissile bond they oxidize: some oxidize C1 (as CBP21), others oxidize C4 or C6, on the downstream side of the scissile bond. Figure reproduced with permission from "NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions" (Aachmann F L et al. 2012) Proc Natl Acad Sci USA 109: 18779-18784

In case of oxidation of glucose at C1 position by AA9 proteins, lactones are produced which are then either spontaneously or enzymatically hydrolysed to aldonic acid. C4 oxidation results in production of ketoaldoses (Dimarogona M et al.2012). The inhibitory effect of C-4 oxidised sugars (4-ketoaldoses) has not been demonstrated so far (Figure 12).

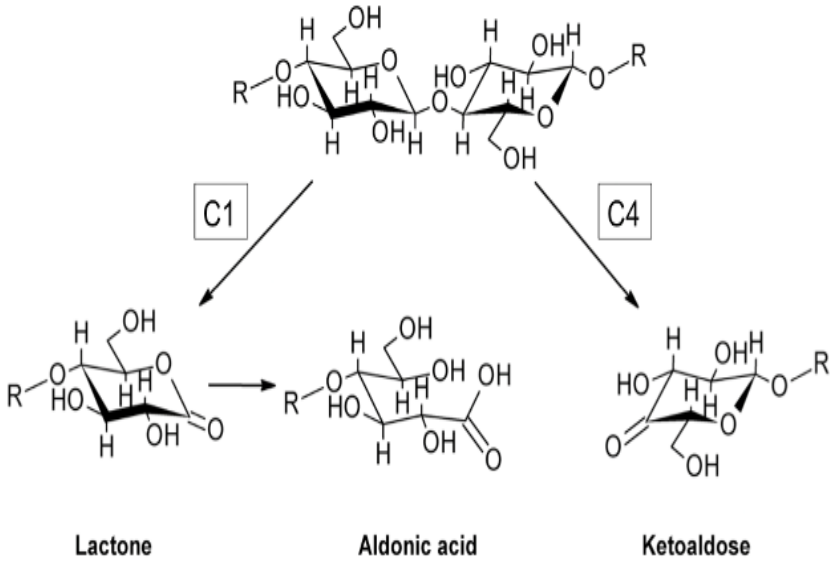


Figure 12. C1 and C4 oxidation by GH61 enzymes. C1 oxidation lactones are produced that spontaneously or enzymatically get hydrolysed to aldonic acid and C4 oxidation results in production of ketoaldoses. Reprinted with permission from "Cellulose degradation by oxidative enzymes" Dimarogona, Maria, Evangelos Topakas, and Paul Christakopoulos. *Computational and Structural Biotechnology Journal* 2 (2012): e201209015. *PMC*. Web. 4 Dec. 2014.

Glucanases need pretreated substrates with disrupted crystallinity whereas AA10 and AA9 enzymes can act on carbohydrate chains at the surface of more compact and inaccessible crystalline materials. Interestingly, *Hypocrea jecorina* (*T. reesei)*, the strain used for the production of commercial cellulase systems expresses only two LPMO/GH61 proteins. To address the low GH61 activity present in the *T. reesei* cellulase system current cellulase preparation, Cellic CTec2 by Novozymes, is supplemented with extra GH61 enzyme. A recent study showed an increased cellulose conversion when LPMO/GH61 containing Cellic CTec2 was applied on pretreated wheat straw under conditions as close as possible to a bioethanol production setup (Cannella D et al. 2012). However, it was found that gluconic acid produced as a result of C1 oxidation had increased inhibitory effect on cellulose conversion. Also, the production of gluconic acid was inhibited at increased temperatures (Cannella D et al. 2012).

Recently, Hemsworth GR et al. 2013 reported a new family of LPMO enzymes (AA11) from *Aspergillus oryzae* (*koji* mold) that have characteristics somewhat in between those of the AA9 and AA10 families. Although the overall fold and metal coordination were basically similar to the known LPMOs, the new family has a slightly convex surface where the copper active site sits. In the future it is likely that auxiliary family enzymes like the AA9, AA10 AA11 enzymes will make a major contribution to reducing the cost of biomass conversion.

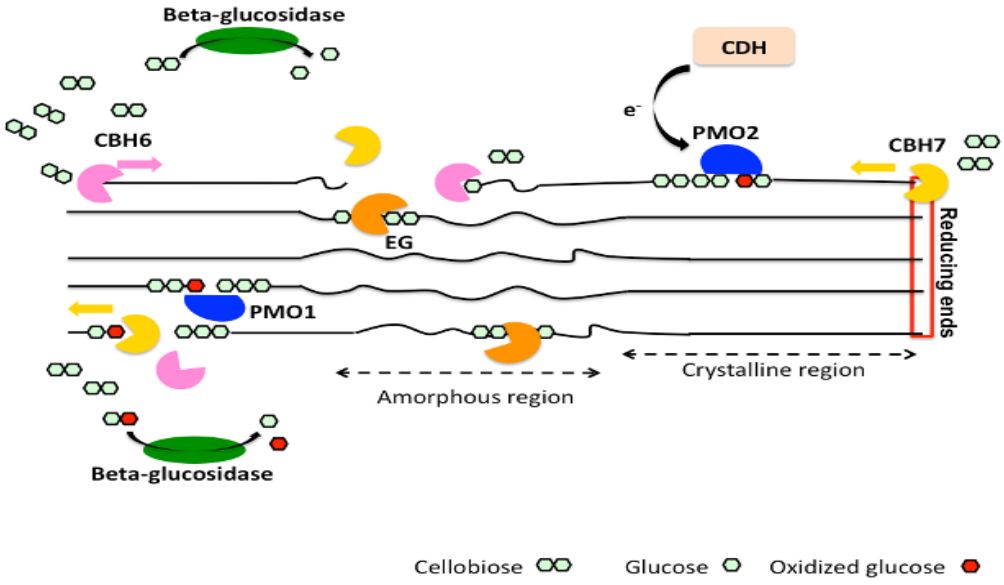


Figure 13. A simplified scheme of the current view on the enzymatic degradation of cellulose, involving cellobiohydrolases (CBH), endoglucanases (Eg), type 1 and type 2 LPMOs (LPMO1 and LPMO2, respectively). Cellobiose dehydrogenase (CDH) is a potential electron donor for PMOs. Egs and PMOs cleave cellulose chains internally releasing two new chain ends that are targeted by CBHs. CBHs generate cellobiose or oxidized cellobiose that is subsequently hydrolyzed by β-glucosidase. Reprinted with permission from "Cellulose degradation by oxidative enzymes" Dimarogona, Maria, Evangelos Topakas, and Paul Christakopoulos. *Computational and Structural Biotechnology Journal* 2 (2012): e201209015. *PMC*. Web. 4 Dec. 2014.

Structural organization of cellulases:

The gene sequence of fungal and bacterial cellulases reveals that many of these proteins have two independent domains, a CM (Catalytic Module) and a CBM (Carbohydrate Binding Module) connected by a highly glycosylated flexible linker peptide (Gilkes NR et al. 1991) (Beguin P 1990) (Figure14).

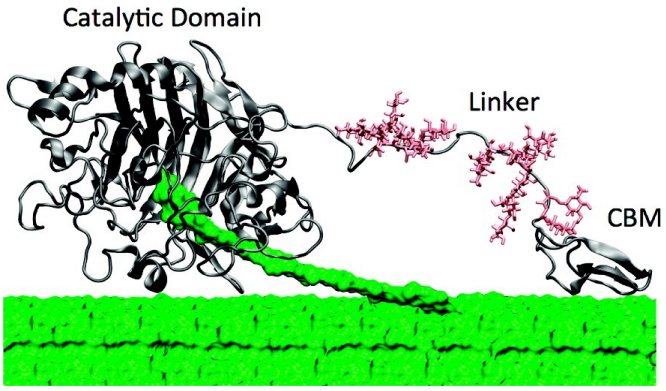


Figure. 14. The cellobiohydrolase I enzyme (Cbh1) from *T. reesei*. Cbh1 consists of three sub-domains: left, the catalytic domain; middle, the linker peptide with O-glycosylations (pink); and right, the carbohydrate binding module. A single cellulose chain from the cellulose microfibril (green) is hypothesized to thread into a tunnel in the CD; the chain is cleaved and the two-sugar product, cellobiose, is expelled from the left end of the tunnel. Reprinted with permission from "Identification of Amino Acids Responsible for Processivity in a Family 1 Carbohydrate-Binding Module from a Fungal Cellulase." (Beckham GT et al. 2010) J Phys Chem B 114: 1447-1453 Copyright (2010) American Chemical Society.

The most recognized function of the CBMs is to bind to the substrate and bring the biocatalyst in the vicinity of its substrate, thereby effectuating carbohydrate hydrolysis. CBMs are thought to have other functions including increasing enzyme concentration at the surface of substrate, the non-hydrolytic disruption of the crystalline substrate, and alignment of the catalytic module with the substrate (Lynd L et al. 2002; Boraston A et al. 2004). CBMs are specific for insoluble cellulose and can be classified into type A CBMs, those that interact with crystalline cellulose, and type B CBMs, those that interact with non-crystalline cellulose. These non-catalytic modules form a complex with the substrate by forming non-covalent

thermodynamically favorable bonds. Reflecting the importance of CBMs, removal of the CBM from *T. reesei* Cbh1 reduces the catalytic activity of enzyme on insoluble substrates (Mansfield SD et al. 1999). Interestingly, removal of CBM did not affect the hydrolysis rate on a soluble substrate (Raghothama S et al. 2000; Irwin DC et al. 1993).

Using infra-red spectroscopy and x-ray diffraction, researchers (Gao PJ et al. 2001; Wang et al. 2008) followed the structural and morphological changes observed on cotton fibres following treatment with a purified CBM derived from *T. reesei* Cbh1. These studies found that the CBM promoted the non-hydrolytic disruption of crystalline cellulose by disrupting the weak interchain bonds in crystalline cellulose. Lee I et al. 2000 using atomic force microscopy showed that elongated holes were formed throughout the surface of cotton fibres following treatment with hexachloropalladate-inactivated *T. reesei* Cbh1. They suggested that this was due to the penetration of the CBM into the cellulose fibres. In contrast, they found no effect on the surface of cotton fibres when they were treated with bacterial cellulase from *Thermotoga maritima*, which lacked a CBM. These observations highlight the importance of the Cbh1 CBM in the non-enzymatic disruption of crystalline cellulose.

CBMs are grouped into approximately 68 sequence-based families on the Carbohydrate–Active Enzyme (CAZY) Web server (http://www.cazy.org/Carbohydrate-Binding-Modules.html). CBM1 is found almost exclusively in fungi and when fungal cellulases possess a CBM it is almost always a family 1 CBM.

The peptide linker connecting the CBM and GH catalytic domain plays an important role in the synergistic function of these two domains. Studies have shown that altering the linker length, changing the amino acid content or separating the two domains by enzymatic cleavage of the linker, reduces the overall rate of hydrolysis (Gilkes NR et al. 1988; Tomme P et al. 1988).

This suggests that altering or cleaving the linker may affect enzyme binding to the substrate thereby lowering the rate of carbohydrate hydrolysis (Shen H et al. 1991; Srisodsuk M et al. 1993).

Sequence comparison between bacterial and eukaryotic linker regions shows that bacterial linkers have more than twice the proline content but fewer putative O-glycosylation sites than eukaryotic linkers (Sammond DW et al. 2012). It has been proposed that the uniform distribution of O-glycosylations along the linker is required for proteolysis protection and the maintenance of an extended structure. The terminal ends of the linker regions are more glycine rich and have fewer potential glycosylation sites suggesting that the linker region terminal regions require flexibility to achieve proper orientation between the CBM and CM (Sawmond DW et al. 2012).

Glycoside Hydrolase families:

Glycoside hydrolases (Figure 15) are a widespread group of enzymes that catalyse glycosidic bond cleavage between two or more carbohydrate moieties or between a carbohydrate and non-carbohydrate moiety. Earlier glycoside hydrolases were classified based on their substrate specificities or the catalytic mechanism employed; however, this failed to reflect the structural features of these enzymes. Classification based on sequence similarity has allowed the glycoside hydrolases to be classified into 131 families, which in turn are grouped into 14 ‘clans’ based on the respective protein folds of their catalytic cores (Figure 16). The number of clans is much smaller than the number of GH families because the catalytic core folds are much more conserved than are the primary structures (www.cazy.org/Glycoside-Hydrolases.html).

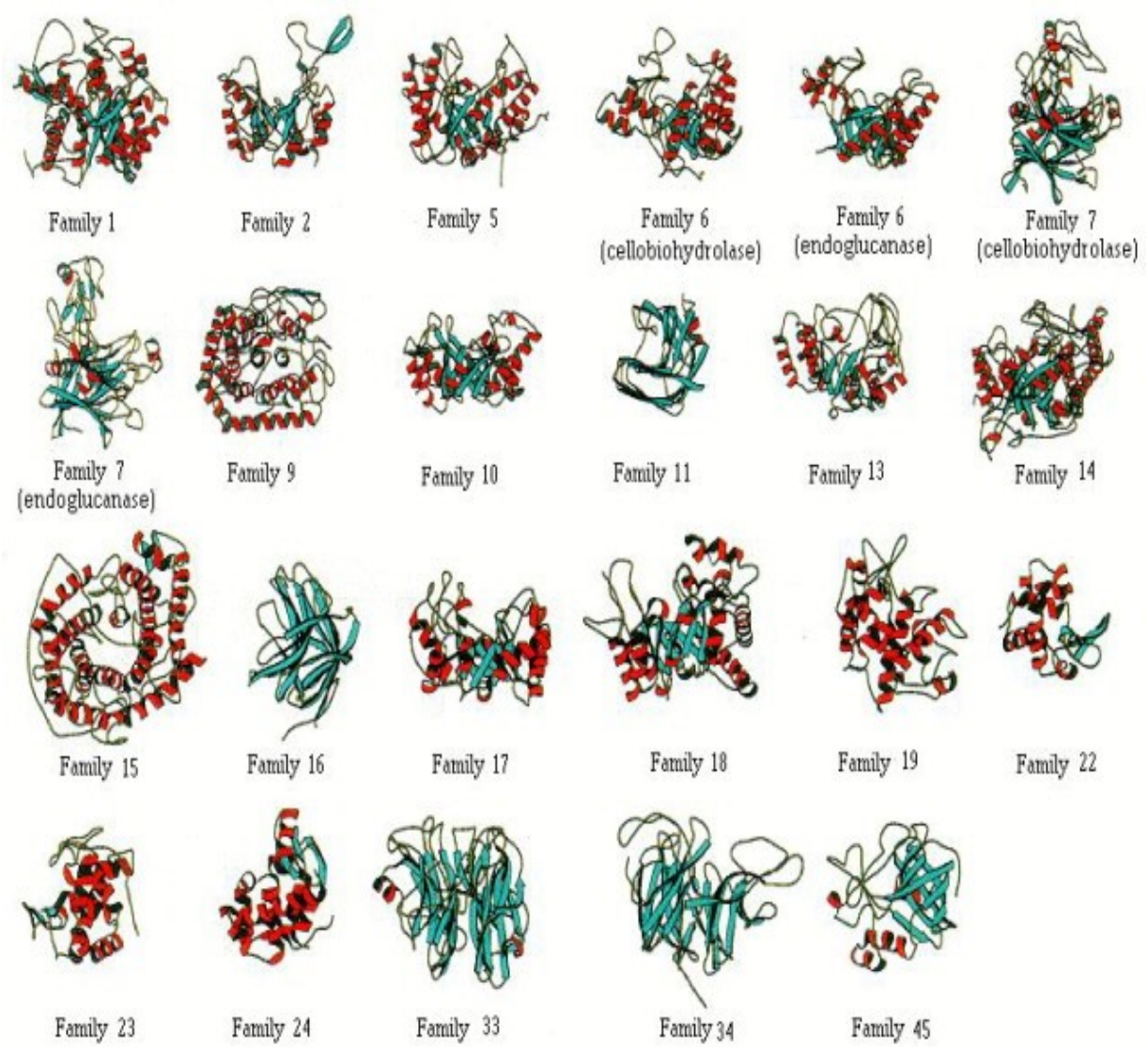


Figure 15: Ribbon representation of the main fold of the catalytic domain in various GH families. Beta strands are shown in cyan and, α helices are in red. Reprinted from "Structure and mechanism of glycosyl hydrolases" (Davies G et al.1995) *Structure 3: 853-859* Copyright (1995), with permission from Elsevier.

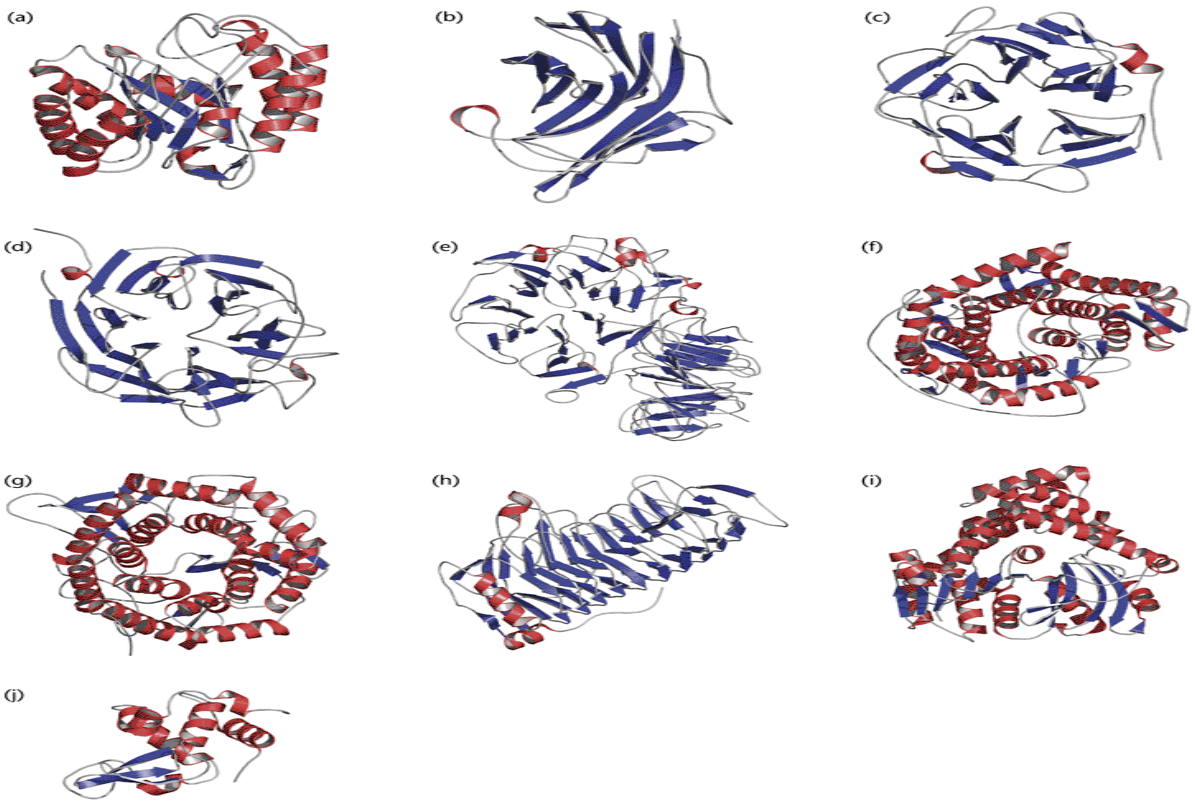


Figure 16. Glycosidase folds for 10 'Clans': (a) $(\beta/\alpha)_8$ barrel; (b) jelly roll; (c) 5-fold propeller; (d) 6-fold propeller; (e) 7-fold propeller; (f) $(\alpha/\alpha)_6$ barrel; (g) $(\alpha/\alpha)_7$ barrel; (h) right-handed β-helix; (i) dehydrogenase-like fold; and (j) lysozyme-like fold. Reprinted from "Glycosidases: Functions, Families and Folds" Hancock, S. M. and Withers, S. G. (2007). eLS. .

Of the 14 'clans,' four of the clans, K, L, M and N have similar folds to that of clans A, G, G and H but differ in the location of their catalytic acid/base or nucleophile or may have insertion or deletions in their classic folds (Hancock SM et al. 2007). An Enzyme classification (EC) number has been assigned to most biochemically characterized enzymes according to the chemical reaction they catalyse. Each enzyme's EC classification serves as accession number for the enzyme database BRENDA (http://www.brenda-enzymes.info/). Different enzymes catalysing the same reaction receive the same EC number and a single enzyme catalysing two different reactions receives two different EC numbers.

Information about enzyme substrates, products, inhibitors and other functional data can be obtained from the database by using the EC number. Names of enzymes can reflect the GH family to which they belong: for example, the *Clostridium thermocellum* Egl cellulase has a catalytic core that belongs to GH family 44 so it is called Cel44A and *T. reesei* Cbh1 cellulase has a catalytic core that belongs to GH family 7 so it is called Cel7A. The "A" signifies that it is the first enzyme belonging to this family to be reported from this organism. Since the division is based on the amino acid sequence of the catalytic core of the enzyme it is possible that enzymes with different catalytic mechanisms are grouped in one family, for example, GH families 6 and 7, both contain cellobiohydrolases and endoglucanases, or that enzymes with similar catalytic mechanisms are grouped in two families, for example GH family 1 and 3, both contain ß-glucosidases.

Regardless of their mechanism three active site topologies have been proposed for different glycoside hydrolases, namely, pocket or crater, cleft or groove, and tunnel (Figure 17) (Davies G. 1995). The pocket or crater active site topology usually identifies enzymes that act on the non-reducing end of a polysaccharide chains. Pocket or crater class active sites are found in enzymes like ß-galactosidases, ß-glucosidases and neuraminidases. They are also found in exopolysaccharidases such as glucoamylase and ß-amylase, which are not very efficient on fibrous substrates but work well on substrates such as starch grains that have large number of free chain ends available at the surface (Davies G 1995). The cleft or groove topology is commonly found in endo-acting polysaccharides such as lysozymes, endocellulases, chitinases, xylanases, alpha-amylases, endo-β-1,3-glucanases, endo-β-1,3-1,4-glucanases and endo-β-1,4-glucanases (Davies G. 1995). The tunnel topology is found in cellobiohydrolases and has apparently evolved from the cleft or groove topology by acquiring a long surface loop that

covers the cleft to form tunnel. The catalytic centres lie enclosed in the tunnel. The tunnel enables the polysaccharide chain to both remain bound to the enzyme and release the hydrolysis product (Davies G 1995)

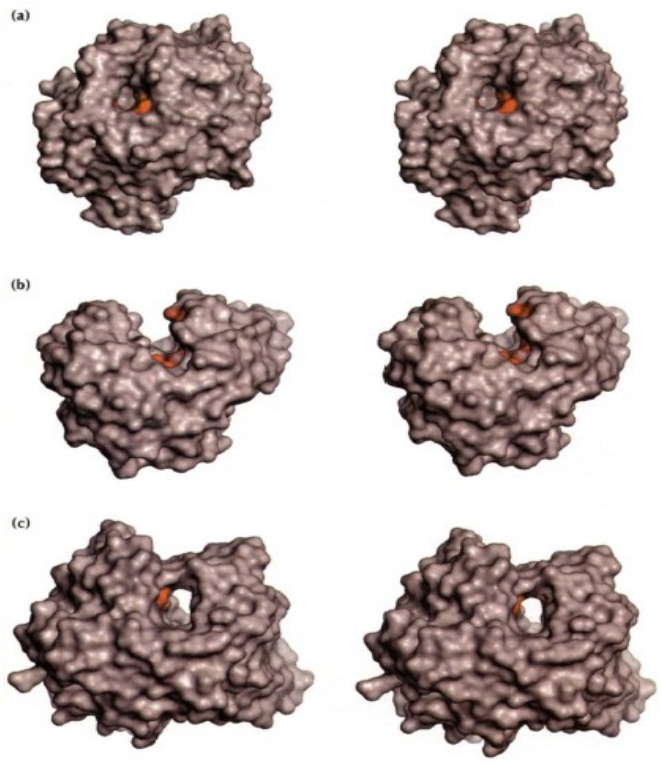


Figure 17. The three types of active sites found in glycosyl hydrolases. (a) The pocket (e.g. ß-glucosidases & glucoamylase). (b) The cleft (e.g. endoglucanases and xylanases). (c) The tunnel (cellobiohydrolase). The proposed catalytic residues are shaded in red. Reprinted from "Structure and mechanism of glycosyl hydrolases" (Davies G et al.1995) *Structure 3: 853-859* Copyright (1995), with permission from Elsevier.

Hydrolytic mechanisms of cellulases:

Acid-base catalysis is used by cellulases for the hydrolysis of glycosidic bonds. Catalysis may be performed by either an inverting or a retaining mechanism. A proton donor (acid) and a nucleophile (base) are two essential residues required for the inverting and retaining mechanisms. Aspartate and glutamate amino acid residues act as acid/base donors in most glycosyl hydrolyse families but are not a conserved feature. GH families 18, 20, 25, 56, 84 and

85 use acetamido or N-glycol group at C-2 of the substrate instead of catalytic nucleophile (Scheltinga A 1995). An activated tyrosine as a catalytic nucleophile is used by the GH 33 and 34 families (sialidases and trans-sialidases) (Watts A et al. 2003 and Amaya M et al. 2004).

The inverting mechanism is a single step displacement mechanism with oxocarbenium-ion like transition state that leads to a change in configuration at the anomeric carbon, C1, from ß to α after hydrolysis. The nucleophile/base deprotonates a water molecule, which in turn attacks the C1 on the glucose ring inverting its configuration from ß to α (Koshland 1953) (Figure 18).

The retaining mechanism is a two-step displacement mechanism with covalent glucosyl-enzyme intermediate and oxocarbenium-ion like transition state which leads to retaining the configuration at anomeric carbon C1 of the substrate after hydrolysis (Koshland 1953; Schulein M 2000) (Figure 18). In the first step of the reaction, the nucleophile/base attacks C1 directly forming a covalent intermediate and releasing glycosyl-enzyme product. In the second step, the catalytic base deprotonates a water molecule as second nucleophile, which in turn attacks the C1 anomeric carbon generating a product having the same stereochemistry as the substrate (Koshland 1953) (Figure 18).

GHs of family 4 and 109 exhibit NAD^+ dependent hydrolysis. NAD is required as a co-factor and remains bound to the enzyme throughout catalysis. The mechanism involves initial oxidation of substrate followed by E1cb elimination with help of enzymatic base. An α, ß–unsaturated intermediate is formed as a result of elimination reaction, which then undergoes addition of water at the anomeric centre. Finally the ketone at C3 is reduced to generate sugar products. Thus although elimination mechanism is employed the final outcome is hydrolysis.

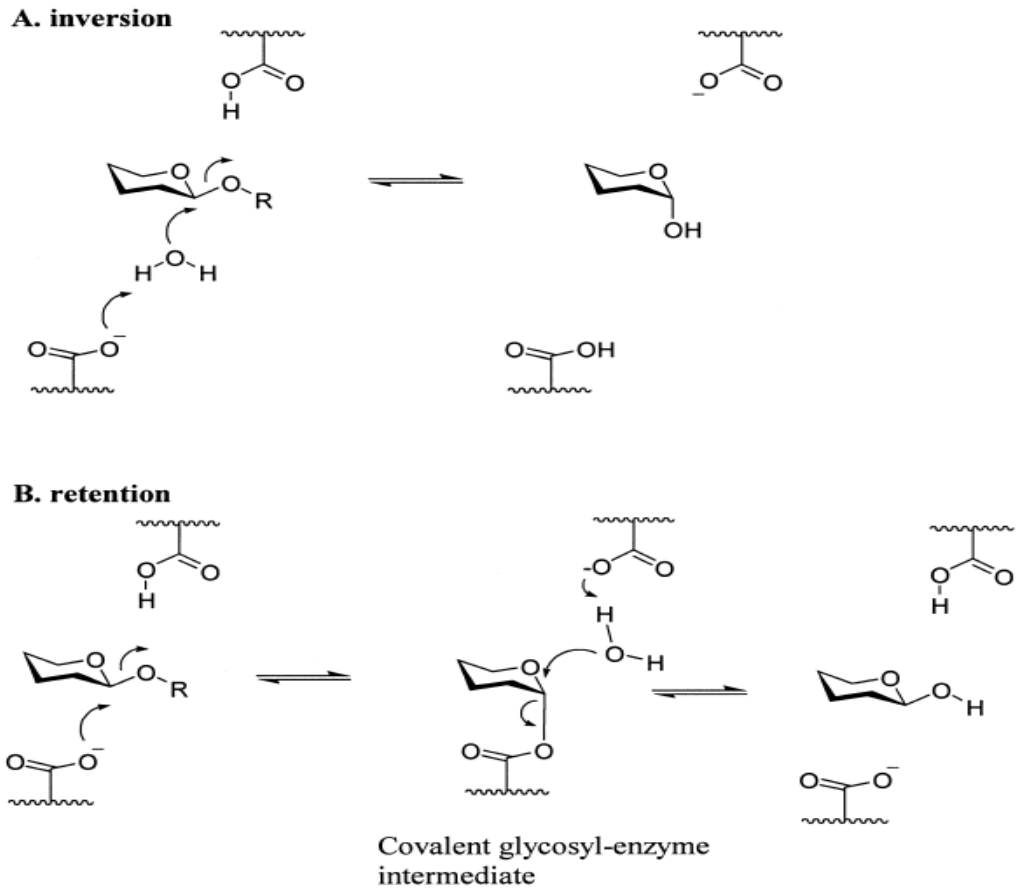


Figure 18. The two major mechanisms used for the enzymatic hydrolysis of glycosidic bonds. (a) Inverting mechanism. (b) Retaining mechanism. Reprinted with permission from "Protein engineering of cellulases" Schulein M (2000). *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* 1543: 239-252, Copyright (2000), from Elsevier.

Cellobiohydrolases:

Cellobiohydrolases have been classified as Cbh1 and Cbh2 enzymes based on their preference for hydrolysis at the reducing or non-reducing ends of cellulose chains. In the *T. reesei* cellulase system, Cbh1 represents about 60% of the total mass of cellulase protein whereas Cbh2 accounts for about 20%. Together these two cellulases show exo-exo synergy. It has been

shown that a least some cellobiohydrolases possess endoglucanase activity and may create new cellulose chain ends by cutting the chains in the middle (Bioset C et al. 2000).

Cbh1 enzymes belong to family GH7. This family includes both exoglucanases and endoglucanases that cleave ß-1-4 glycosidic bonds. GH7 exoglucanases use a double displacement or retention mechanism and release mainly cellobiose units from the reducing end of cellulose polymers. Cbh1 enzymes consist of a large catalytic core connected to a small CBM via a 6 to 109 residue glycosylated polypeptide linker. The crystal structure of the Cbh1 catalytic domain reveals a β-jellyroll folded framework in which a highly curved β-sandwich is formed by two ß-sheets, each with seven and eight antiparallel ß- strands, placed face to face (concave and convex faces) (Figure 19). Some loops from the inner β-sandwich extend further and bend to form part of the tunnel (~50 Å) around the active site (Divne C et al. 1998, 1994) and (Figure 15 and 19).

The substrate binding site located within the tunnel can accommodate 6 to 10 glucosyl units. The length of the tunnel and location of active site at its far end makes it possible for the enzyme to proccessively split off and release cellobiose units from the reducing end of the chain whereas glucosyl units upstream of the cleavage site remain attached to the enzyme (Divne C et al. 1998).

Cbh1 structure:

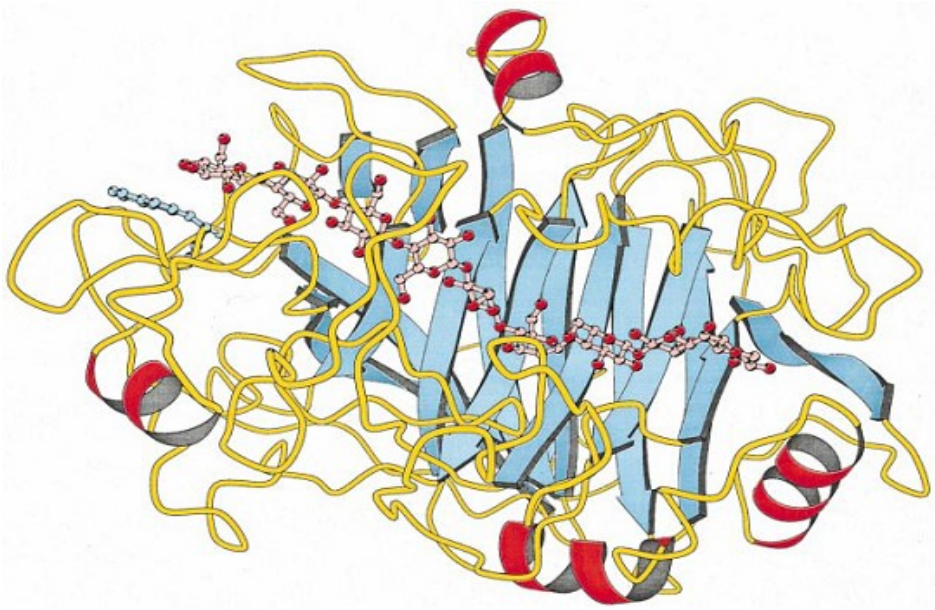


Figure 19. Schematic representation of Cbh1 catalytic domain with a bound cellooligomer. Secondary structure elements are coloured as follows: β-strands are in blue arrows, α-helices are in red spirals, and loop regions are in yellow coils. The cellooligomer is shown in pink as a ball and stick object. "High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *T. reesei"* Divne C, Stahlberg J, Teeri TT, Jones TA (1998) *J Mol Biol 275: 309*-325, copyright 1998, Elsevier.

The *T. reesei* Cbh1 has 11 glucosyl binding sites. Ten sites (i.e. -7 through +3) are located within the tunnel and one site (+4) is outside the tunnel. The sides of the tunnel are made of a complex network formed of hydrogen bonds and salt links and are rich in amino acids especially tryptophan (4 residues) that can interact with sugars. These tryptophan residues are evenly spread throughout the tunnel and are major players for formation of glucosyl binding sites at -7, -4, -2 and +1 (Figure 20).

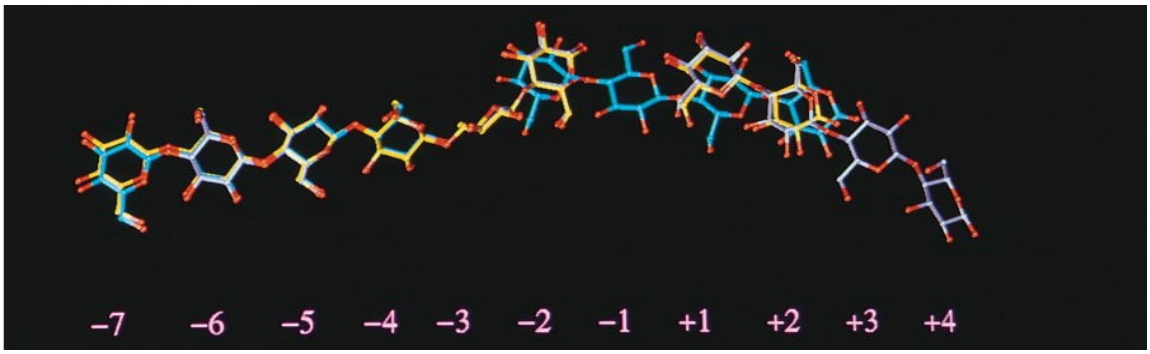


Figure 20. Summary of a crystallographic study showing binding of a cellulose chain in the cellulose-binding tunnel of cellobiohydrolase I (Cbh1). X-ray crystallographic study of long natural oligosaccharides complexed with three mutated enzymes each lacking an essential amino acid, (Glu212, Asp214 and Glu217). The blue, yellow and pink colours show superposition of the cellooligomers obtained in each experimental study. Oligosaccharide binding sites in the Cbh1 tunnel (-7 to +3) where +1,-7,-4 and -2 represent major binding sites for the enzyme to the substrate and the reducing end of the substrate is to the right of the figure. Reprinted with permission from "High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*" Divne C, Stahlberg J, Teeri TT, Jones TA (1998) *J Mol Biol 275: 309*-325, copyright 1998, Elsevier.

The catalytic residues in the tunnel are placed close to each other in a consensus motif Glu-X-Asp-X-X-Glu-, where the first glutamic acid (Glu212) acts as the catalytic nucleophile and Glu217 acts as the general acid/base. Hydrolysis results in inversion or retention of the anomeric carbon atom configuration depending on whether the distance between these catalytic units is about 5 Å or about 10 Å (Divne C et al. 1994; Rouvinen J et al. 1990).

The arrangement of catalytic residues Glu212 and Glu217, in the Cbh1 catalytic site is such that the two residues are place perpendicularly i.e. above and below the glycosidic linkage to be cleaved. For the glycosidic bond cleavage to occur the cellulose chain has to be positioned in such a way that the glycosidic bond at +1/-1 is pointing upward (productive binding mode (Figure 21). The cellulose chain undergoes a right handed twist from entry into the tunnel (-7) towards the exit, by 140°, flipping it upside down. Cel7A enzyme studies have shown that Asp214 is important for catalysis because it maintains the correct positioning and protonation state of the catalytic nucleophile, Glu212 (Divne C et al. 1998; Stahlberg J 1996).

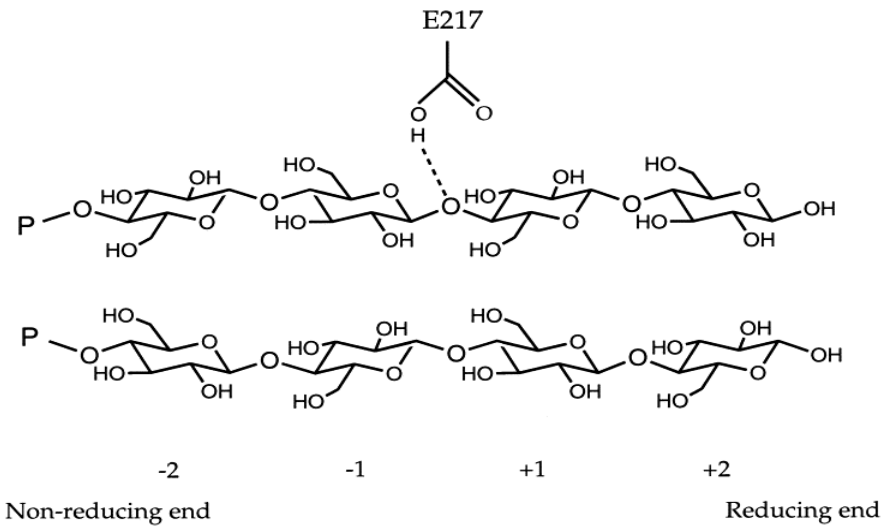


Figure 21. The difference between productive binding (upper panel) and non-productive binding (lower panel) of a cellulose chain. The position of the proton donor Glu217 in respect to the glycosidic linkage at -1/ + 1 is indicated. For clarity, only the glucosyl residues in sites -2 to +2 are shown. P denotes the continuation of the non-reducing end of the cellulose chain beyond site -2. The two saccharide chains have the same directionality with their reducing ends to the right, and the non-reducing ends to the left. The two orientations are related by a 180º rotation. "High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*" Divne C, Stahlberg J, Teeri TT, Jones TA (1998) *J Mol Biol 275: 309*-325, copyright 1998, Elsevier.

The substrate binding sites at the entrance of the tunnel are optimised to permit sliding of cellulose chains through the tunnel. Substrate sliding is aided because the tunnel surface is lined with water molecules. For the cellulose chain to be threaded through the tunnel it has to be lifted from the surface of the crystalline cellulose. The tryptophan residue at position -7 of the substrate binding site forms an overhang outside the tunnel and plays an important role in lifting the chain from the surface of the crystalline cellulose and preventing it from reattaching to the cellulose surface. Passage of the chain through the tunnel is further assisted by the enzyme undergoing a series of conformational changes (Divne C et al. 1998).

The main product formed during Cbh1 hydrolysis, due to the structure of the enzyme and the substrate, is cellobiose. The glycosidic bonds in the cellulose chain have an alternating

orientation and two glucosyl units are required to reach the product site at +1 and +2 in the tunnel where the glycosidic bond linking the glucosyl residues in sites -1 and +1 are presented in suitable orientation for cleavage before the next catalytic event. In the +2 site the guanido group of Arg394 recognises the incoming free reducing end of the substrate group, and forms bidentate (when two ligands are attached on to the central atom) hydrogen bonding with the O6 and O1 hydroxyl groups. Productive binding mode (Figure 21) is a prerequisite state for the incoming group to be recognized at the +2 site (Divne et al., 1994). The cleavage produces cellobiose as a product. Binding in the non-productive mode can result in production of cellotriose, although cellobiose production predominates.

Cbh2 structure:

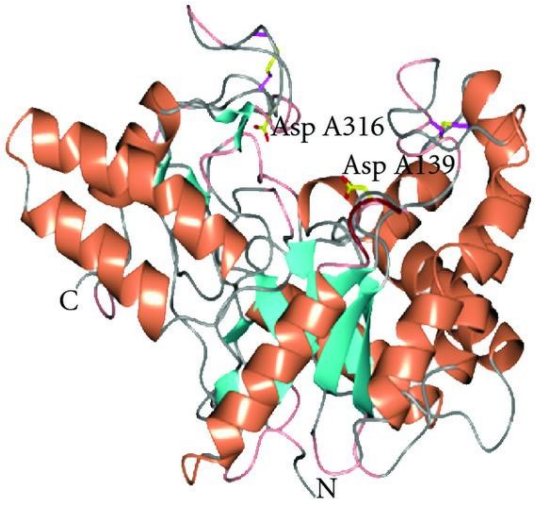


Figure 22: Schematic representation of Cbh2 from *Humicola insolens*. *H. insolens* family 6 endoglucanase Cel6B (PDB id 1DYS). The catalytic residues, Asp A316 (acid) and Asp A139 (potential base), are shown in ball-and-stick representation. Reprinted with permission from "Cellulases from Thermophilic Fungi: Recent Insights and Biotechnological Potential," Duo-Chuan Li, An-Na Li, and Anastassios C. Papageorgiou, Enzyme Research, vol. 2011, Article ID 308730, 9 pages, 2011. doi:10.4061/2011/308730e

Cbh2 enzymes belong to family GH6. The catalytic module of *H. jecorina* Cbh2 (CEL6A) was the first cellobiohydrolase structure to be solved (Rouvinen et al. 1990). This family includes both exoglucanases and endoglucanases and the catalytic mechanism used is net inversion of the anomeric configuration, which results in production of cellobiose. There is no overall sequence homology between the catalytic domains of *H. jecorina* cellobiohydrolases CEL6A and CEL7A.

The crystal structure of the CEL6A catalytic domain shows modified α/β barrel folds with seven β-strands forming the central β-barrel as opposed to the classical $(\beta/\alpha)_8$ "TIM" barrel structure. A 20 Å long tunnel is formed by two extensive loops at the carboxyl terminal end of the β-barrel. The sides of the tunnel are formed by complex hydrogen bonds, salt links

and are lined by water molecules. The tunnel is also rich in amino acids that interact with sugars. There are three tryptophan residues in the tunnel and four substrate binding sites (A, B, C, D) that are positioned inside the tunnel. The tunnel is flattened and the final product cellobiose produced remains enclosed in the tunnel (Divne C et al. 1994). There are six cysteine residues that form two disulphide bridges that stabilize the loops forming the tunnel. The site of glycosyl bond cleavage is between subsites B and C and involves residues Asp175 (Asp139 in Figure 22) and Asp221 (Asp316 in Figure 22) (Rouvinen J et al. 1990). Asp221 can act as proton donor for the catalytic cleavage, although Asp 401 may also act as a general base.

Restricted space in the tunnel and alternate arrangement of glycosidic bonds on the cellulose chains permits formation of correct conformation for hydrolysis at every second glycosyl bond as the cellulose polymer moves proccessively through the tunnel. This results in cellobiose as the principal product released by successive cleavages near the nonreducing end of the substrate.

Cbh2 is less processive than Cbh1 and exhibits endoglucanase activity (Boisset C et al. 2000). This could be attributed to the movement of the loops that result in opening and closing of the tunnel roof. Opening of the tunnel roof has been shown to increase endoglucanase activity (Zou J et al. 1999; Varrot A et al. 1999). There are significant sequence homologies between the catalytic domains of Cbh2/Cel6A cellobiohydrolases and GH6 endoglucanases, and between Cbh1/Cel7A cellobiohydrolases and Eg1/GH7 endoglucanases. The loops forming the active site tunnels in Cel6A and Cel7A are absent in GH6 and GH7 endoglucanases which makes their active site more open and is believed responsible for their increased endoglucanase activity (Meinke A et al. 1995).

Synergism:

Cellulase synergism occurs when two or more cellulases working together show higher activity levels than the sum of the activities of the individual cellulases working under the same conditions (Zhang YHP et al. 2007). Synergism can be binary for example, synergy between endoglucanases and exoglucanases (Nidetzky B et al. 1993; Kim Y et al. 2009). Synergism can also be ternary for example, synergy between exoglucanases, endoglucanases and β-glucosidases (Zhang YHP et al. 2007). The degree of synergism by a cellulase mixture is the ratio between the measured activity of the cellulase mixture and the sum of the activity exhibited by each of the cellulases present in the mixture (Banerjee G et al. 2010). The degree of synergism has been shown to be affected by type of substrate and the type of cellulase (source from which it is obtained).

Synergism between the endo-and exo-glucanases is brought about by co-operative activity, where endoglucanases work on random internal sites creating more reducing and non-reducing ends, which are then hydrolysed by exoglucanases to produce cellobiose and glucose (Zhang YHP et al. 2006). Recent studies on endo-exo synergism using radio labelled bacterial cellulose to measure synergy (Jalak J et al. 2012), showed that *T. reesei* Cbh1 (Cel7A) gets stalled in the amorphous regions of the bacterial cellulose. However, degradation of amorphous cellulose by adding endoglucanases prevents Cbh1 stalling. Valjamae P et al. 1999, showed that a short hydrochloric acid (HCl) treatment of crystalline cellulose increases Cbh1 activity but reduces that of Eg1 (Cel7B). However, a longer HCl treatment gave the opposite effect. Furthermore, increasing the length of the crystalline cellulose pretreatment with HCl reduced the amount of hydrolysis obtained using a combination of Cbh1 and Eg1 (Valjamae P et. al. 1999).

These results suggest that, although CBHs can hydrolyse cellulose chains on the crystalline cellulose surface, their processivity gets obstructed with erosion of the crystalline surface and that endoglucanases can remove this obstacle by degrading the amorphous regions. Additionally, several studies have shown that endocellulases are able to initiate hydrolysis at internal sites thereby increasing access sites for exocellulases (Parsiegla G et al. 1998; Boisset C et al. 2000).

It has been shown that *T. reesei* Cbh1 is able to completely digest some forms of cellulose but *Termobifida fusca* exocellulase (Cel48A) is unable to do so. Both these enzymes hydrolyse from the reducing ends and act synergistically with both Cbh2 and Egl type enzymes (Barr BK et al. 1996). Unlike Cel48A, the *T. fusca* (Cel9A), a processive endoglucanase, was able to synergize with both types of exocellulase and endocellulases (Irwin et al. 1993). Apparently *T. fusca* Cel9A exhibits properties intermediate between endo and exo cellulases (Sakon J et al. 1997). This suggested that the Cel9A could bind anywhere along the cellulose chain and perform cleavage. Following cleavage, the non-reducing end fragment bound to the CBM moves into the active site cleft -1 to -4 thereby allowing processive cleavage of cellotetrose from the non-reducing end (Sakon J et al. 1997). The weak binding between the CBM and the crystalline cellulose substrate may enable the catalytic domain of the enzyme to interact with several substrate molecules near the active site and thereby perform multiple cycles of endo and exo hydrolysis.

In compact crystalline structure, the available surface area for enzyme interaction is limited. Various substrates like PASC, cotton, filter paper, Avicel and bacterial microcrystalline cellulose (BMCC) are heterogeneous and have variable crystallinity. The crystallinity index (CI) is a measure that reflects average crystalline value of the substrate and can be used to compare data obtained by using different substrates. Other than crystalline and amorphous cellulose there is

also intermediate paracrystalline cellulose, which accounts for over 30% of the cellulose in cotton and wood (Larsson P et al. 1997). The presence of paracrystalline cellulose results in failure to obtain the expected changes in the CI of cellulose substrates after enzyme digestion (Park S et al. 2010). Parameters like particle size and surface to volume ratio may influence the CI. Furthermore, CI value determinations can show variations of over 30% between samples of the same substrate. Valjamme P et al. 1999, observed that there is lower synergism on highly ordered cellulose (e.g. Valonia cellulose). According to the recent mechanistic models, the rate limiting step for hydrolysis is the placement of individual chains in active sites (Levine SE et al. 2010).

Cbh1 exocellulases, expansins and swollenins contain polysaccharide binding surfaces that help in non-hydrolytic disruption of the cellulose fibril network (Arantes V and Saddler JN 2010) thereby providing access to the unexposed polymeric chains. It is believed that synergism can only occur when two cellulases attack different regions of the cellulose fibrils.

The degree of synergism is affected by factors such as crystallinity, degree of polymerization of the substrate, the amount of enzyme loading, and the reaction time (Hall M et al. 2010). Lower available surface area for enzyme activity can cause crowding and obstruct enzyme motion. Each cellulase enzyme that gets adsorbed on the cellulose surface will obscure a number of binding sites. Footprints created by *T. reesei* Cbh1 have been measured using small angle X ray scattering (Abuja P et al. 1988 a, b). It is assumed that each glucose unit occupies about 0.25 nm^2 and that the catalytic domain and CBM of Cbh1 occupy the equivalent of 156 glucose units. Enzyme crowding effect studies with *T. reesei* Cbh1 and Eg2 shows that when the cellulose surface area available for binding is low relative to the binding area potential of the cellulases, the binding sites for Eg2 tend to become obscured. As a result, Eg2 cleavage of cellulose chains

is inhibited, which prevents the creation of new chain ends for Cbh1 action, ultimately breaking the synergistic relationship between these two enzymes (Levine SE et al. 2010).

In contrast, when the surface area available for binding is high relative to the binding area potential of the cellulases, synergy between enzymes is maintained. The rate of hydrolysis and degree of synergism are both dependent on the component cellulases present in the multi-enzyme mixture (Bayer E et al. 2004, Detroy RW et al. 1982, and Henrissat B et al. 1985). This makes the optimization of enzyme mixtures an important step in achieving increased hydrolysis rates and glucose production (Din N et al. 1994).

Synergism also occurs between β-glucosidases and endoglucanases and between β-glucosidases and exoglucanases. Addition of Bgl from Taiwanese fungus *Chaetomella raphigera* to *T. reesei* cellulases (endo and exo) showed a high level (8.9 fold) of synergy (Ng I et al 2010). Addition of a *C. raphigera* Bgl to a reaction mixture containing *T. reesei* cellulases (endo and exo) also synergistically enhanced cellulose hydrolysis. Similarly, Chir JL et al. 2011 showed that a 50-100% increased CMC and filter paper hydrolysis was obtained as a result of synergism between both, a wild type or mutant *T. fusca* Bgl (S319C) with Cel9A from the same organism. Cellulase Cel9A of *T. fusca* exhibits dual properties of endo and exo cellulases (Wilson DB 2004). It is believed that synergism observed between Bgls and endo- and exo-cellulases occurs because the Bgls prevent product inhibition by converting the cello-oligomers to glucose.

In recent years, a number of disruptive proteins that belong to families GH61 and CBM33 (discussed earlier) have been discovered. They show potential of working synergistically with other cellulases and degrading recalcitrant biomass. Determining the molecular mechanisms

underlying the synergistic function of these proteins may help us further unravel the yet unknown aspects of synergistic biomass hydrolysis.

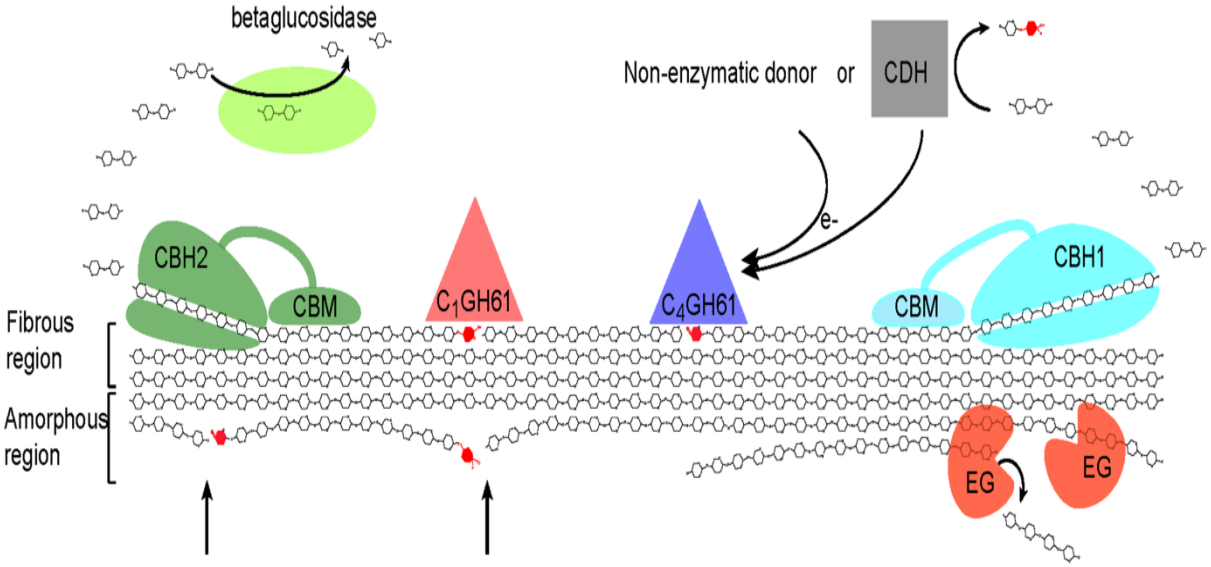


Figure 23. Current view on fungal enzymatic degradation of cellulose. Abbreviations: Eg, endoglucanase; CBH, cellobiohydrolase; CDH, cellobiose-dehydrogenase; CBM, carbohydrate-binding module. The C1-C4 oxidizing action of GH61 (red and purple) is shown which results in producing non-oxidized ends for the Cbh1 (blue) and Cbh2 (green) action producing cellobiose which is acted upon by β-glucosidase to produce glucose (oxidized sugars are colored red). Arrows indicate new attacking points for CBHs. Reprinted with permission from "Novel enzymes for the degradation of cellulose." Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH (2012) Biotechnology for Biofuels 5: 45.

PASC (Phosphoric Acid Swollen Cellulose) as a model substrate:

PASC is prepared by treating crystalline cellulose with 85% phosphoric acid. This results in swelling of the crystalline cellulose without impacting the degree of polymerization (Hall M et al. 2012). Swelling does not drastically alter the overall structure of the cellulose polymers but it does significantly change physical properties of the crystalline cellulose. There is a significant increase in the volume of cellulose sample due to uptake of swelling agent (Zhang YHP et al. 2006). Phosphoric acid treatment turns crystalline cellulose amorphous. Amorphous cellulose has a greater capacity of retaining water as compared to crystalline cellulose (Hall M et al. 2012) and water molecules within the cellulose matrix greatly enhances enzymatic hydrolysis (Hall M et al. 2012). PASC is widely used as an amorphous substrate to check the ability of an enzyme or enzyme system to hydrolyze a cellulosic substrate (Andersen N et al. 2008; Arantes V and Saddler JN 2010). Interpreting the results obtained when comparing the efficiency of cellulose hydrolysis by defined combinations of cellulases is easier if a well-defined substrate like PASC is used.

Thesis objective

The objectives of this work were: I and ii) clone and screen a library of novel heterologous Cbh genes that could be functionally expressed by *A. niger*; iii) establish benchmark cellulase hydrolysis assays using the model substrate PASC that could be used to screen the library of novel cellulases for enzymes that enhanced cellulose hydrolysis by a commercial cellulase system; iv) assess the ability of the functionally expressed novel heterologous Cbh enzymes to synergistically enhance cellulose hydrolysis rates and/or the efficiency of cellulose hydrolysis by a commercially available cellulase system.

Materials and Methods

Materials and Methods

Materials:

Chemicals and other reagents:

Protein quantitation 2-D Quant Kit (80-6483-56) was purchased from GE Life Sciences (the kit includes Bovine Serum Albumin (BSA) protein standard, Precipitant, Co-precipitant, Copper solution and Colour reagents A and B). 4-Nitrophenyl β-D-cellobioside (pNPC) was obtained from Sigma-Aldrich. Phosphoric acid swollen cellulose (PASC) was prepared from Avicel (Sigma) as described previously (Wood 1988) with the following modifications. The 600 mL suspension of 20 grams of Avicel in phosphoric acid was kept on ice for 3 hours with occasional grinding with a pestle and mortar. After centrifugation for 10 minutes at 14,000 g, the Avicel pellet was suspended in two litres of iced cold water and rinsed at 4°C for 15 minutes. The washes were repeated until the pH was between 5 and 7. Finally, the washed PASC was suspended in 10 mM citrate buffer pH 5.0. All the other reagents used were of analytical grade and from commercial sources.

Enzymes:

Purified commercial enzymes, Cbh1, Cbh2, Eg1 and Eg2 from *T. reesei* were provided by Iogen Inc. along with other *T. reesei* cocktails. The cocktails included; the complete *T. reesei* cellulase system, where the major cellulase system enzymes Cbh1, Cbh2, EgI, EgII, EgIII, EgIV, EgV, BglI and BglII represent greater than 97% of the total protein; the cellulase system produced by a *T. reesei* strain with the gene encoding *Cbh1* deleted, where the major cellulase

system enzymes Cbh2, EgI, EgII, EgIII, EgIV, EgV, BglI and BglII represent about 92% of the total protein; the cellulase system produced by a *T. reesei* strain with the gene encoding *Cbh2* deleted, where the major cellulase system enzymes Cbh1, EgI, EgII, EgIII, EgIV, EgV, BglI and BglII represent about 95% of the total protein; the cellulase system produced by a *T. reesei* strain with the gene encoding Egl deleted, where the major cellulase system enzymes Cbh1, Cbh2, EgII, EgIII, EgIV, EgV, BglI and BglII represent about 95% of the total protein; the cellulase system produced by a *T. reesei* strain with its *Eg2* gene deleted, where the major cellulase system enzymes Cbh1, Cbh2, EgI, EgIII, EgIV, EgV, BglI and BglII represent about 96% of the total protein. In addition, cellobiohydrolase I (Cbh1) Lot: 40201 and cellobiohydrolase 2 (Cbh2) Lot: 20201 were purchased from Megazyme (UK) and β-glucosidase (NOVOZYME 188 C-2605) was purchased from Sigma-Aldrich (USA). Accellerase 1500 was obtained as a complimentary sample from Genencor (DuPont). β-glucosidase from *Aspergillus niger*, was also provided by Genencor (now DuPont).

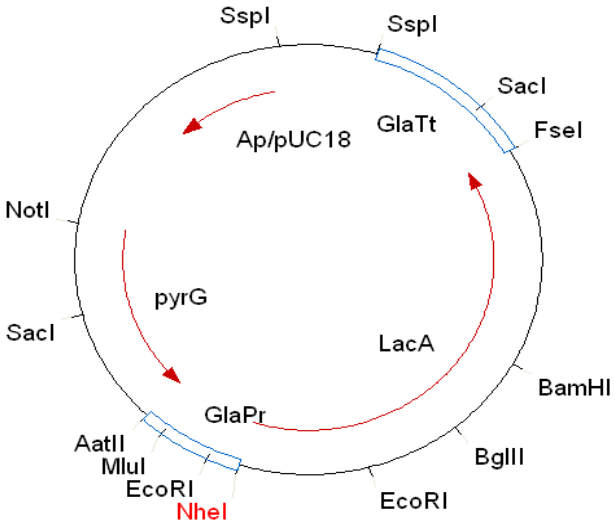
Genetically engineered *A. niger* strains were used to produce the second group of enzymes. These engineered strains harboured genes amplified by PCR from fungal genomic DNA or cDNA cloned into expression vectors ANIp5 or ANIp7G . Genes potentially encoding cellobiohydrolases were identified by searching publicly-available fungal genome sequences using the *T. reesei CbhI* and *Cbh2* sequences as queries (. Sydenham R and. Tambor H 2012 personal communication). In addition, some of the genes were cloned using cDNA libraries of two different species of fungi (https://fungalgenomics.concordia.ca).

Recombinant ANIp5 and ANIp7G plasmids (Figure 24) harbouring exoglucanases were transformed into *A. niger* strain N593 glaA::hisG (Storms R et al. 2005). Recombinant pGBFIN-GTW plasmids harbouring exoglucanases were transformed into *A. niger* strain CBS 513.88

(FGSC A1513) (Tambor JH et al. 2012). Individual transformants harbouring recombinant ANIp5 and ANIp7G plasmids were grown in Erlenmeyer flasks (500 ml) with 100 ml of MMJ medium (Master ER et al. 2008) with 15% glucose at 30 ºC and shaking (200 rpm) for 7 days. Individual transformants of CBS 513.88 transformed with pGBFIN- GTW harbouring fungal Cbhs were grown in 100 ml of STIPT liquid medium as described previously (Tambor JH et al. 2012). Recombinant Cbhs (Sthe437, TreeCbh1; TreeCbh2; Ledo11599; Asn3751; Ccin11396; and Afu6g11610) were kindly provided by Racha Cheikh-Ibrahim). The other recombinant Cbhs were expressed in strain A2P5 as described in (Sydenham R et al. 2014).

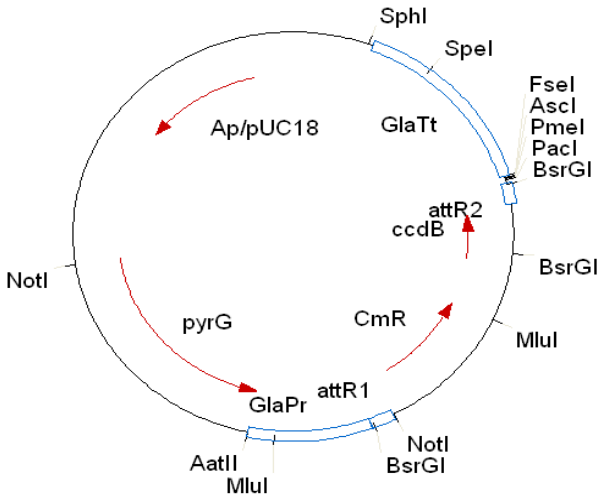
Culture medium, separated from cell biomass by centrifugation at 10,000 X g for 10 min, was used to screen for expression and for enzyme assays. The gene and protein sequences of the recombinant cellobiohydrolases can be found in (Appendix 1A). The primer sequences used to amplify cDNAs and gDNAs encoding the recombinant cellobiohydrolases are presented in Appendix 1B.

A. ANIp5



SspI	424
SacI	1152
FseI	1488
BamHI	2977
BglII	3531
EcoRI	4122
NheI	4929
EcoRI	5126
MluI	5444
AatII	5590
SacI	6377
NotI	6994
SspI	8817

B.ANIp7G



Name	Pos
SphI	405
SpeI	760
FseI	1488
AscI	1496
PmeI	1504
PacI	1512
BsrGI	1535
BsrGI	1937
MluI	2334
NotI	3100
BsrGI	3220
MluI	3750
AatII	3896
NotI	5300

Figure 24. Expression vectors ANIp5 (Storms R et al. 2005) and ANIp7G (Tambor JH et al. 2012) were used to express gDNA and cDNA clones encoding cellobiohydrolases.

Methods:

SDS-PAGE Gel Electrophoresis:

Denaturing SDS-PAGE analysis was performed using 12% polyacrylamide gels as described elsewhere (Laemmli UK 1970) .Protein samples were diluted 4:1 (v/v) into SDS loading dye buffer, incubated for 5 min in a boiling water bath at 95ºC and then kept on ice. Prior to loading onto the SDS-PAGE gels the protein samples were centrifuged at 10,000 X g for 1 min. Gels were run at 80 V in 1x Tris-Glycine buffer until the samples migrated through the stacking gel, then the voltage was maintained 130 V as the samples migrated through the resolving gel. Molecular weight protein standards (Amersham Low Molecular Weight Calibration Kit for SDS Electrophoresis) obtained from GE Healthcare were used to estimate the molecular weights of the expressed proteins. Proteins were visualized by staining gels with Coomassie Brilliant Blue R-250.

Sugar removal from culture filtrates:

Sugars present in the growth medium were removed prior to some activity assays. Sugars in culture filtrates were removed using ULTRAFREE-0.5 Centrifugal Filter Devices (10 K cut-off) from Millipore. This device has two parts, a concentrator tube which has a membrane and a receiving tube below it. Citrate buffer (10 mM pH 5.0) was used as the exchange buffer. Sample solutions of 250-500 µl were added to the concentrator, which was then centrifuged at 10,000 rpm for 30 min at 4 ºC. The contents of the receiving tube were discarded, 250-500 µl of exchange buffer were added, and the device was again centrifuged at 10,000 rpm for 30 min. The contents of the receiving tube were again discarded and buffer exchange was repeated 4 times. The volume of filtrate remaining in the concentrator was diluted with the exchange buffer to make the final volume equal to the original volume of the sample.

Estimation of protein concentration:

Protein concentrations were estimated using the 2-D Quant Kit (GE Healthcare life sciences) and the method provided by the supplier.

Chromogenic enzyme activity assays:

Chromogenic enzyme assays were used to screen culture filtrates for enzyme activity. Enzyme activities toward chromogenic substrates were generally assayed at 50 ºC , pH 5.0 for 60 min. Assays were performed with p-nitrophenol-cellobioside (pNPC) (0.17 mM) (SIGMA-N-5759 with the Bgl inhibitor D-L gluconolactone 33 µg/ml), sodium citrate (10 mM pH 5.0) and an appropriately diluted enzyme sample. Typically, reaction mixtures were performed in a final volume of 300 µl prepared by combining pNPC substrate (10 µl of 5 mM pNPC), sodium citrate buffer (15 µl of 0.2 M pH 5.0), H_20 (250 µl) and enzyme (25 µl). A standard curve was prepared using the same reaction conditions except that the substrate and enzyme samples were replaced with appropriately diluted pNP-only (35 µl in citrate buffer 10 mM pH 5.0). Reactions were carried out in 96-well microplates with sealing films to prevent evaporation. Reactions were stopped by transferring to ice and the addition of 50 µl of Na_2CO_3 (1 M) to ensure enough alkalinity to ionize *p*-nitrophenol (PNP) to fully develop the yellow color. Each reaction mixture (80 µl) was then transferred to a clean microtitre plate and the absorbance was measured at 410 nm using a BIO-TEC Power Wave HT microplate reader. Each enzyme activity measurement was carried out in triplicate. One unit of enzyme activity is defined as the amount of enzyme that liberated 1 µmol of p-nitrophenol in 1 min. Enzymes samples showing at least three times more activity than the negative control, a culture filtrate prepared from an *A. niger* transformant harbouring the expression vector without an insert, were selected for further characterization.

Experimental designs used to characterize enzymes:

There were two experimental designs employed to characterise enzymes based on their ability to hydrolyse cellulosic substrate PASC. All experiments were done in duplicate and triplicate assays were performed for each time point.

Design # 1: Enzyme supplementation at 24 hours.

The aim of experimental design #1 was to test whether the various recombinant enzymes could enhance PASC hydrolysis by Accellerase 1500. Accellerase 1500 (80 mg/ml) equals 9.7E-05 IU, where units of activity were measured using Accellerase 1500 (0.08 mg/ml), PASC (0.1%) and a 1 h hydrolysis at 37ºC.

Design 1 enzyme assays were carried out using citrate buffer (67 mM), PASC (0.25%, 1.5% or 2%) and various amounts of cellulase enzyme. Every 1,000 µl of reaction mixture consisted of 200 µl of appropriately diluted cellulases in 10 mM citrate buffer, 300 µl of citrate buffer (200 mM pH 5) and 500 µl of PASC (0.5%, 3% or 4% in 10 mM citrate buffer). For the enzyme background control, the reaction solution was prepared in the same manner substituting citrate buffer (200 mM pH 5.0) for the substrate. For the substrate background control substrate was prepared in the same manner but substituting an equal volume of buffer for enzyme. The control absorbances were subsequently subtracted from the absorbance reading of each reaction mixture.

The enzymes and Accellerase 1500 used were diluted using citrate buffer (10 mM pH 5). Enzyme assays were performed in round bottom tubes (10 ml) that could be closed using caps. Enzyme assays were incubated on a rotator at 80 rpm in a 37ºC incubator. The rotator was tilted at an angle but keeping the reaction mixture from entering the cap. After 24 hours aliquots (500

µl) were transferred to new reactions tubes that had been supplemented with 100 µl of test enzyme or Accellerase 1500 in citrate buffer (10 mM) and the reaction mixture was continued up to 72 hours. One unsupplemented reaction tube was also included. Aliquots were transferred to tubes (1.5 ml) at 0 h, 24 h, 48 h and 72 h, reactions stopped by placing the tubes on ice for 10 min, and stored at -20ºC until they were processed further using bicinchoninic acid assay (BCA) to measure the amount of reducing sugar ends in the reaction. The amount of PASC hydrolysis (reducing sugar ends in the reaction) was expressed as a percentage of the theoretical yield of reducing sugar ends expected if 100% of the PASC was converted to glucose monomers.

Calculations for Design 1 (1.5% PASC): For a 2,000 µl reaction volume using 1.5% PASC as the substrate, the amount of PASC added would be 30 mg, and therefore each of the 500 µl reactions would have 30 mg/ml x 500 µl /2000 µl or 7.5 mg PASC per reaction. Since PASC is composed of anhydro glucose units 7.5 mg of PASC would contain (7.5 mg anhydro glucose/162.2 grams per mole of anhydro glucose units or 4.6×10^{-5} moles of anhydroglucose), which if completely converted by enzymatic hydrolysis of glucose would yield 4.6×10^{-5} moles of glucose.

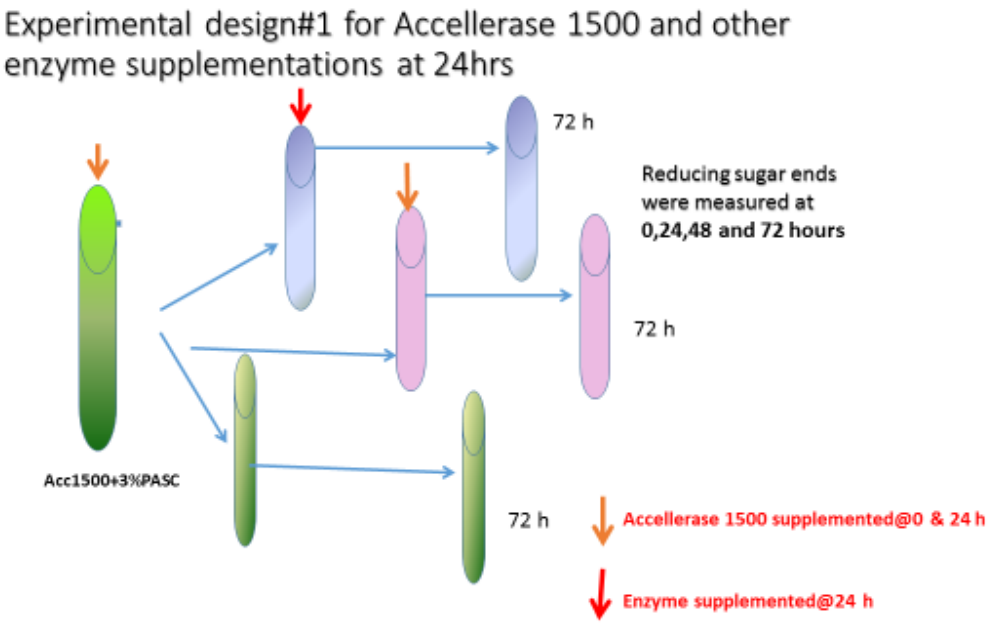


Figure 25. **Experimental design #1. Assessing how enzyme supplementation affects cellulose hydrolysis rates and yields by Accellerase 1500**. The large green tube represents the initial reaction mixture (PASC and Accellerase 1500 in 200 mM citrate buffer pH 5). The smaller tubes blue and pink represent aliquots of the original reaction mixture that have been supplemented with various test enzymes or Accelerase1500. Green tube represent unsupplemented sample. Samples were removed for reducing sugar assays at 0 h, 24 h, 48 h and 72 h.

Design # 2: Enzyme supplementation at 0 hours.

The aim of experimental design #2 was to test whether the various recombinant enzymes could enhance PASC hydrolysis by Accellerase 1500. Design 2 enzyme assays were carried out using citrate buffer (67 mM), PASC (2%) and cellulase enzyme. Every 500 µl of reaction mixture consisted of 100 µl of appropriately diluted cellulases in 10 mM citrate buffer, 150 µl of citrate buffer (200 mM pH 5) and 250 µl of PASC (4% in 10 mM citrate buffer). The unsupplemented reaction consisted of (Accellerase 1500 80 µl + 20 µl 10 mM citrate buffer) while the supplemented reactions consisted of Accellerase 1500 (100 µl) or Accellerase 1500 80

µl + 20 µl recombinant enzyme. For the enzyme background control, the reaction solution was prepared in the same manner substituting citrate buffer (200 mM pH 5.0) for the substrate. For the substrate background control substrate was prepared in the same manner but substituting an equal volume of buffer for enzyme. The control absorbances were subsequently subtracted from the absorbance reading of each reaction mixture.

The enzymes and Accellerase 1500 used were diluted using citrate buffer (10 mM pH 5). Enzyme assays were performed in round bottom tubes (1 ml) that could be closed using caps. These tubes were inserted in larger 10 ml tubes that could fit in the rotator. Enzyme assays were incubated on a rotator at 80 rpm in a 37ºC incubator. The rotator was tilted at an angle but keeping the reaction mixture from entering the cap. Reaction tubes were removed from the rotator after 0 h, 24 h, 48 h and 72 h and the reaction was stopped by placing the tubes on ice for 10 min followed by freezing -20ºC until they were processed further using bicinchoninic acid assay (BCA) to measure the amount of reducing sugar ends in the reaction. The amount of PASC hydrolysis (reducing sugar ends in the reaction) was expressed as a percentage of the theoretical yield of reducing sugar ends expected if 100% of the PASC was converted to glucose monomers. Calculations for Design 2 (2% PASC): For a 500 µl reaction volume using 2% PASC as the substrate, the amount of PASC added would be 10 mg per reaction. Since PASC is composed of anhydro glucose units 10 mg of PASC would contain (10 mg anhydro glucose/162.2 grams per mole of anhydro glucose units or 6.17×10^{-5} moles of anhydroglucose), which if completely converted by enzymatic hydrolysis to glucose would yield 6.17×10^{-5} moles of glucose.

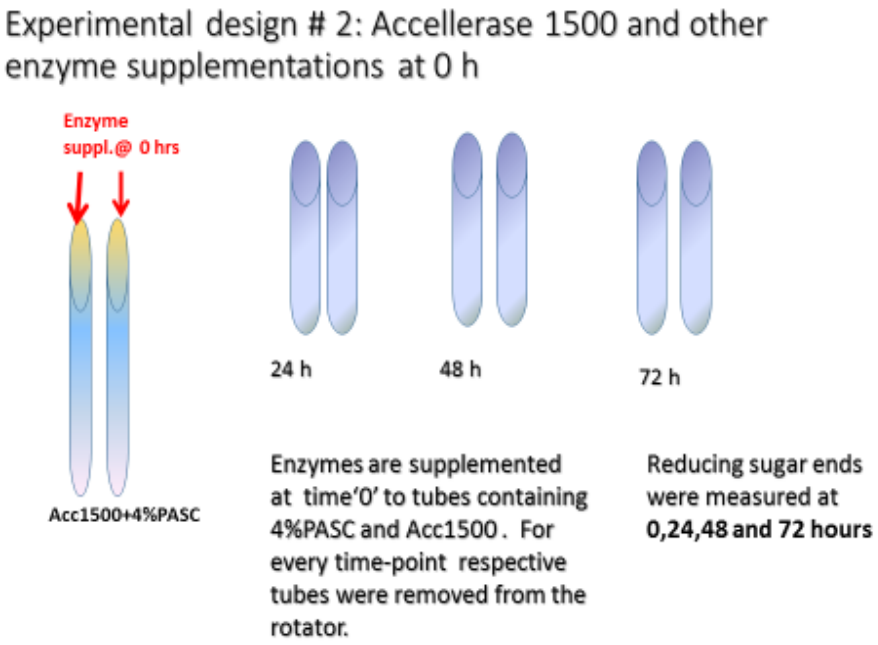


Figure 26. Experimental design #2 used to characterise enzymes based on their capacity to hydrolyse cellulosic substrate PASC. The large tube represents initial reaction mixture with PASC, 200 mM Citrate pH 5, Accellerase1500 and enzymes to be screened. The smaller tubes (blue) represent reaction mixture supplemented with Accellerase1500 and enzymes to be tested. Samples were removed for reducing sugar assays at 0 h, 24 h, 48 h and 72 h respectively.

Non-chromogenic enzyme activity assays:

When using PASC as the substrate glycosidic bond cleavage was followed using the bicinchoninic acid assay (BCA) adapted for 96-well microplates (Grishutin S G et al. 2004). A 10 µl sample of a reaction mixture or a glucose standard was transferred into the well of a PCR plate kept on ice, containing 100 µl of BCA working solution reagents (A+B) (Doner, L.W *et al*., 1992) (Zorov, I.N. *et al*., 1997) and 90 µl of distilled water. Assay plates were sealed to prevent evaporation, incubated at 80ºC for 30 min cooled on ice for 5 min, reaction mixture (80 µl) was transferred to a clean microplate plate, and the absorbance measured at 562 nm using a BIO TEK Power Wave HT microplate reader (it should be noted that to be in the linear range of the BCA

assay the absorbance readings could not be greater than 0.4). A calibration curve was generated using glucose as the standard. The amount of reducing sugar ends in the reaction supernatant and whole reaction mixture were expressed as a percentage of the theoretical yield of reducing sugar ends obtained if 100% of the PASC was converted to glucose monomers.

Bioinformatics analysis of cellobiohydrolase sequences:

Protein sequences of the six Cbhs and the primers used for PCR amplification are presented in Appendix 1A. The sequences were aligned using Clustal W2 and all subsequent test were done using "complete deletion" where gaps and missing data is ignored. Determination of conserved protein domains was performed using NCBI (CDART) Conserved Domain Architecture Retrieval Tool, (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi?cmd=rpsand), and the NCBI Conserved Domain Database (CDD), (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The protein sequences were used to construct a Phylogenetic tree using Clustal W2.

End product analysis using ion chromatography:

End product analysis was performed on 72-hour incubations of reactions containing the enzymes listed in Table 2 as well as Afu6g11610 and (3% PASC) as a substrate. Sample and standard analyses were carried out using a Dionex ICS-5000 coupled to an AS-AP autosampler. All analyses were conducted at 30°C with isocratic elution using a CarboPac PA20 (Dionex) guard (3 mm x 30 mm) and analytical (3 mm x 150 mm) columns connected in series. Elution conditions were 100 mM NaOH for 10 minutes at a flow rate of 0.5 ml/min. Reaction products were appropriately diluted and aliquots (10 µl) were injected into the Dionex system. Glucose and cellobiose standards were prepared in a concentrations range of 0.5 mM to 0.00024 mM. The

results obtained showed two peaks; the first at 4 minutes corresponded to glucose, and the second at 10.8 minutes corresponded to cellobiose. Peak height and peak area values of the standards obtained from the system were plotted and a linear regression was obtained. Millimolar concentrations of glucose and cellobiose (y) in the samples were calculated using the recorded values (or signal values) of peak height / peak area (x) using equation ($y= x^2 + x$). Dilution factor was included in the equation.

Results

1. *Benchmarking PASC (0.25 %) hydrolysis with Accellerase 1500*

I began my cellulase research by performing three sets of experiments aimed at benchmarking the hydrolysis of PASC by Accellerase 1500. The objective of the first set of experiments was to determine whether a significant portion of the glycosidic bonds hydrolyzed by Accellerase 1500 were associated with insoluble cellulose polymers (i.e. polymers with a DP greater than about 8 residues). The objective of the second set of experiments was to identify a ratio of Accellerase 1500 to PASC (enzyme loading/substrate concentration) that resulted in cellulose hydrolysis plateauing when about 50% of the potential glycosidic bonds had been cleaved (i.e. 50% of the potential reducing sugar ends generated). Once a ratio of Accellerase 1500 to PASC that resulted in about 50% hydrolysis was identified, I would use it to screen for enzymes that could enhance the rate and/or yield of reducing sugar equivalents. The objective of the third set of experiments was to use the assay conditions identified in the second set of experiments to identify the Accellerase 1500 component(s) that limited hydrolysis rates and/or sugar yields.

1.i Total versus soluble reducing sugar equivalents

The results of the first set of experiments are presented in Figure 27. To follow the production of soluble reducing sugar equivalents in the PASC hydrolysis reactions aliquots (100 μl) taken (note: hydrolysis reactions were vigorously vortexed prior to sampling) at various times were transferred to a microfuge tube and subjected to centrifugation at 10,000 g for 1 minute. Following centrifugation a supernatant sample (10 μl) was used to determine the amount of soluble reducing sugar equivalents generated using the BCA assay. To determine the total

amount of reducing sugar equivalents (soluble and insoluble) present in PASC hydrolysis reactions, aliquots (100 µl) taken from the reaction mixture immediately after vigorous mixing 10µl were assayed for reducing sugar equivalents using the BCA assay.

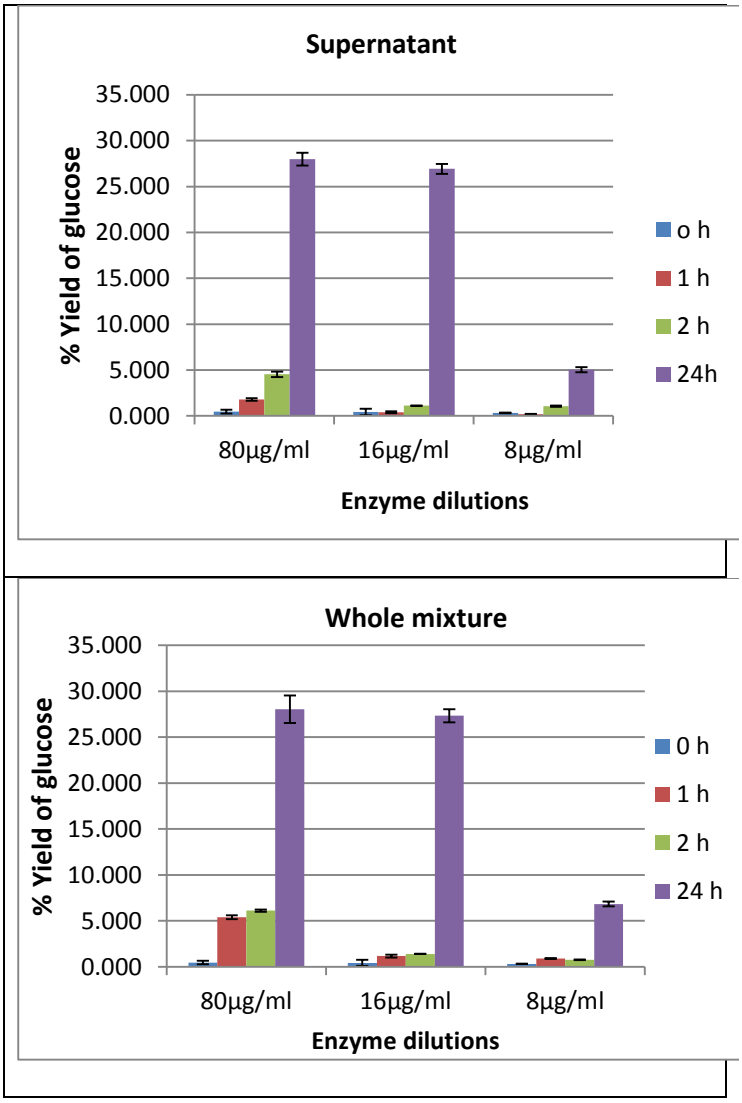


Figure 27: Hydrolysis of PASC (0.25%) using various concentrations of Accellerase 1500 (16, 3.2 or 1.6 µg/ml). Hydrolysis reactions (2 ml), consisted of 1,000 µl PASC (0.5%), 600 µl of citrate buffer (200 mM pH 5) and 400 µl Accellerase 1500 (80, 16 or 8 µg/ml). The yield of soluble reducing sugar equivalents (top panel) and total reducing sugar equivalents (lower panel) are presented as a percentage of the reducing sugar equivalents generated relative to the amount expected if 100% of the PASC (0.25%) was converted to glucose monomers. Complete hydrolysis of 2 ml of 0.25% PASC would yield 3.09 x 10^{-5} moles of glucose).

The results in Figure 27 (summarized in Table 1) show that the amount of reducing sugar equivalents present in the supernatant after 2 h of hydrolysis with 80 μg/ml, 16 μg/ml or 8 μg/ml of Accellerase 1500, was significantly less than the amount of reducing sugar equivalents present in the whole reaction. Furthermore, reducing sugar equivalents present in the supernatant and whole reaction were not significantly different after 24 h of hydrolysis with Accellerase 1500 (16 μg/ml and 3.2 μg/ml) but were significantly different when hydrolysis was done with Accellerase at 1.6 μg/ml. Since the amount of reducing sugar ends associated with the insoluble fraction was significantly greater after a 2 h hydrolysis and after 24 h when the using the lowest concentration of Accellerase 1500. Based on these results, I concluded that the hydrolysis of PASC should be assayed by following the production of total (soluble and insoluble) reducing sugar equivalents.

Table 1: Summary of the results presented in Figure 27.

Whole Reaction	80µg/ml	16µg/ml	8µg/ml	Supernatant	80µg/ml	16µg/ml	8µg/ml
Time in hours	Percent yield (%)			Time in hours	Percent yield (%)		
0	0.468 ±0.208	0.432 ±0.338	0.324 ±0.029	0	0.468 ±0.208	0.432 ±0.338	0.324 ±0.029
1	5.396 ±0.206	1.173 ±0.158	0.899 ±0.036	1	1.799 ±0.136	0.396 ±0.115	0.18 ±0.038
2	6.215 ±0.108	1.603 ±0.022	1.072 ±0.065	2	4.532 ±0.310	1.115 ±0.007	0.755 ±0.029
24	28.05 ±1.494	27.338 ±0.718	6.835 ±0.266	24	27.993 ±0.690	26.932 ±0.539	5.036 ±0.267

The amount of reducing sugar equivalents present in the whole mixture (left side) versus the supernatant (right side) when PASC (0.25 %) hydrolysis was performed using 16, 3.2 or 1.6 µg/ml of Accellerase 1500.

I.ii Developing an assay where the amount of Accellerase 1500 limits the rate and amount of PASC hydrolysis

The second set of experiments (Figure 28 and 29) followed the production of total reducing sugar equivalents during 110 h hydrolysis reactions using PASC (0.25%) as the substrate. The hydrolysis reactions were initiated using Accellerase 1500 at a concentration of 16 µg/ml and some reactions were supplemented with additional Accellerase 1500 at 72 h (Figure 28), 24 hours (Figure 29) and 24 h and 48 h (Figure 29). About 60% of the potential reducing sugar equivalents were released when hydrolysis was performed with 16 µg/ml of Accellerase 1500 for

110 h (Figure 28 and 29). Supplementing with additional Accellerase at 24 h or 72 h or 24 h and 48 hours increased the amount of hydrolysis obtained from about 60% to about 80%, 90% and 96 %, respectively. These results show that the hydrolysis of 0.25 % PASC using 16 µg/ml Accellerase 1500 plateaued at about 60% of the theoretical yield at about 72 h and that supplementation with additional enzyme increased both the rate and yield of hydrolysis. Furthermore, the results suggest that hydrolysis plateaus at about 60% of the theoretical yield when using 16 µg/ml of Accellerase 1500, because one or more than one of the cellulase system components have been inactivated.

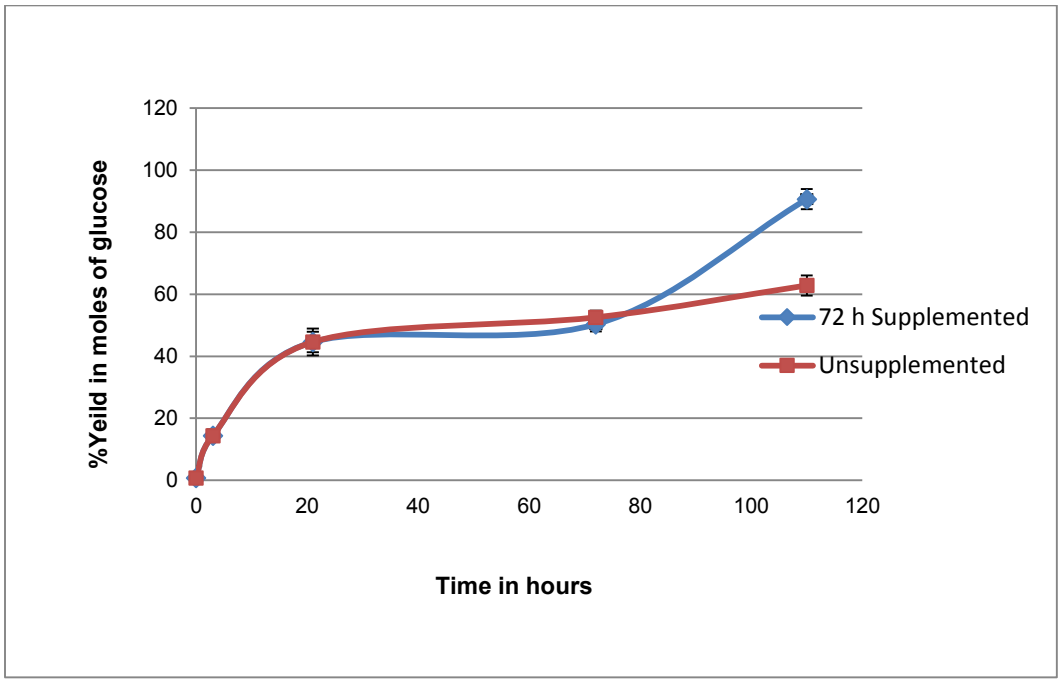


Figure 28: Hydrolysis of PASC (0.25%) with Accellerase 1500 (16 µg/ml added at t = 0 h) and Accellerase 1500 (16 µg/ml added at t = 0 h and the amount of Accellerase 1500 doubled by supplementation at t = 72 h). Red curve; a 1 ml hydrolysis reaction consisting of 500 µl of PASC (0.5 %), 300 µl of citrate buffer (200 mM pH 5) and 200 µl Accellerase 1500 (80 µg/ml) added at 0 h. Blue curve; a 1 ml hydrolysis reaction consisting of 500 µl of PASC (0.5 %), 300 µl of citrate buffer (200 mM pH 5) and 200 µl Accellerase 1500 (80 µg/ml) added at 0 h and Accellerase 1500 (200 µl of 80 µg/ml added at 72 h). The yield of reducing sugar equivalents was determined as described in Figure 27.

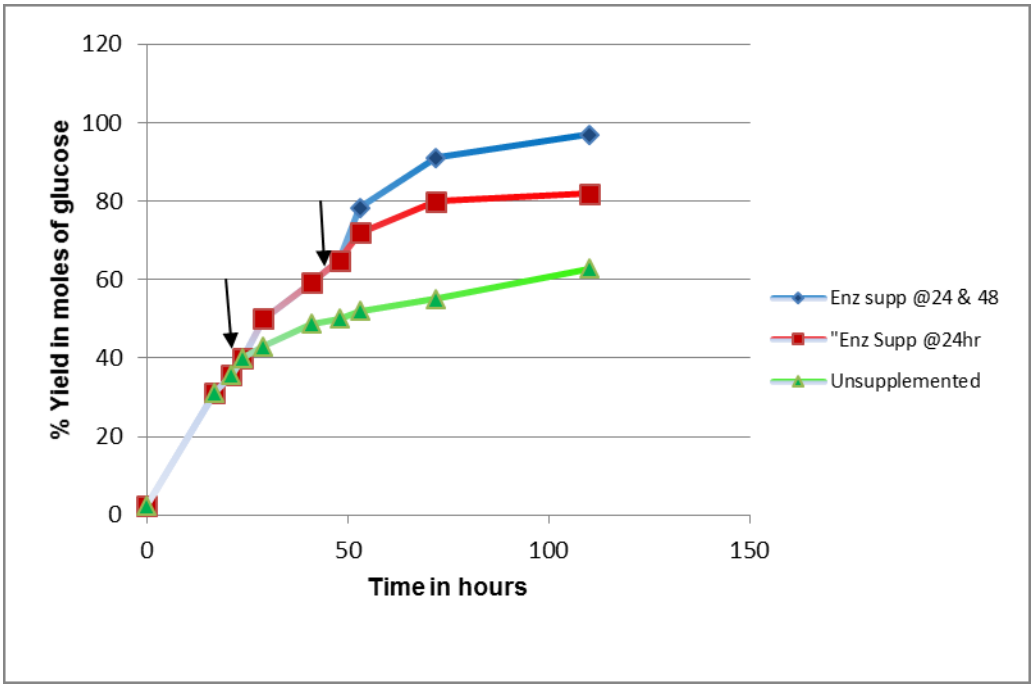


Figure 29: The effect of stepwise supplementation at 24 h and 48 h on PASC (0.25%) hydrolysis by Accellerase 1500 (16 µg/ml). At t= 0 h a 1.5 ml hydrolysis reaction was prepared, consisting of 750 µl PASC (0.5%), 450 µl of citrate buffer (200 mM pH 5) and 300 µl Accellerase 1500 (80 µg/ml)) (light blue curve). At 24 h the reaction was split into two reaction tubes, one with 0.5 ml (green curve) and one with 1.0 ml (purple curve). The 1.0 ml reaction was spiked with additional Accellerase 1500 (16 µg). At 48 h two 0.5 ml aliquots were transferred from the 1 ml reaction tube into 2 new tubes, one was spiked with Accellerase 1500 (8 µg) (dark blue curve) the other tube was not subjected to further supplementation (red curve). Samples were taken at 0, 17, 21, 24, 29, 41, 48, 53, 72 and 110 h. The yield of reducing sugar ends present in each sample was determined as described in Figure 27. Vertical arrows indicate the time-points at which the reactions were supplemented with Accellerase 1500.

I.iii Developing a PASC hydrolysis assay using 1.5% PASC and a rate limiting concentration of Accellerase 1500

The third set of experiments (Figure 30, 31, 32, 33, 34, 35, 36 37 and 38) was performed to determine which if any of the major glycosyl hydrolases present in Accellerase 1500 was the component(s) whose inactivation caused PASC hydrolysis to plateau when about 60% of the glycosidic bonds had been hydrolyzed (Fig. 28 and 29). Using my earlier results that identified

substrate to enzyme ratios that resulted in limited hydrolysis yields and the fact that lignocellulose saccharification for the production of bioethanol require higher cellulose loading (Banerjee G et al. 2011), I wanted to devise a hydrolysis assay that could be used to identify cellulases that could enhance the hydrolysis of PASC by Accellerase 1500 using 1.5% PASC rather that 0.25 % PASC. Since 1.5% PASC is 6 times more concentrated than the 0.25% PASC I used previously, I followed the hydrolysis of 1.5% PASC using different concentrations of Accellerase 1500 with the aim of finding a concentration of Accellerase 1500 that would yield about 50 % hydrolysis after 72 hours. The results of this experiment (Figure 30) showed that Accellerase 1500 at 26.4 µg/ml was able to hydrolyze about 50% of the glycosyl bonds in a 1.5% PASC reaction.

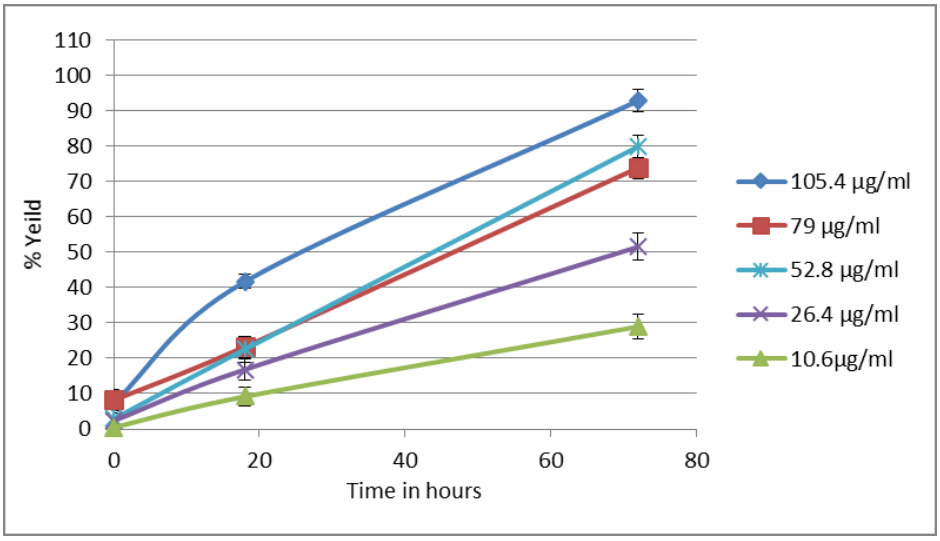


Figure 30: The effect on PASC (1.5%) hydrolysis of varying the concentrations of Accellerase 1500 (10.6, 26.4, 52.8, 79, and 105.4 µg/ml). Hydrolysis reactions (1 ml) consisted of 500 µl of PASC (3%), 300 µl of citrate buffer (200 mM pH 5), and 200 µl of Accellerase 1500 (53, 132, 264, 395 and 527 µg/ml). The percentage yield of reducing sugar ends present in each sample was determined as described in Figure 27 using aliquots taken at 0 h, 18 h and 72 h.

I.iv Determining whether the cellulose hydrolysis assay developed in I.iii could be used to identify the Accellerase 1500 component that was limiting the rate and/or yield of PASC hydrolysis.

To identify the Accellerase 1500 component or components that were limiting hydrolysis rates and/or yields when the reactions contained 1.5% PASC and 26.4 μg/ml Accellerase 1500, I supplemented hydrolysis reactions at 24 h individually with Accellerase 1500, Cel7A/Cbh1, Cel6B/Cbh2, Cel6A/Eg1, Cel5A/Eg2 or Bgl1 and various binary combinations of these enzymes. The results from these experiments are presented in (Figures 31, 32, 33, 34, 35, 36, 37 and 38) and summarized in Table 2.

Cbh1 consistently enhanced the yield of hydrolysis products by about 14%, Cbh2 consistently enhanced the yield of hydrolysis products by about 33% and Bgl consistently enhanced the yield of hydrolysis products by about 8%. In contrast, Eg1 supplementation did not affect yields and Eg2 supplementation inhibited the yield by about 7%.

Hydrolysis reactions supplemented with binary combinations of Cbh1, Cbh2, Eg1, Eg2 and Bgl showed that the binary combinations Cbh2 and Cbh1 (Figure 34), Cbh2 and Bgl (Figure 32), Cbh1 and Eg1 or CbH1 and Eg2 (Figure 35), Cbh2 and Eg1 or Cbh2 and Eg2 (Figure 35) and Bgl and Eg1 or Bgl and Eg2 (Figure 37) did not enhance hydrolysis significantly more than the amount expected if supplementation had been with only the component that enhanced hydrolysis the most (Figure 31 and 38). Interestingly, one binary combination (Cbh1 and Bgl) (Figure 33) enhanced hydrolysis rates by about 38%, which is significantly more than the 8% and 14% increase obtained when individually supplementing with Bgl (Figure 31, 32, 33 and 37)

or Cbh1 (Figure 31, 33, and 35) and the increase expected (about 7%) if their combined effect on hydrolysis was additive. This experiment was repeated four times.

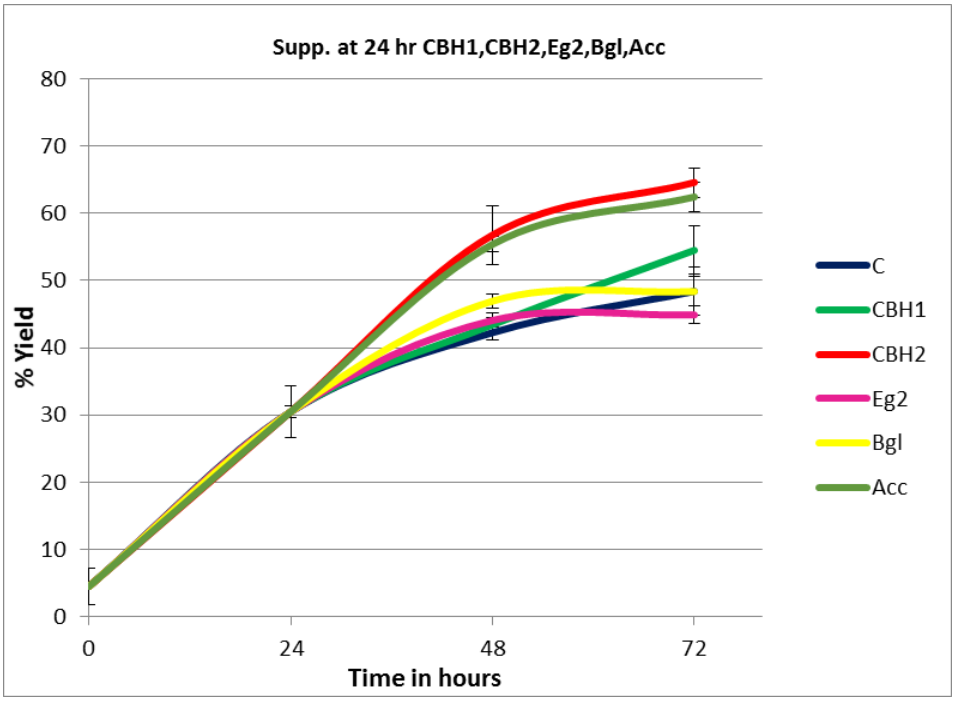


Figure 31: PASC (1.5%) hydrolysis by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with purified cellulase system enzymes. A 3 ml PASC hydrolysis reaction (1500 µl PASC (3%), 900 µl of citrate buffer (200 mM pH 5), and 600 µl of Accellerase 1500 (132 µg/ml)) was divided into six 0.5 ml reactions at 24 h and five of the tubes were supplemented with 100 µl of Accellerase 1500, Cbh1, Cbh2, Eg2 or Bgl (132 µg/ml). The 6th tube (C) was not supplemented. The yield, percentage reducing sugar equivalents generated relative to the amount expected if the PASC was completely hydrolyzed, was determined as described in Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

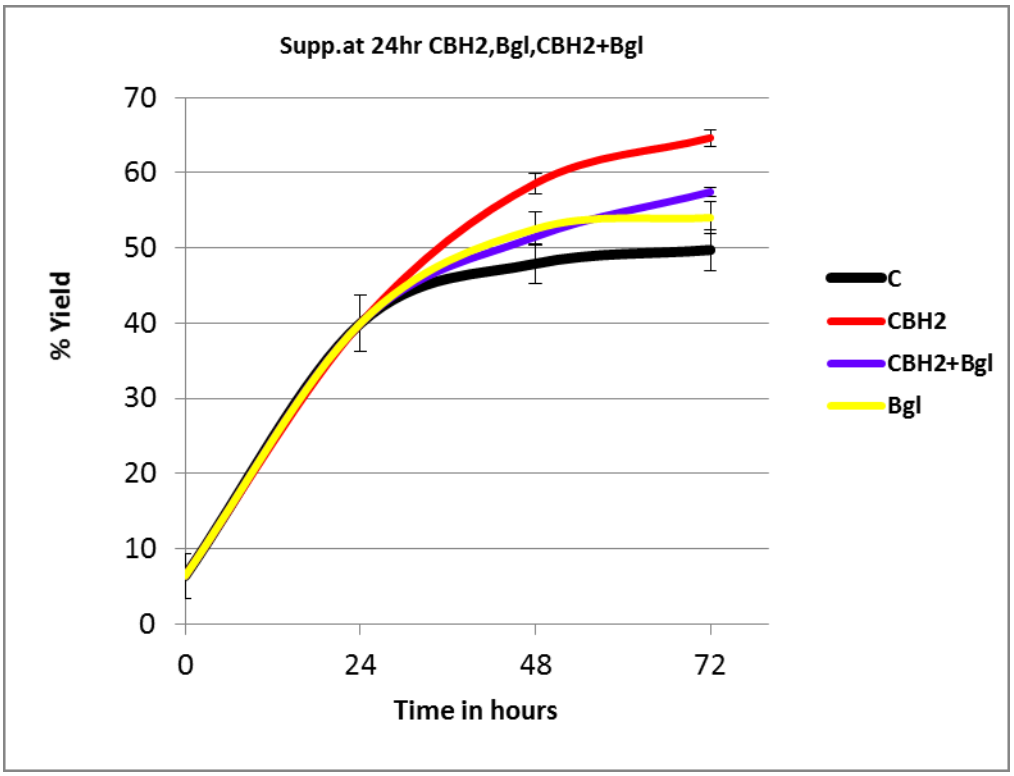


Figure 32: Hydrolysis of PASC (1.5%) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with Cbh2, Bgl or Cbh2 + Bgl. A 2 ml hydrolysis reaction (1000 µl of PASC (3 %), 600 µl of citrate buffer (200 mM pH 5), and 400 µl of Accellerase 1500 (132 µg/ml)) was divided into four 0.5 ml reactions at 24 hours. One tube (C) was not supplemented whereas 3 of the four tubes were supplemented with 100 µl of either Bgl (132 µg/ml), Cbh2 (132 µg/ml), or 50 µl of Bgl (66 µg/ml) and 50 µl Cbh2 (66 µg/ml). The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

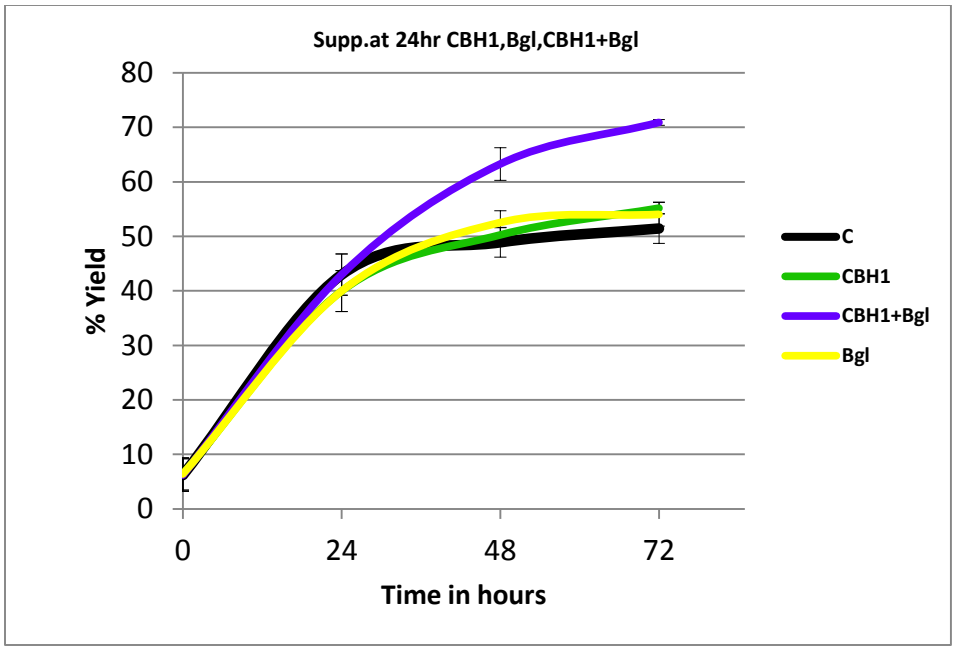


Figure 33: Hydrolysis of PASC (1.5%) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with Cbh1, Bgl or Cbh1 and Bgl. A 2 ml hydrolysis reaction (1000 µl of PASC (3%), 600 µl of citrate buffer, and 400 µl of Accellerase 1500 (132 µg/ml)) was divided into four 0.5 ml reactions at 24 hours. One tube was supplemented with (100 µl of Bgl (132 µg/ml), one tube with 100 µl Cbh1 (132/ml µg) and one tube with 50 µl of Bgl1 and 50 µl of Cbh1 (66 µg/ml of each enzyme). The 4th tube (C) served as the no supplementation control. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

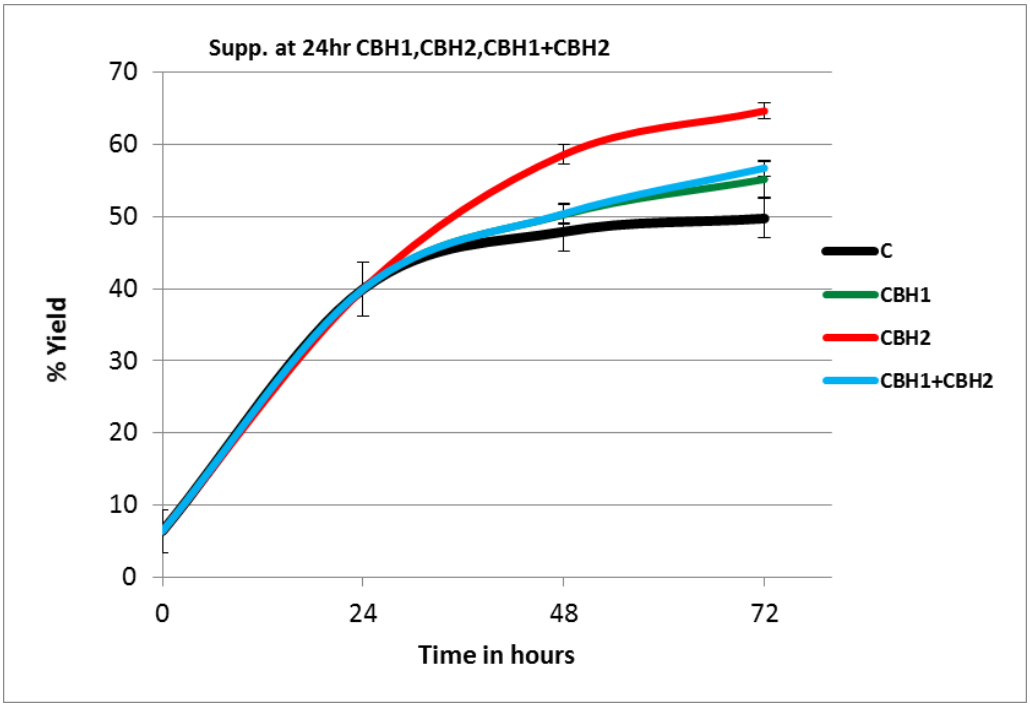


Figure 34: Hydrolysis of PASC (1.5%) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with Cbh1, Cbh2 or Cbh1 and Cbh2. The reactions were performed as described for Figures 32 and 33. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

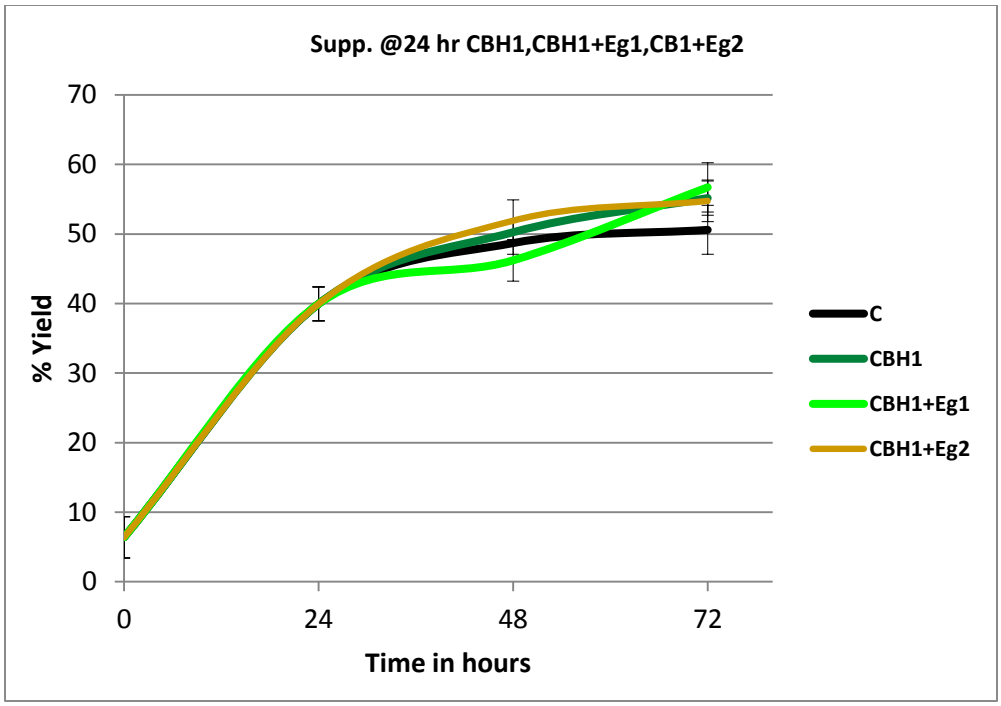


Figure 35: Hydrolysis of PASC (1.5%) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with Cbh1, Cbh1 and Eg1 or Cbh1 and Eg2. The reactions were performed as described for Figures 32 and 33. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

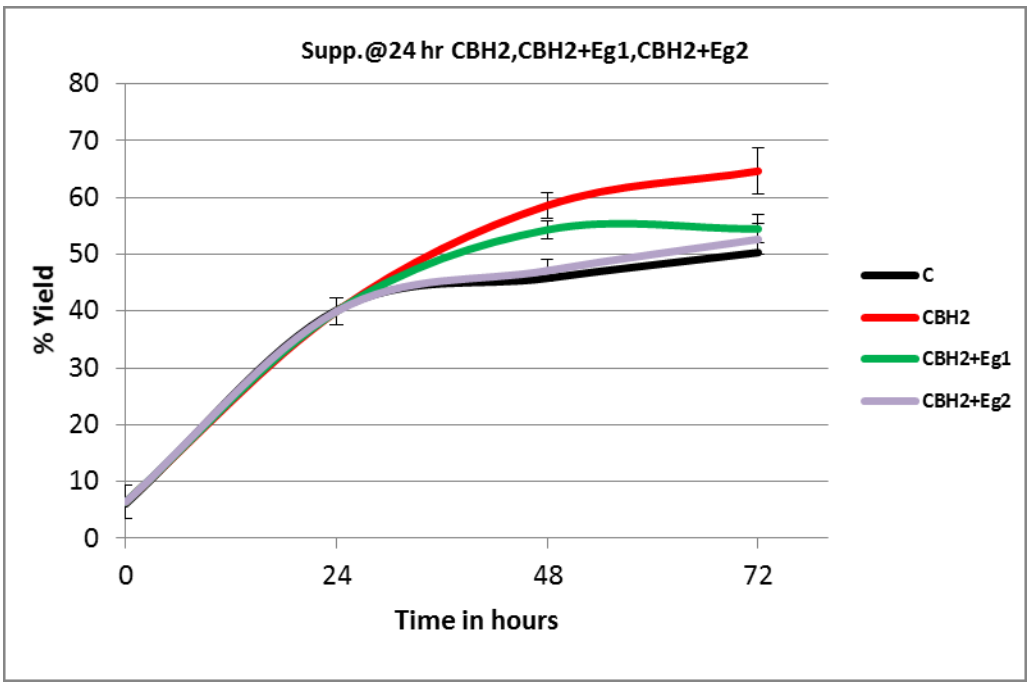


Figure 36: Hydrolysis of PASC (1.5 %) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with a binary combinations, Cbh2 + Eg1 and Cbh2 + Eg2. The reactions were performed as described for Figures 32 and 33. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

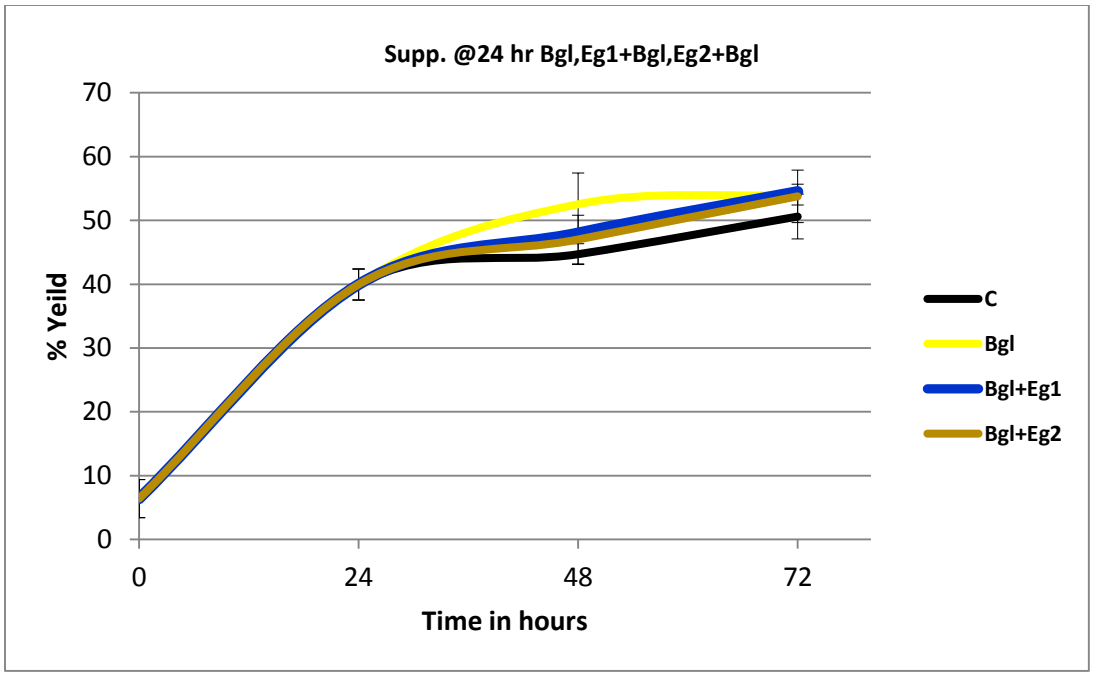


Figure 37: Hydrolysis of PASC (1.5%) by Accellerase 1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with Bgl, Bgl and Eg1 and Bgl and Eg2. The reactions were performed as described for Figures 32 and 33. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

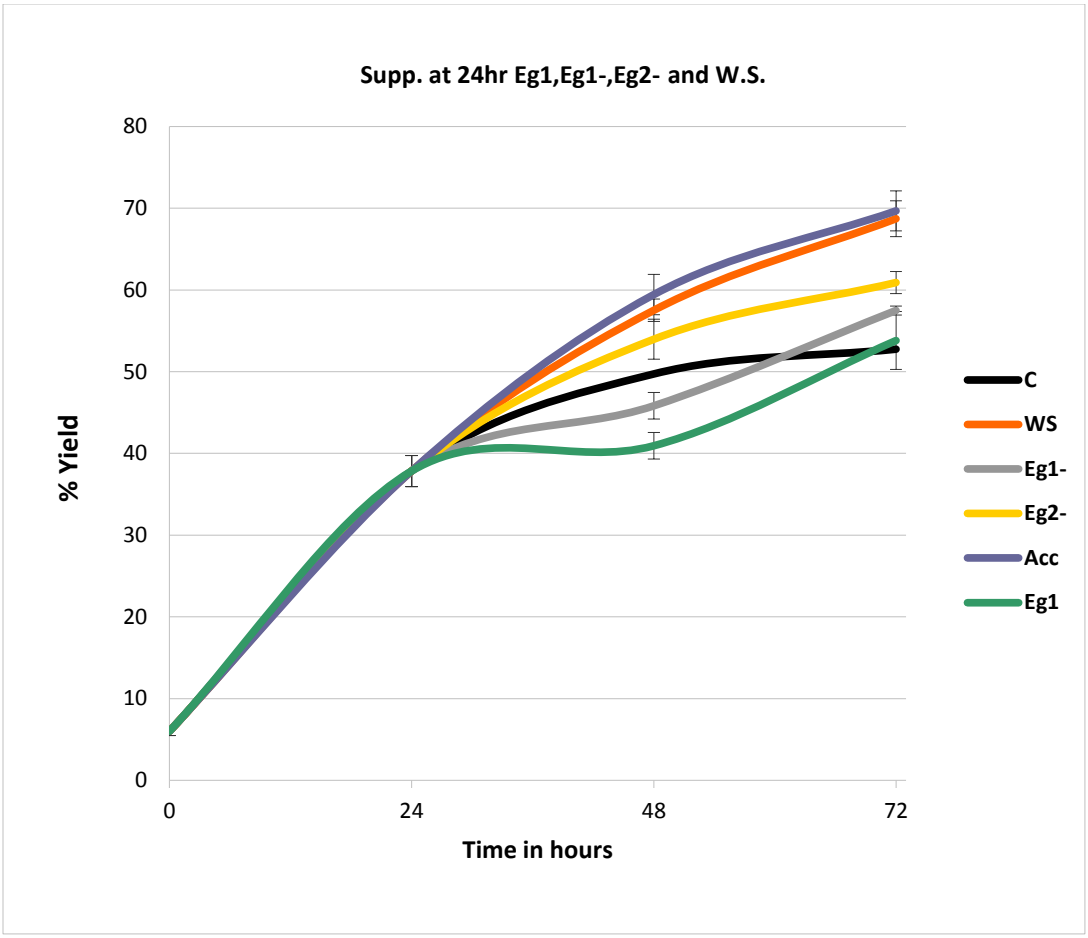


Figure 38: The effect on PASC (1.5%) hydrolysis by Accellerase1500 (26.4 µg/ml), and Accellerase 1500 (26.4 µg/ml) supplemented with purified cellulase system enzymes with Accellerase 1500, Iogen *T. reesei* cellulase system (WS), WS without Eg2, WS without Eg1 or Eg1,The reactions were performed as described for Figure 31. The yield of reducing sugar equivalents was determined as described for Figure 27 using aliquots taken at 0 h, 24 h, 48 h and 72 h.

Table 2: Summary of the effect on PASC hydrolysis at 72 h of supplementation at 24 h with additional Accellerase 1500, WS, WS-Eg1, WS-EG2, WS-Cbh1, WS-Cbh2, Eg1, Eg2, Cbh1, Cbh2, Bgl and various binary combinations of the major cellulases system components.

	Unsupplemented Accellerase1500 (C)[1]	Supplementing component(s) 72h[2]	Percentage (%) hydrolysis of PASC to glucose equivalents[1]
1	48.4	Cbh1	55.5 /14.4%
2	48.4	Cbh2	64.6 /33.5%
3	48.4	Eg1	53.8/ 11.2%
4	48.4	Eg2	44.9 /-7.2%
5	48.4	Acc1500	62.5/ 29.1%
6	49.7	Bgl	54.0 /8.7%
7	50.6	Cbh1+Eg1	56.7 /12.1%
8	50.6	Cbh1+Eg2	56.8/ 12.3%
9	51.3	Cbh1+Bgl	70.9/ 38.2%
10	50.3	Cbh2+Eg1	54.5 /8.3%
11	50.3	Cbh2+Eg2	52.6 /4.6%
12	49.7	CB2+Bgl	57.45/ 15.6%
13	49.7	Eg1+Bgl	54.5/ 9.7%
14	49.7	Eg2+Bgl	53.8 /8.2%
15	49.7	Cbh1+Cbh2	56.7/14%
16	52.8	W.S.	68.7/30.2
17	52.8	W.S Eg1^{-}	57.5/8.9
18	52.8	W.S. Eg2^{-}	60.9/15.3
19	49.4	VTO	48.0/ -2.8%

[1] Yield of glucose equivalents after a 72 h hydrolysis using unsupplemented Accellerase 1500 (26.4 µg/ml). .

[2] Accellerase 1500 (26.4 µg/ml) hydrolysis reactions supplemented with the indicated component (26.4 µg/ml) or binary components (13.2 µg/ml for each component).

[3] The first of the two percentage values indicates the yield of reducing sugar equivalents obtained after a 72 h hydrolysis performed using Accellerase 1500 and the supplement(s) indicated; whereas, the second percentage value indicates the amount of enhancement obtained with the indicated supplementation.

II.i. Activity and expression screening of novel cellulases

Racha Cheikh-Ibrahim a previous graduate student screened a library of Cbh1 and Cbh2 enzymes and identified 7 enzymes (Sthe437, TreeCbh1; TreeCbh2; Ledo11599; Asn3751; Ccin11396; and Afu6g11610) that could be functionally expressed by *A. niger* (Racha Cheikh-Ibrahim MSc Thesis, Expression and Characterization of Diverse Cellobiohydrolases Encoded in Fungal Genomes 2009 Concordia University). I wanted to determine whether these enzymes could enhance the hydrolysis of cellulose by the commercial cellulase system Acellerase1500. I began by preparing culture filtrates containing each of these cellulases expressed from ANIp7G that had been transformed into *A. niger* expression strain PY11.

I examined the level of expression obtained for each of the seven genes using SDS-PAGE analysis (Figure 39). I also used the 2-D-Quant kit to determine the level of total secreted protein in each culture filtrate (Table 3). I noticed a high level of background protein in the filtrates prepared using *A. niger* strain PY11 as the expression host.

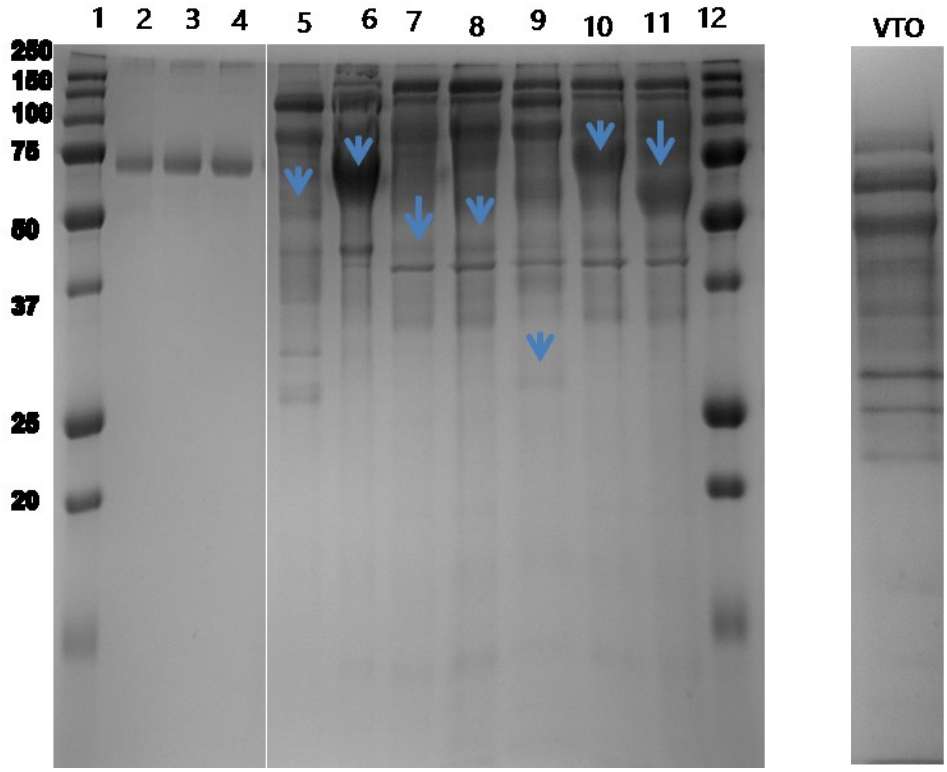


Figure 39: Protein expression levels determined using SDS-PAGE (12%) analysis. Lanes 1 and 12, molecular weight standards with molecular sizes indicated on the left; lanes 2, 3 and 4, BSA standards loaded with 5 µg, 7.5 µg and 10 µg, respectively; lanes 5 to 11, culture filtrates (10 µl) from PY11 transformants expressing Asn3751#1, Afug11610#28, Sthe437#1, Ledo11599#7, Ccin11396#34, TreeCbh1#11, and TreeCbh2#25; lane VTO, culture filtrate (10 µl) produced using PY11 transformed with expression vector ANIp7G without an insert. Arrows indicate putative recombinant enzyme bands.

II.ii. Expression of enzymes using new expression system A2P5

Due to the relatively high levels of background protein present in the PY11 culture filtrates I also attempted to express the 7 previously identified cellulases using *A. niger* strain A2P5 and plasmid ANIp7MS as the expression vector. This strain has been engineered to produce much lower levels of contaminating native secreted proteins. Only three of the above seven genes, Afug11610, Asn3751 and TreeCBHI, were expressed at levels that enabled detection of a protein band on SDS-PAGE (Figure 40). Furthermore only two of the above

seven genes, Afug11610 and Sthe437, expressed secreted cellobiohydrolase with detectable activity when activity assays were performed using pNPC as the substrate (Table 3). Culture filtrates of these two enzymes, prepared using A2P5 transformants, were used for further hydrolysis enhancement studies. The total protein expression levels, target Cbh protein levels and cellulase activity levels obtained with culture filtrates of the PY11 and A2P5 transformants harbouring Sthe437, TreeCBHI; TreeCBHII; Ledo11599; Asn3751; Ccin11396; and Afu6g11610 are presented in Figures 39, Figure 40 and Table 3. As expected significant bands corresponding to the expected Cbhs were observed in all 7 PY11 culture filtrates (Figure 39). A dark protein band was obtained with the Afug11610#28 filtrate and light protein bands were observed with the Asn3751 and TreeCbh1 culture filtrates obtained when using strain A2P5 as the expression host. However, recombinant protein bands were not observed in the culture filtrates obtained with A2P5 transformants harbouring Sthe437#1, Ccin11396#34, Ledo11599#7 and TreeCbh2#25. As compared to PY11 very low levels of background protein were observed in the culture filtrates produced with A2P5 (Compare Figures 39 and 40 and the summary data in Table 3).

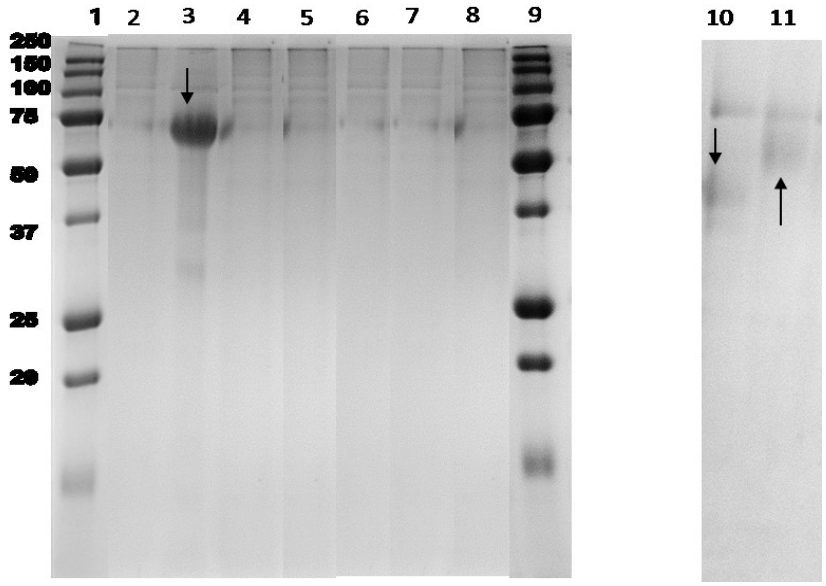


Figure 40: SDS-PAGE analysis of culture filtrates. Lanes 1 and 9, molecular weight standards with sizes in kDa indicated on the left; lanes 2 and 8, culture filtrates (10 µl) produced by A2P5 transformed with ANIp7MS without an insert, lanes, 3 to 7, and 10 and 11 culture filtrate (10 µl) produced by A2P5 transformed with Afug11610#28, Sthe437#1, Ccin11396#34, Ledo11599#7, TreeCbh2#25, Asn3751#1 and TreeCbh1#11. Arrows indicate putative recombinant protein bands.

Although light bands of Asn3751 and TreeCbh1 were observed (Figure 40), there was no enzyme activity obtained using the substrate pNPC to assay cellobiohydrolase activity. Hence these enzymes were not considered for further characterization. A strong protein band was observed, see lane 3 of Figure 40, with the culture filtrate for Afug11610. Although a protein band was not observed for Sthe437 (lane 3 of the same figure), on account of its high activity as shown using the pNPC assay, Sthe437 was considered along with Afug11610 for further characterization.

II.iii. Activity and expression screening of 15 additional Cbhs

I also screened a library of 15 novel fungal Cbhs for their ability to be functionally expressed by *A. niger* strain A2P5 (Table 3). I began by screening each culture filtrate for

cellobiohydrolase activity using the chromogenic substrate pNPC at pH value of 5, in microtitre plates. As a control I also included culture filtrates prepared using *A. niger* transformed with ANIp5, the expression vector used for cloning the 15 novel fungal Cbhs, lacking an insert. The filtrates with highest activity were then chosen for further analysis and characterization.

Table 3: Screening fungal exoglucanases for functional expression in *A. niger*

	Clone ID[1] and Enzyme	Total secreted protein (mg/ml) [2]	Target protein expression (mg/ml) SDS-PAGE [4]	Molecular mass observed (O) and predicted (P)	pNPC activity [5]
1.	PY11 ANIp5 (VTO) /A2P5 ANIp5 (VTO)	1.6	NAp [3]/NAp	NAp/NAp	1 ±0.027/1±0.001
2.	Afu6g11610 # 28/Cbh1	2.46	1.0 PY11/0.8 A2P5	65.5 (O) PY11 / 53.7 (P) PY11	4 ±0.016/26±0.002**
3.	TreeCbh1 transformant # 11	1.66	0.8 PY11/>0.2 A2P5	65 (O) PY11/ 52.3 (P)	5 ±0.044/1±0.016
4.	TreeCbh2 transformant # 25	1.28	0.8 PY11/ND A2P5	55 (O) PY11 / 47.9 (P)	6 ±0.011/1±0.016
5.	Sthe437 transformant # 1/Cbh1	0.56	0.02 PY11/ND A2P5	52 (O) PY11 / 54.0 (P) PY11	5 ±0.017/8±0.001*
6.	Ledo11599 transformant #7 /Cbh2	1.60	0.03 PY11/ND A2P5	55 (O) PY11 / 44.3 (P) PY11	6 ±0.073/1±0.024
7.	Ccin11398 transformant #34 34/Cbh2	0.32	0.01 PY11/ND A2P5	55 (O) PY11/ 42.0 (P) PY11	3 ±0.059/1±0.013
8.	Asn3751 transformant # 1 (Cbh2)	1.37	.02 PY11/>0.2 A2P5	50 (O) PY11 / 46.2 (P)	ND
9.	Seq 235 SnodCbh6B-2 (Cbh2)	ND [3]	NDe[3] A2P5	NDe (O) / 46.7 (P)	1 ±0.015
10.	Seq 271 NdisCbhGH7B-2 (Cbh1)	ND	NDe A2P5	NDe (O) / 48.9 (P)	1 ±0.006
11.	Seq 273 NdisCbhGH6C-2 (Cbh2)	ND	ND A2P5	95 (O) / 37.0 (P)	2 ±0.006*
12.	Seq 274 NdisCbhGH6A-1 (Cbh2)	0.24	NDe A2P5	95 (O) / 49.0 (P)	4 ±0.02**
13.	Seq 276 TvirCbhGH6A-1 (Cbh2)	ND	NDe A2P5	95 (O) / 47.8 (P)	2 ±0.041*
14.	Seq 279 VdahCbhGH7C-1 (Cbh1)	ND	NDe A2P5	95 (O) / 55.5 (P)	1 ±0.011
15.	Seq 284 VdahCbhGH7A (Cbh1)	ND	NDe A2P5	98 (O) / 55.0 (P)	1 ±0.007

16.	Seq 285 AoryCbhGH7A (Cbh1)	2.8	0.5 A2P5	85,80 (O) / 46.9 (P)	18 ±0.044**
17.	Seq 268 SnodCbhGH7A (Cbh1)	ND	NDe A2P5	95 (O) / 43.0 (P)	1 ±0.004
18.	Seq 270 NdisCbhGH7A (Cbh1)	0.26	0.3 A2P5	80 (O) / 53.2 (P)	6 ±0.002**
19.	Seq 275 TvirCbhGH7A (Cbh1)	ND	0.2 A2P5	80 (O) / 51.3 (P)	2 ±0.018*
20.	Seq 277VdahCbhGH7D (Cbh1)	ND	ND A2P5	55 (O) / 46.8 (P)	2 ±0.016*
21.	Seq 278 VdahC bhGH7B (Cbh1)	ND	ND A2P5	57 (O) / 52.3 (P)	1 ±0.012
22.	Seq 282 TterCbhGH7A (Cbh1)	0.52	ND A2P5	72 (O) / 43.4 (P)	27 ±0.035**
23.	Seq 283 VdahCbhGH6A (Cbh2)	ND	ND A2P5	75 (O) / 48.8 (P)	1 ±0.013
24.	VTO	ND	ND A2P5	NAp (O) / NAp (P)	1 ±0.003

[1] Cbh genes with IDs beginning in four letters and Asn3751 were cloned into expression vector ANIp7G (Masters et al.) and transformed into *A. niger* strains PY11 and A2P5. When using PY11 as the expression host 20 to 75 independent transformants were screened for cellobiohydrolase activity and the transformant that expressed the most Cbh activity was selected for characterization (e.g.: transformant # 28 for Afug11610). When using strain A2P5 as the expression host 10 to 30 independent transformants were pooled and conidia from the pooled transformants were used to inoculate a single culture for Cbh production. Cbh genes with IDs beginning with Seq were cloned into expression vector ANIp5, transformed into *A. niger* strain A2P5 and culture filtrates prepared with a pool of 10 to 30 independent transformants.

[2] Total secreted protein (mg/ml) determined for culture filtrates prepared using PY11 transformants was determined using the 2-D-Quant kit.

[3] ND, not determined; NAp, not applicable; and NDe, protein band not detected.

[4] Heterologous protein expression was estimated by SDS-PAGE using a BSA standard and densitometry to determine relative band intensities.

[5] pNPC activity levels as the fold-increased enzyme activity obtained using the *A. niger* transformants indicated relative to the activity observed with the same expression strain transformed with the expression vector without an insert. If there are 2 values, the first value is the relative increased activity observed with strain PY11 transformants and the second value is the relative activity obtained with the pooled A2P5 transformants. Moderate levels of expression * (> 2 times to < 7 times the level of expression obtained with the VTO control): Moderate levels of expression ** (> 7 times the level of expression obtained with the VTO control).

Transformants that expressed at least 8 times the activity levels obtained with the VTO, AoryCbhGH7A, TterCbhGH7A, Afu6gCbhGH7A and Sthe437CbhGH7A, were selected for further characterization.

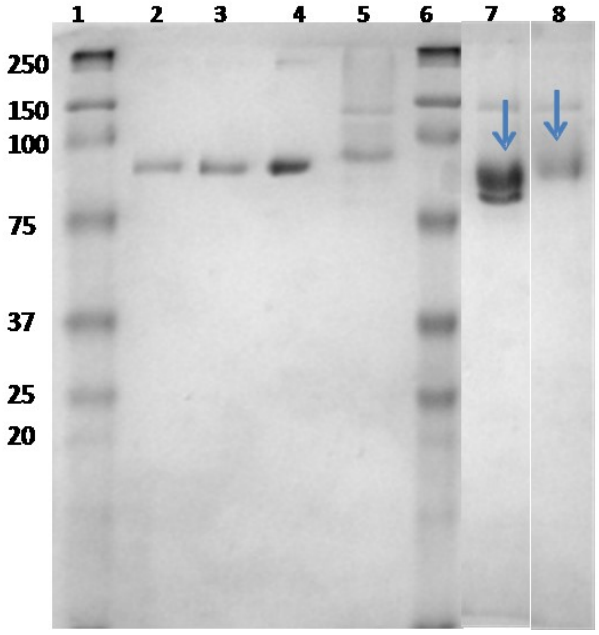


Figure 41: Protein expression levels determined using SDS-PAGE (12%) analysis. Lanes 1 and 6, molecular weight standards with sizes in kDa indicated on the left; lanes 2, 3 and 4 are BSA standards loaded with 5 µg, 7.5 µg and 10 µg of BSA, respectively, lanes 5, 7 and 8 are loaded with 10 µl of culture filtrate prepared from the VTO, AoryCbhGH7A and TterCbhGH7A, respectively.

Protein bands were obtained for both the enzymes TterCbhGH7A and AoryCbhGH7A. A diffused band was seen for TterCbhGH7A. A double band was observed in case of AoryCbhGH7A. One of the bands was 80 kDa and the other 85 kDa, both of which are higher than the predicted weight 46.9 KDa, suggesting differential glycosylation of the protein. Very light protein bands can be observed for the VTO in lane 5.

II.iv. Developing a PASC hydrolysis assay for 2% PASC with Accellerase 1500 activity rate limiting for hydrolysis rates and product yield

In order to work on more concentrated substrate, based on earlier results I devised another set of experiment using 2% PASC instead of 1.5% PASC. Since 2% PASC is 1.33 times more concentrated than 1.5% PASC, I designed an experiment using different concentrations of Accellerase 1500 with an aim of finding the Accellerase 1500 concentration that would yield about 50% hydrolysis at 72 hours. I previously determined that using 26.4 µg/ml of Accellerase 1500 would result in about 50% hydrolysis when using 1.5% PASC as the substrate (Figure 30). When I kept the substrate to enzyme loading ratio the same but used PASC (2 %) and Accellerase 1500 (35.2 µg/ml) I obtained only 42.7% hydrolysis (data not shown). The reduced hydrolysis obtained with the higher concentration of PASC (2%) maybe due to variations in the substrate (PASC), which has to be manually ground and mixed during its preparation (*Material & Method*). Hence I decided to determine the hydrolysis levels obtained with two higher concentrations of Accellerase 1500 (44 µg/ml and 88 µg/ml). The reactions were carried out at 37ºC using a rotator as described in Design # 1 *Materials and Methods*. This experiment (Figure 42) showed that I could use hydrolysis reactions containing 2% PASC and 44 µg/ml Accelerase1500 to assess the ability of the four novel cellulases identified above to synergistically enhance cellulose hydrolysis rates and/or yields.

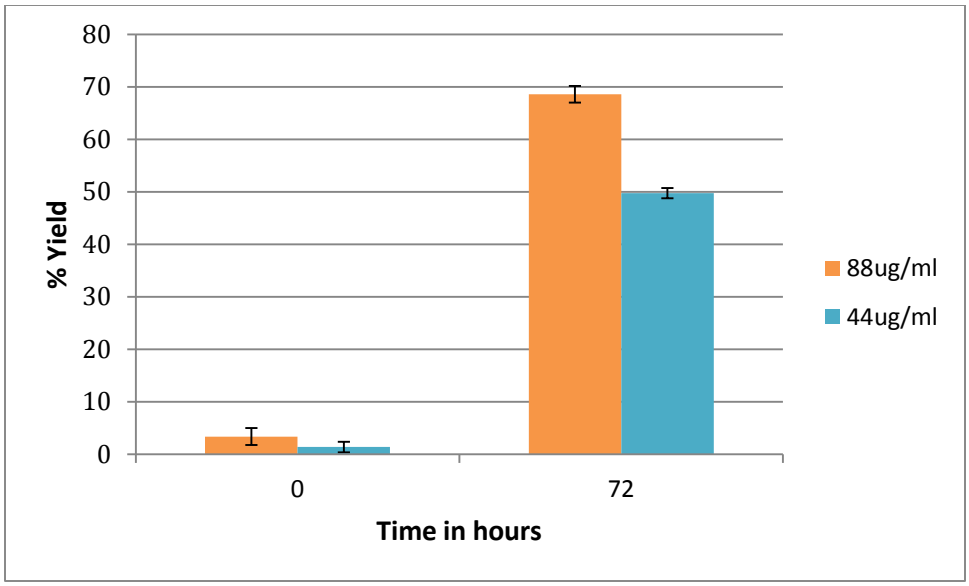


Figure 42: Hydrolysis of PASC (2%) with Accellerase 1500 (44 µg/ml) blue or Accellerase 1500 (88 µg/ml) orange. PASC hydrolysis was performed for 72 h at 37ºC and 80 RPM as described in the *Materials and Methods*. Each of the 1 ml reaction tubes consisted of 500 µl of PASC (4%), 300 µl of citrate buffer (200 mM pH 5), and 200 µl of Accellerase 1500 (220 µg/ml and 440 µg/ml). The yield of reducing sugar ends expressed as a percentage of the theoretical yield if 100% of the 2% PASC was converted to glucose monomers was determined as described for Figure 27. Aliquots were taken at 0 h and 72 h.

II.v. Hydrolysis enhancement studies using the novel enzymes:

The four selected CBH enzymes; Afu6gCbhGH7A, Sthe437CbhGH7A, TterCbhGH7A and AoryCbhGH7A were tested for their ability to enhance the hydrolysis of PASC (2%) by Accellerase 1500 using hydrolysis strategies Design #1 and Design #2 (Figure 43 and 44) as described above.

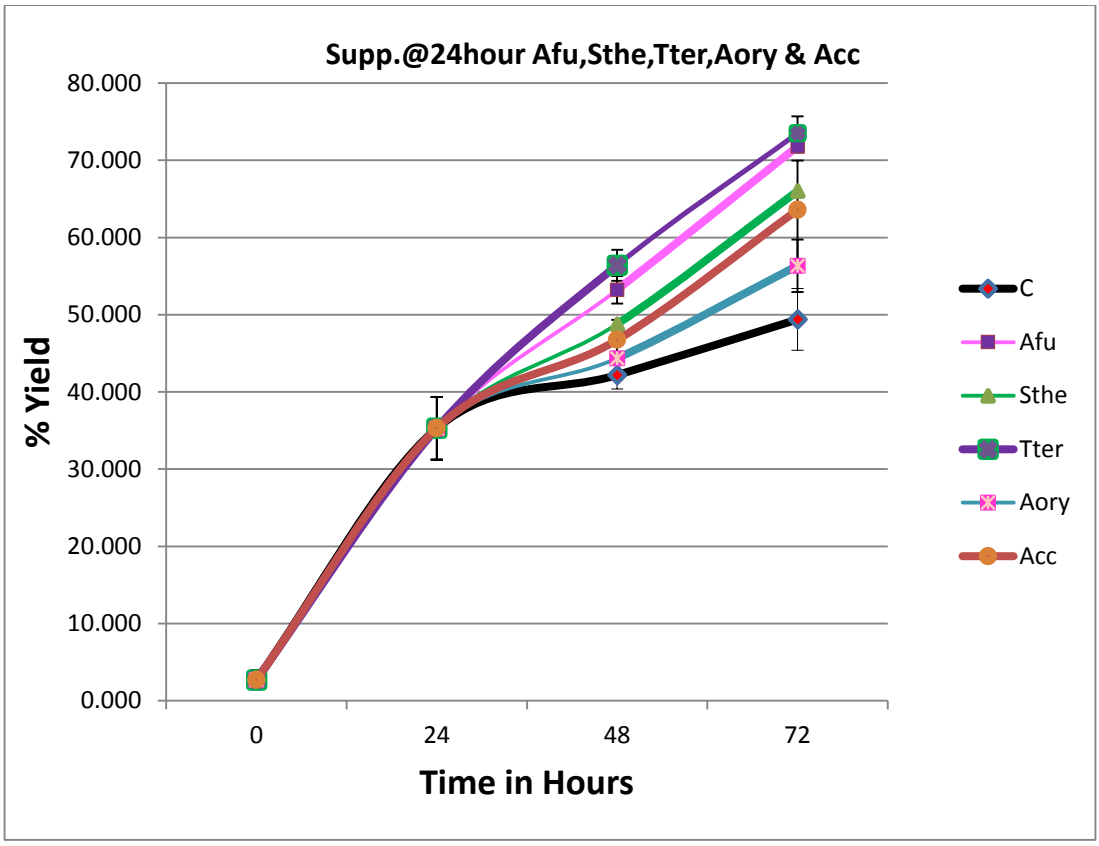


Figure 43: Design #1, enhancement of PASC (2%) hydrolysis by Accellerase 1500 (44 µg/ml) by supplementation at 24 h with Afu6gCbhGH7A, Sthe437CbhGH7A, AoryCbhGH7A and TterCbhGH7A. At 0 h a 3 ml PASC hydrolysis reaction (PASC (2%) and Accellerase 1500 (44 µg/ml) prepared by combining 1,500 µl of PASC (4%), 900 µl of citrate buffer (200 mM pH5), and 600 µl of Accellerase 1500 (220 µg/ml) was incubated at 37ºC and 80 RPM. At 24 h the 3 ml reaction was divided into six 0.5 ml reactions and the indicated enzymes (22 µg in 100 µl) were added. The yield of reducing sugar ends expressed as a percentage of the theoretical yield if 100% of the 2% PASC was converted to glucose monomers was determined as described for Figure 27. The yield of reducing sugar equivalents was determined using aliquots taken at 0, 24, 48 and 72 h.

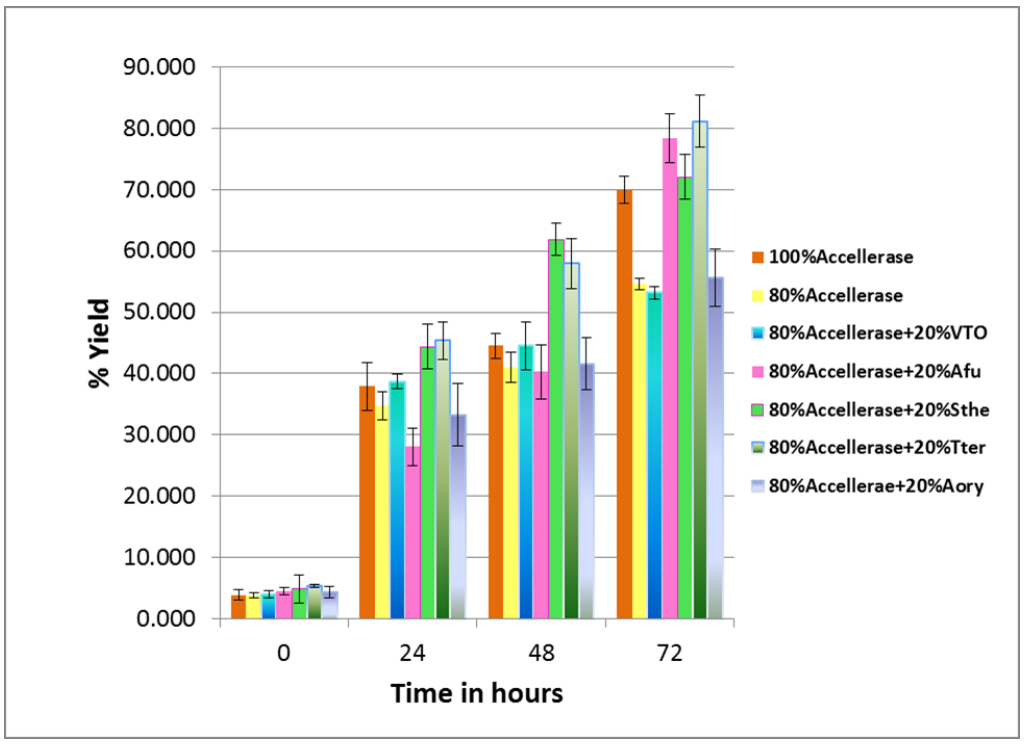


Figure 44: Design #2, enhancement of PASC (2%) hydrolysis by Accellerase 1500 (70.4 µg/ml) by supplementation with novel Cbhs, Afu6gCbhGH7A, Sthe437CbhGH7A, AoryCbhGH7A and TterCbhGH7A (17.6 µg/ml). Hydrolysis reactions (500 µl) were prepared by combining 250 µl of PASC (4%), 150 µl of citrate buffer (200 mM pH 5), 80 µl of Accellerase 1500 (440 µg/ml) and 20 µl of Afu6gCbhGH7A, Sthe437CbhGH7A, AoryCbhGH7A, TterCbhGH7A or Accellerase 1500 (440 µg/ml) or 10 mM citrate buffer (10 mM pH 5). PASC hydrolysis assays were incubated for 72 h at 37ºC and 80 RPM as described for Design #2. The yield of reducing sugar ends expressed as a percentage of the theoretical yield if 100% of the 2% PASC was converted to glucose monomers was determined as described for Figure 27.

In Design#1, supplementation with enzyme TterCbhGH7A increased hydrolysis by about 24% and that by Afu6gCbhGH7A increased hydrolysis about 22% as compared to the unsupplemented reaction. Sthe437CbhGH7A and Accellerase 1500 supplementation increased hydrolysis by approximately 14% and 16%, whereas enzyme AoryCbhGH7A supplementation increased hydrolysis the least, about 7%. Using Design#2, I observed that supplementation at T= 0 h with enzymes TterCbhGH7A and Afu6gCbhGH7A increased hydrolysis by 27% and 24% respectively, supplementing with Sthe437CbhGH7A and Accellerase 1500 increased hydrolysis

by 18% and 15% respectively, whereas supplementing the hydrolysis reaction with enzyme AoryCbhGH7 did not increase PASC hydrolysis significantly. A summary comparing the results obtained using Design#1 and Design#2 is presented in Table 4.

Table 4: Summary of percentage hydrolysis of PASC (2%) at 72 h using novel cellulases and screening methods Designs #1 and Design #2

	Design 1		Design 2	
No.	**Enzymes used**	**Percentage hydrolysis of PASC [1]**	**Sample 72h**	**Percentage hydrolysis of PASC**
1	Accellerase 1500 (44 µg/ml)	49.4 NA[3]	Accellerase 1500 (70.4 µg/ml)	54.5 NA
2	Accellerase 1500 (44 µg/ml) + Accellerase1500 (1 X) [2]	63.6 28.7%/0% [4]	Accellerase 1500 (88 µg/ml)	69.3 27.1%/0% [5]
3	Accellerase 1500 (44 µg/ml) + Afu6gCbhGH7A (1 X)	71.8 45.3%/12.9%	Accellerase 1500 (70.4 µg/ml) + Afu6gCbhGH7A (17.6 µg/ml)	78. 3 43.7%/13%
4	Accellerase 1500 (44 µg/ml) Sthe437CbhGH7A (1 X)	66.0 33.6%/3.8%	Accellerase 1500 (70.4 µg/ml) Sthe437CbhGH7A (17.6 µg/ml)	72.0 32. 1%/3.9%
5	Accellerase 1500 (44 µg/ml) + AoryCbhGH7A (1 X)	56.4 14.2%/-11.3%	Accellerase 1500 (70.4 µg/ml) AoryCbhGH7A (17.6 µg/ml)	55.6 2.1%/-19.8%
6	Accellerase 1500 (44 µg/ml) + TterCbhGH7A (1 X)	73.5 48.8%/15.6%	Accellerase 1500 (70.4 µg/ml) + TterCbhGH7A (17.6 µg/ml)	81.1 48.8%/17.0%
7	Accellerase 1500 (44 µg/ml) + VTO(1 X)	47.9 -3.1%/-24.7%	Accellerase 1500 (70.4 µg/ml) + VTO (17.6 µg/ml)	53.2 -2.3%/-23.3%

[1] Yield of reducing sugar ends expressed as a percentage of the theoretical yield of reducing sugar ends expected if 100 % of the 2% PASC was converted to glucose monomers. The yield of total reducing sugar equivalents was determined for Design#1and 2, as described for Figure 43 and 44. PASC hydrolysis was

performed for 72 h at 37 ºC and 80 RPM as described in *Materials and Methods*. Aliquots were taken at time 0, 24, 48 and 72 hours and reducing sugar ends in the whole mixture were determined.

[2] PASC hydrolysis reactions reaction (500 µl) with Accellerase 1500 (44 µg/ml) and supplementation at 24 h with 100 µl of the indicated enzyme or enzyme system (22 µg) final volume of 600 µl of hydrolysate.

[3] NA indicates not applicable.

[4] The first of the two percentage values indicates the hydrolysis enhancement obtained when supplementing with the indicated test enzyme relative to that obtained with only 44 µg/ml or Accellerase 1500 added at t = 0 h, whereas the second percentage value indicates the enhancement observed relative to that obtained when 44 µg/ml Accellerase 1500 was supplemented with Accellerase 1500 at 44 µg/ml at t = 24 h.

[5] The first of the two percentage values indicates the hydrolysis enhancement obtained when supplementing with 17.6 µg/ml the indicated test enzyme relative to that obtained with only 70.4 µg/ml of Accellerase 1500 added at t = 0 h, whereas the second percentage value indicates the enhancement observed relative to that obtained with 70.4 µg/ml Accellerase 1500 supplemented with 17.6 µg/ml of Accellerase 1500 at t = 0 h.

III.i. End product analysis using the Dionex ICS-5000 Ion Chromatography System

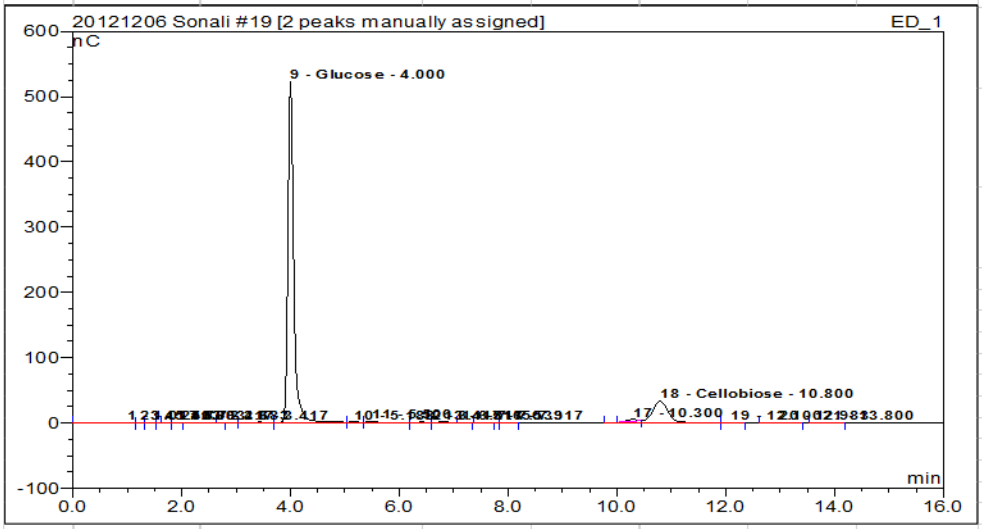
For end product analysis, I used samples obtained at 72 h from experiments performed by using the enzymes in Table: 2 and novel enzyme Afu6gCbhGH7A using Design#1. Dionex CarboPac column 10 was used to analyse each hydrolysis reaction sample (10 µl**)**. Standards were prepared by using 0.000244 mM to 0.50 mM concentrations of glucose and cellobiose. Figure 19 presents the analysis obtained from the Dionex system for sample Eg1-72 (Iogen *T. reesei* system (WS) without Eg1) at 72h. The results obtained showed two peaks. The first large peak indicated presence of glucose and second peak indicated cellobiose. Peak height and peak area values of the standards obtained were plotted on graphs. The graphed standards were used to determine the amount of glucose and cellobiose in each of the 21 different 72 h hydrolysis reactions presented in Table 5.

The major end products obtained were glucose and cellobiose. The concentration of glucose was found to be significantly higher than that of cellobiose. As observed when measuring hydrolysis rates using the BCA assay, the highest levels of hydrolysis observed with Dionex analysis were Accellerase 1500 supplemented with Afu6gCbhGH7A (>73%), followed by Accellerase 1500 supplemented Cbh2 (>67%), followed by binary combination Cbh1 and Bgl (>61%) and Cbh2 + Bgl (57.8%). The lowest glucose yields were obtained with the VTO (45.8%), Eg1 (46.6) and Eg2 (42.5%) respectively (Table 5). The lowest levels of cellobiose production were obtained in samples supplemented with only Bgl or the binary combinations that included Bgl with two exceptions. One exception was the binary combination of Bgl and Eg2, which showed that cellobiose accounted for greater than 2 % of the input PASC and when supplementation was with the VTO where cellobiose was less than 0.1 % of the input PASC. The above results suggest that Afu6gCbhGH7A is a strong candidate for synergistic action with

commercial cellulase systems and that the VTO produced by expression strain A2P5 contains significant Beta-glucosidase activity.

19 Eg1-72

Sample Name:	Eg1-72	*Injection Volume:*	10.0
Vial Number:	GC7	*Channel:*	ED_1
Sample Type:	unknown	*Wavelength:*	n.a.
Control Program:	50% 200 mM NaOH, 0,5 ml per min, 16 mi	*Bandwidth:*	n.a.
Quantif. Method:	IQ FINAL PM	*Dilution Factor:*	1.0000
Recording Time:	12/7/2012 17:37	*Sample Weight:*	1.0000
Run Time (min):	16.00	*Sample Amount:*	1.0000



No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	1.02	n.a.	0.595	0.359	0.44	n.a.	BMb
2	1.25	n.a.	0.607	0.078	0.10	n.a.	bM
3	1.43	n.a.	0.832	0.142	0.17	n.a.	M
4	1.70	n.a.	0.035	0.002	0.00	n.a.	Ru
5	1.88	n.a.	0.027	0.002	0.00	n.a.	Ru
6	2.42	n.a.	0.933	1.012	1.24	n.a.	M
7	2.68	n.a.	0.141	0.010	0.01	n.a.	Rd
8	3.42	n.a.	1.667	0.449	0.55	n.a.	M
9	4.00	Glucose	522.829	65.550	80.34	0.798	M ^
10	5.18	n.a.	1.397	0.336	0.41	n.a.	M
11	5.50	n.a.	2.501	0.760	0.93	n.a.	M
12	6.42	n.a.	0.963	0.261	0.32	n.a.	M
13	6.82	n.a.	1.217	0.421	0.52	n.a.	M
14	7.15	n.a.	0.039	0.005	0.01	n.a.	Rd
15	7.53	n.a.	0.056	0.011	0.01	n.a.	Rd
16	7.92	n.a.	0.052	0.011	0.01	n.a.	MB
17	10.30	n.a.	2.615	0.576	0.71	n.a.	Ru
18	10.80	Cellobiose	32.831	11.560	14.17	0.029	BMb^
19	12.10	n.a.	0.021	0.005	0.01	n.a.	bMB
20	12.98	n.a.	0.073	0.028	0.03	n.a.	BMB
21	13.80	n.a.	0.032	0.013	0.02	n.a.	BMB
Total:			569.464	81.591	100.00	0.828	

Figure 45: Sample analysis chart presented by Dionex ICS 5000 analysis of a 72 h PASC (2%) hydrolysis reaction with WS – EG1. Top panel includes the following; the sample name/number, the sample injection volume, the flow rate of eluent and the run time in minutes. **Middle panel,** the first tall peak indicating glucose is seen at 4.0 min and a very short peak indicating cellobiose at 10.35 min. The **lower panel** shows **the** retention time, peak height, peak area, relative area and relative amount of the peaks with the only significant peaks correspond to glucose and cellobiose.

Table 5: Summary of percentage hydrolysis of PASC at 72 h.

(Results obtained using Dionex ICS 5000 for sample analysis at 72 h)

No.	Accellerase 1500 (C) or Accellerase supplemented with the indicated samples 72h	Percentage hydrolysis of PASC (%) to Glucose	Percentage hydrolysis of PASC (%) to Cellobiose
1	C	47.5	2.6
2	Cbh1	55.8	2.5
3	Cbh2	67.3	2. 8
4	Eg1	46.6	1.7
5	Eg2	42.6	2.1
6	Acc1500	60.8	1.4
7	Bgl	56.3	0.1
8	Cbh1+Eg1	52.9	2.8
9	Cbh1+Eg2	54.8	2.2
10	Cbh1+Bgl	61.0	0.9
11	Cbh2+Eg1	50.9	2.7
12	Cbh2+Eg2	42.7	2.4
13	Cbh2+Bgl	57.8	0.4
14	Eg1+Bgl	54.0	0.6
15	Eg2+Bgl	51.5	2.2
16	Cbh1+Cbh2	51.4	2.1
17	W.S.	63.5	0.6
18	$Eg1^{-}$	58.7	2.5
19	$Eg2^{-}$	57.8	2.5
20	VTO	45.8	0.1
21	* Afu6gCbhGH7A	72.9	2.7

Table 5: Summary of percentage hydrolysis of PASC at 72 h using Dionex ICS 5000. All the samples analysed for glucose and cellobiose concentrations were obtained using the 72 h hydrolysis reaction using Accellerase 1500 and the indicated supplementations. For each hydrolysis reaction a 10 µl sample was analysed using Dionex CarboPac column 10. Yield of reducing sugar ends expressed as a percentage of the theoretical yield of reducing sugar ends expected if 100% of the 1.5% or 2%(*) PASC was converted to glucose monomers or cellobiose. The yield of total reducing sugar equivalents for 1.5% PASC was determined as described for Figure 31 and for 2% PASC was determined as for Figure 43. The

yield of total reducing sugar equivalents (cellobiose) for 1.5% or 2% PASC was determined as complete hydrolysis of 1.5% or 2% PASC would yield $3.5x10^{-5}$ moles or $5.8x10^{-5}$ of cellobiose dimers (theoretical yield).

IV.Bioinformatic analysis of cellobiohydrolase sequences

I identified six Cbh genes *Afu6g11610CbhGH7A, Sthe437CbhGH7A, Seq274NdisCbhGH6A, Seq-285AoryCbhGH7A, Seq270NdisCbhGH7A and Seq-282TterCbhGH7A* that could be functionally expressed by *A. niger* strain A2P5 (Table 3). I used all six novel enzymes for phylogenetic analysis. I began the phylogenetic analysis by using the protein sequences encoded by these six genes, which are available in Appendix Table 1A, to perform a ClustalW alignment. The ClustalW alignments were then used to generate a phylogram (Figure 46). The six cellobiohydrolases, (Afu6gCbhGH7A, StheCbhGH7A, NdisCbhGH6A, AoryCbhGH7A, NdisCbhGH7A and TterCbhGH7A) fell into three distinct groups. Afu6gCbh1GH7A, NdisCbh1GH7A, StheCbh1GH7A and AoryCbh1GH7A, which all belong to GH subfamily GH7A, formed one group; whereas, TterCbhGH7A, apparently belongs to its own group as did NdisCbh2GH6A, which belongs to GH family GH6A. The fact that TterCbhGH7A did not group with the other GH7A enzymes was unexpected. I therefore used the predicted amino acid sequence of TterCbhGH7A to do a protein/protein Blast search of the non-redundant protein database at NCBI. The results obtained indicated that TterCbhGH7A not does not belong to the GH7A/Cbh1 subgroup but rather belongs to the GH7B/Eg1 subgroup of GH7. Protein sequence alignments also showed that Afu6gCbhGH7A, CbhGH7A, NdisCbhGH7A and NdisCbhGH6A, include a family one carbohydrate binding module (CBM1) domain, where the three GH7A members have the CBM1 at their C-terminal ends and the lone GH6A member has a

CBM1 at its N-terminal end. Neither AoryCbhGH7A nor TterCbhGH7A, which is apparently a Cbh7B, have a CBM1.

Phylogram:

A phylogram is a branching diagram showing the inferred evolutionary relationships between various biological species or other entities based on physical and/or genetic similarities and differences. Only the enzymes that were successfully expressed at "high levels" were used to make the Phylogram (Figure 46).

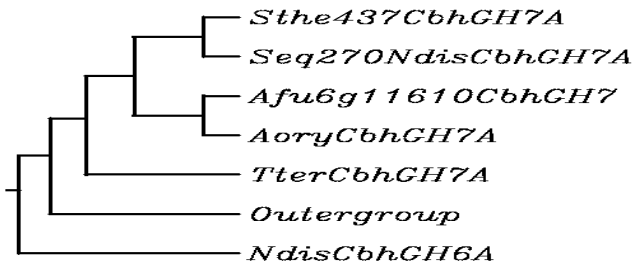


Figure 46. Molecular Phylogenetic tree constructed using ClustalW.

The phylogenetic tree is based on the degree of similarities between the aligned sequences. It was built using ClustalW. Phylogenetic tree was rooted using an outgroup. Outgroups are similar to the sequences of interest but belong to a species that diverged earlier than any of the members that are being compared. For cellobiohydrolases expressed in fungal species, the outer group used was endoglucanase from bacteria Ruminiclostridium thermocellum BC1.

DISCUSSION

The main objective of this study was to screen a library of fungal cellobiohydrolases for enzymes that could enhance the hydrolysis the model cellulosic substrate PASC by the commercial cellulase system Accellerase 1500. In the first set of experiments I followed the total amount of reducing sugar ends produced and the amount of soluble reducing sugar ends produced during the hydrolysis of PASC by Accellerase 1500. The results of this experiment showed that the amount of reducing ends present in the supernatant was significantly less than the amount present in the whole reaction after 1 h and 2 h of incubation; however, after incubation for 24 h the amount of reducing sugar ends in the supernatant and whole reaction were not significantly different.

In order to identify which component or components of the Accellerase1500 cellulase system were limiting the production of and or yield of reducing sugar ends during the hydrolysis of PASC by Accellerase 1500, I began by establishing hydrolysis conditions where the amount of Accellerase 1500 used limited the hydrolysis rate and yield when 1.5% PASC was used as the substrate. Based on the results from these experiments (Figures 28, 29 and 30) I decided to use a 72 hour hydrolysis reaction using PASC (1.5% w/v), 32 µg/ml of Accellerase 1500 at 50 °C, because this combination resulted in a limiting hydrolysis rate and yield. These results, presented in Figures 28, 29 and 30, also showed that supplementing with additional Accellerase 1500 enhanced not only the hydrolysis rates but also the yield suggesting that supplementation replaces an Accellerase 1500 component that is inactivated during the hydrolysis. Banerjee G et al. 2010, 2011 and Gao D et al. 2010, in their studies found that optimal cellulase system composition varied depending on the type of substrate used. Although my results show that one or more components of the cellulase system were limiting hydrolysis they did not show which

component(s) were limiting. To identify which component(s) was limiting hydrolysis rates I supplemented my rate limiting hydrolysis reaction (72 hour hydrolysis reaction performed at 50 °C using PASC (1.5% w/v) as the substrate and 132 µg/ml of Accellerase 1500 at 24 hours with 132 µg/ml of each of the four major *T. reesei* cellulases Cbh1, Cbh2, Eg1 and Eg2, a Bgl from *A. niger* and Accellerase1500 (Table 2). The results indicated that supplementing the reaction mixture with either Accellerase 1500, or Cbh2 enhanced the amount of reducing sugar ends produced by the most, 29% and 33%, respectively. Cbh1, Eg1 and Bgl enhanced the yields by about 14%, 11%, and 9%, respectively. Interestingly, Eg2 supplementation actually reduced the amount of reducing sugar ends produced by about 7%. I also supplemented the reaction with binary combinations of Cbh1, Eg1 Eg2 and Bgl (a 66 µg/ ml of each enzyme). With one exception, all binary combinations reduced the amount of reducing ends produced relative to the amount obtained when supplementation was with 132 µg/ml of only the top performing cellulase system component. Supplementation with the binary combination Cbh1 and Bgl enhances hydrolysis by 38%. This was significantly greater than the sum of the amount obtained when supplementation was with only Cbh1 (14.4%) or only Bgl (8.7%). This result suggests that Cbh1+Bgl showed synergism.

My results with the binary combination, Cbh1 and Bgl are consistent with the previous past reports of synergism between exocellulases Cbh1 and Cbh2, between endocellulases and exocellulases and between Bgls and cellulases (Andersen N et al. 2010, Gao PJ et al. 2010, Woodward J et al. 1991).

I screened a library of Cbh enzymes that were transformed into *A. niger* host strain PY11. I noticed a high protein background in the PY11 filtrates (Fig13). Hence, the recombinant proteins were expressed again, using A. *niger* strain A2P5, which produced much lower levels

of background activity (Fig 14). However, only two of the above seven enzymes were successfully expressed Afu6g11610CbhGH7A and Sthe437CbhGH7A. Both the enzymes were expressed at high levels but only Afu6gCbhGH7A showed a protein band on 12% SDS-PAGE gel (Fig 14). These two enzymes were characterised further using hydrolysis enhancement studies (data discussed later).

Culture filtrates were also prepared using *A. niger* strain A2P5 transformed with expression vector ANIp7MS harbouring 15 different Cbh cDNA or gDNA clones. The 15 culture filtrates were screened for cellobiohydrolase activity using pNPC as the substrate. CBHs are able to hydrolyze the agluconic bond between the nitrophenyl and carbohydrate moiety in pNPC producing *p*-nitrophenol, which can be measured using a spectrophotometer (Table 3). Four of the total 15 culture filtrates exhibited high very significant levels of cellobiohydrolase activity as compared to the negative control (VTO). These enzymes, AoryCbhGH7A, NdisCbhGH6A-1, NdisCbhGH7A and TterCbhGH7A were further examined for protein expression using SDS-PAGE (Figure15) and 2-D-Quant kit analysis (Table 4). No band was observed for enzyme NdisCbhGH6A. A double band was observed in case of AoryCbhGH7A suggesting heterogeneous glycosylation.

I selected the four enzymes that were expressed at the highest levels by A2P5, Afu6g11610CbhGH7A, Sthe437CbhGH7A, AoryCbhGH7A and TterCbhGH7A for hydrolysis enhancement studies using PASC (2% w/v). I decided to use two different designs, Design#1 and Design#2 to study PASC hydrolysis enhancement. As explained earlier in Design#1, Accellerase 1500 was added at the beginning (T= 0 h) to the reaction with 2% PASC and at 24 hours (T=24 h) reactions was supplemented with equal amount (220 µg/ml) of each novel cellulases. Adding Accellerase 1500, at the beginning to the reaction might permit loosening of

the compact chains, which then could be acted upon by the supplemented enzymes that were being tested. The advantage of this design was that variations that arise within single experiment could be reduced by using a common starting stock of hydrolysate containing Accellerase 1500 for testing the enzymes. However, when the stock hydrolysate was transferred into different tubes there was a possibility that unequal proportion of PASC and Accellerase1500 was being distributed within the tubes causing variations in the results. In order to overcome this shortcoming of the experimental Design#1, Design#2 was also used. In the Design#2 method, the commercial cellulase and the novel cellulase being tested were added at the beginning to the reaction. Separate tubes were prepared for each time point thus the transfer of hydrolysate between tubes was avoided. Testing the enzymes on different starting stock caused some variations in the results within the same experiment but gave greater resolution with respect to enzyme action on PASC.

Supplementation of hydrolysis reaction with enzyme TterCbhGH7A increases hydrolysis by about 24% and that by Afu6gCbhGH7A with 22% respectively, in Design#1 whereas by 27% and 24% in Design#2. Sthe437CbhGH7A and Accellerase1500 supplementation showed an increase by approximately 16% and AoryCbhGH7A by 6% in Design#1, whereas in Design#2 the Sthe437CbhGH7A and Accellerase1500 supplementation showed an approximate increase by 15% and 18% respectively. No significant increase was shown by AoryCbhGH7A.

Supplementation of the hydrolysis reaction with enzymes TterCbhGH7A and Afu6gCbhGh7A enhanced hydrolysis by 48.8% and 45.3% in Design#1 and 48.8% and 43.7%, in Design#2. StheCbhGH7A, AoryCbhGH7A and Accellerase 1500 supplementation increased hydrolysis by 33.6%, 14.2% and 28.7% in Design#1, and 45.5%, 2.2% and 30.5% in Design#2.

The above results clearly show that both the designs are capable of screening enzymes based on their hydrolytic potentials. Design#1 however has the advantage of being able to screen a larger number of enzymes at the same time because fewer tubes are required (Design#1 would need 12 tubes for 1 enzymes whereas Design#2 would require 29). Enzymes Afu6gCbhGh7A, TterCbhGH7A and StheCbhGH7A all look like promising candidate enzymes since they enhanced the hydrolysis rates and yields to a greater extent than did Accellerase 1500.

Product analysis was performed on PASC hydrolysis reactions using Dionex ICS 5000. Product analysis was performed on samples harvested after 72-hour of hydrolysis of using the enzymes listed in Table 1 as well as Afu6gCbhGH7A. The results showed that the major soluble product detected in all cases was glucose and that cellobiose was present at very low detectable levels. This may be due to the glucose resistant β-glucosidase from *A. niger*, which is included in Accellerase 1500, preventing the accumulation of cellobiose.

The ion chromatography results show that the glucose levels obtained in the hydrolysis reactions are similar to the levels of reducing sugar ends determined using the BCA assay (compare the results in table 4 with table 2). Hydrolysis reactions where Accellerase 1500 was supplemented with Afu6gCbhGH7A produced the highest of glucose followed by the reactions supplemented with Cbh2, Cbh1+Bgl, and then Accellerase1500.

The bioinformatics analysis of the protein sequences of the six screened enzymes show that Sthe437, NdisCbhGH7A, Afu6g11610 and AoryCbhGH7A belong to Glycosyl Hydrolase Family 7 (GH7), whereas NdisCbhGH6A belongs to Glycosyl Hydrolase Family 6 (GH6). TterCbhGH7A belonged to its own group. Sthe437, NdisCbhGH7A, Afu6g11610 and NdisCbhGH6A all include a conserved family one carbohydrate binding module (CBM1) or

fungal carbohydrate binding domain (fCBD), with the three GH7 enzymes having the CBM at its C-terminal end and the GH6 enzyme having its CBM at its amino terminal end.

Enzymes that belong to GH7 may have three different activities using retaining mechanism: endoglucanase activity, reducing-end cellobiohydrolase activity, or chitosanases activity. Indeed the cloned *400-Seq275TterCbhGH7A* gene does not encode a C-terminal CBM and shows a high degree of similarity both GH7 member endoglucanases and exoglucanases. Whether this enzyme is an endoglucanase or exoglucanase awaits further analysis including sequence alignments to determine whether it like the exoglucanases possesses the inserts for formation of the loops that form the substrate tunnel.

REFERENCES

Aachmann FL, Sorlie M, Skjak-Braek G, Eijsink VGH, Vaaje-Kolstad G (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. Proc Natl Acad Sci U S A 109: 18779-18784

Abuja P, Pilz I, Claeyssens M, Tomme P (1988) Domain-Structure of Cellobiohydrolase-Ii as Studied by Small-Angle X-Ray-Scattering - Close Resemblance to Cellobiohydrolase-i. Biochem Biophys Res Commun 156: 180-185

Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH, Westereng B (2014) Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. Proc Natl Acad Sci U S A 111: 6287-6292

Alonso DM, Wettstein SG, Dumesic JA (2012) Bimetallic catalysts for upgrading of biomass to fuels and chemicals. Chem Soc Rev 41: 8075-8098

Amaya M, Watts A, Damager I, Wehenkel A, Nguyen T, Buschiazzo A, Paris G, Frasch A, Withers S, Alzari P (2004) Structural insights into the catalytic mechanism of Trypanosoma cruzi trans-sialidase. Structure 12: 775-784

Andersen N, Johansen KS, Michelsen M, Stenby EH, Krogh KBRM, Olsson L (2008) Hydrolysis of cellulose using mono-component enzymes shows synergy during hydrolysis of phosphoric acid swollen cellulose (PASC), but competition on Avicel. Enzyme Microb Technol 42: 362-370

Arantes V, Saddler JN (2010) Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. Biotechnology for Biofuels 3: 4

Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Walton JD (2010) Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. Biotechnology for Biofuels 3: 22

Banerjee G, Car S, Scott-Craig JS, Hodge DB, Walton JD (2011) Alkaline peroxide pretreatment of corn stover: effects of biomass, peroxide, and enzyme loading and composition on yields of glucose and xylose. Biotechnology for Biofuels 4: 16

Banerjee G, Scott-Craig JS, Walton JD (2010) Improving Enzymes for Biomass Conversion: A Basic Research Perspective. Bioenergy Research 3: 82-92

Barr B, Hsieh Y, Ganem B, Wilson D (1996) Identification of two functionally different classes of exocellulases. Biochemistry (N Y) 35: 586-592

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

Beckham GT, Matthews JF, Bomble YJ, Bu L, Adney WS, Himmel ME, Nimlos MR, Crowley MF (2010) Identification of Amino Acids Responsible for Processivity in a Family 1 Carbohydrate-Binding Module from a Fungal Cellulase. J Phys Chem B 114: 1447-1453

Beguin P (1990) Molecular-Biology of Cellulose Degradation. Annu Rev Microbiol 44: 219-248

Berlin A, Maximenko V, Gilkes N, Saddler J (2007) Optimization of enzyme complexes for lignocellulose hydrolysis. Biotechnol Bioeng 97: 287-296

Bhowmick R, Ghosal A, Das B, Koley H, Saha DR, Ganguly S, Nandy RK, Bhadra RK, Chatterjee NS (2008) Intestinal Adherence of Vibrio cholerae Involves a Coordinated Interaction between Colonization Factor GbpA and Mucin. Infect Immun 76: 4968-4977

Boisset C, Fraschini C, Schulein M, Henrissat B, Chanzy H (2000) Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from Humicola insolens and its mode of synergy with cellobiohydrolase Cel7A. Appl Environ Microbiol 66: 1444-1452

Boraston A, Bolam D, Gilbert H, Davies G (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382: 769-781

Borjesson P (2009) Good or bad bioethanol from a greenhouse gas perspective - What determines this? Appl Energy 86: 589-594

Brett C (2000) Cellulose microfibrils in plants: Biosynthesis, deposition, and integration into the cell wall. International Review of Cytology - a Survey of Cell Biology, Vol 199 199: 161-199

Bu L, Beckham GT, Crowley MF, Chang CH, Matthews JF, Bomble YJ, Adney WS, Himmel ME, Nimlos MR (2009) The Energy Landscape for the Interaction of the Family 1 Carbohydrate-Binding Module and the Cellulose Surface is Altered by Hydrolyzed Glycosidic Bonds. J Phys Chem B 113: 10994-11002

Buranov AU, Mazza G (2008) Lignin in straw of herbaceous crops. Industrial Crops and Products 28: 237-259

Buschiazzo A, Alzari PM (2008) Structural insights into sialic acid enzymology. Curr Opin Chem Biol 12: 565-572

Cannella D, Hsieh CC, Felby C, Jorgensen H (2012) Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. Biotechnology for Biofuels 5: 26

Chanzy H, Imada K, Mollard A, Vuong R, Barnoud F (1979) Crystallographic Aspects of Sub-Elementary Cellulose Fibrils Occurring in the Wall of Rose Cells Cultured Invitro. Protoplasma 100: 303-316

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25: 759-761

Chedin J M. Progress in understanding of liquid reaction mediums and interpretations of their reactions with cellulosic fibres. A. Chim. Ind (Paris) 1954 71: 55

Chir J, Wan C, Chou C, Wu A (2011) Hydrolysis of cellulose in synergistic mixtures of beta-glucosidase and endo/exocellulase Cel9A from Thermobifida fusca. Biotechnol Lett 33: 777-782

Chundawat SPS, Bellesia G, Uppugundla N, Sousa LdC, Gao D, Cheh AM, Agarwal UP, Bianchetti CM, Phillips GN,Jr., Langan P, Balan V, Gnanakaran S, Dale BE (2011) Restructuring the Crystalline

Cellulose Hydrogen Bond Network Enhances Its Depolymerization Rate. J Am Chem Soc 133: 11163-11174

Curran MA (2007) Co-product and input allocation approaches for creating life cycle inventory data: A literature review. International Journal of Life Cycle Assessment 12: 65-78

Davies G, Henrissat B (1995) Structures and Mechanisms of Glycosyl Hydrolases. Structure 3: 853-859

Demirbas A (2008) Biofuels sources, biofuel policy, biofuel economy and global biofuel projections. Energy Conversion and Management 49: 2106-2116

Detroy R, Cunningham R, Bothast R, Bagby M, Herman A (1982) Bioconversion of Wheat Straw Cellulose Hemicellulose to Ethanol by Saccharomyces-Uvarum and Pachysolen-Tannophilus. Biotechnol Bioeng 24: 1105-1113

Dimarogona M, Topakas E, Christakopoulos P (2012) Cellulose degradation by oxidative enzymes. Computational and structural biotechnology journal 2: e201209015-e201209015

Din N, Damude H, Gilkes N, Miller R, Warren R, Kilburn D (1994) C-1-C-X Revisited - Intramolecular Synergism in a Cellulase. Proc Natl Acad Sci U S A 91: 11383-11387

Ding S, Liu Y, Zeng Y, Himmel ME, Baker JO, Bayer EA (2012) How Does Plant Cell Wall Nanoscale Architecture Correlate with Enzymatic Digestibility? Science 338: 1055-1060

Ding S, Rincon M, Lamed R, Martin J, McCrae S, Aurilia V, Shoham Y, Bayer E, Flint H (2001) Cellulosomal scaffoldin-like proteins from Ruminococcus flavefaciens. J Bacteriol 183: 1945-1953

Divne C, Stahlberg J, Reinikainen T, Ruohonen L., Pettersson, G., Knowles, J K & Jones T A (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from Trichoderma reesei. Science 265(5171): 524-528.

Divne C, Stahlberg J, Teeri TT, Jones TA (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 angstrom long tunnel of cellobiohydrolase I from Trichoderma reesei. J Mol Biol 275: 309-325

Fairley P (2011) Next generation biofuels. Nature 474: S2-S5

Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC (2011) Nanostructure of cellulose microfibrils in spruce wood. Proc Natl Acad Sci U S A 108: E1195-E1203

Fontes CMGA, Gilbert HJ (2010) Cellulosomes: Highly Efficient Nanomachines Designed to Designed to Deconstruct Plant Cell Wall Complex Carbohydrates. Annual Review of Biochemistry, Vol 79 79: 655-681

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei. J Biol Chem 278: 31988-31997

Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, Sorlie M, Horn SJ, Eijsink VGH (2011) Cleavage of cellulose by a CBM33 protein. Protein Science 20: 1479-1483

Fuchs G, Boll M, Heider J (2011) Microbial degradation of aromatic compounds - from one strategy to four. Nature Reviews Microbiology 9: 803-816

Fuchs RL, Mcpherson SA, Drahos DJ (1986) Cloning of a Serratia-Marcescens Gene Encoding Chitinase. Appl Environ Microbiol 51: 504-509

Gao PJ, Chen GJ, Wang TH, Zhang YS, Liu J (2001) Non-hydrolytic disruption of crystalline structure of cellulose by cellulose binding domain and linker sequence of cellobiohydrolase I from Penicillium janthinellum. Acta Biochimica Et Biophysica Sinica 33: 13-18

Gardiner ES, Sarko A (1985) Packing Analysis of Carbohydrates and Polysaccharides .16. the Crystal-Structures of Cellulose-Ivi and Cellulose-Ivii. Canadian Journal of Chemistry-Revue Canadienne De Chimie 63: 173-180

Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Warren RAJ (1991) Domains in Microbial Beta-1,4-Glycanases - Sequence Conservation, Function, and Enzyme Families. Microbiol Rev 55: 303-315

Gilkes N, Warren R, Miller R, Kilburn D (1988) Precise Excision of the Cellulose Binding Domains from 2 Cellulomonas-Fimi Cellulases by a Homologous Protease and the Effect on Catalysis. J Biol Chem 263: 10401-10407

Habibi Y, Lucia LA, Rojas OJ (2010) Cellulose Nanocrystals: Chemistry, Self-Assembly, and Applications. Chem Rev 110: 3479-3500

Hall M, Bansal P, Lee JH, Realff MJ, Bommarius AS (2010) Cellulose crystallinity - a key predictor of the enzymatic hydrolysis rate. Febs Journal 277: 1571-1582

Hancock, S. M. and Withers, S. G. 2007. Glycosidases: Functions, Families and Folds. eLS. .

Harris PV, Welner D, McFarland KC, Re E, Poulsen JN, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L (2010) Stimulation of Lignocellulosic Biomass Hydrolysis by Proteins of Glycoside Hydrolase Family 61: Structure and Function of a Large, Enigmatic Family. Biochemistry (N Y) 49: 3305-3316

Helbert W, Sugiyama J, Ishihara M, Yamanaka S (1997) Characterization of native crystalline cellulose in the cell walls of Oomycota. J Biotechnol 57: 29-37

Hemsworth G R, Henrissat B, Davies G J & Walton P H (2014). Discovery and characterization of a new family of lytic polysaccharide monooxygenases. Nature chemical biology 10(2): 122-126.

Henrissat B, Driguez H, Viet C, Schulein M (1985) Synergism of Cellulases from Trichoderma-Reesei in the Degradation of Cellulose. Bio-Technology 3: 722-726

Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH (2012) Novel enzymes for the degradation of cellulose. Biotechnology for Biofuels 5: 45

Howard S, He S, Withers S (1998) Identification of the active site nucleophile in jack bean alpha-mannosidase using 5-fluoro-beta-L-guloxyl fluoride. J Biol Chem 273: 2067-2072

Irwin DC, Spezio M, Walker LP, Wilson DB (1993) Activity Studies of 8 Purified Cellulases - Specificity, Synergism, and Binding Domain Effects. Biotechnol Bioeng 42: 1002-1013

Jalak J, Kurasin M, Teugjas H, Vaeljamaee P (2012) Endo-exo Synergism in Cellulose Hydrolysis Revisited. J Biol Chem 287: 28802-28815

Jarvis MC (2000) Interconversion of the I alpha and I beta crystalline forms of cellulose by bending. Carbohydr Res 325: 150-154

Jarvis MC (2013) Cellulose Biosynthesis: Counting the Chains. Plant Physiol 163: 1485-1486

Jeoh T, Ishizawa C I, Davis M F, Himmel M E, Adney W S and Johnson D K (2007), Cellulase digestibility of pretreated biomass is limited by cellulose accessibility. Biotechnol Bioeng 98: 112–122.

Karlsson J, Siika-aho M, Tenkanen M, Tjerneld F (2002) Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of Trichoderma reesei. J Biotechnol 99: 63-78

Kim S, Dale B (2005) Life cycle assessment of various cropping systems utilized for producing biofuels: Bioethanol and biodiesel. Biomass & Bioenergy 29: 426-439

Kim Y, Mosier NS, Ladisch MR (2009) Enzymatic Digestion of Liquid Hot Water Pretreated Hybrid Poplar. Biotechnol Prog 25: 340-348

Kless H, Sitrit Y, Chet I, Oppenheim AB (1989) Cloning of the Gene Coding for Chitobiase of Serratia-Marcescens. Molecular & General Genetics 217: 471-473

Koshland D E (1953). Stereochemistry and the mechanism of enzymatic reactions. Biological Reviews, 28(4): 416-436.

Kosugi A, Amano Y, Murashima K, Doi R (2004) Hydrophilic domains of scaffolding protein CbpA promote glycosyl hydrolase activity and localization of cellulosomes to the cell surface of Clostridium cellulovorans. J Bacteriol 186: 6351-6359

Kosugi A, Murashima K, Tamaru Y, Doi R (2002) Cell-surface-anchoring role of N-terminal surface layer homology domains of Clostridium cellulovorans EngE. J Bacteriol 184: 884-888

Koyama M, Helbert W, Imai T, Sugiyama J, Henrissat B (1997) Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. Proc Natl Acad Sci U S A 94: 9091-9095

Kumar A (2009) A conceptual comparison of bioenergy options for using mountain pine beetle infested wood in Western Canada. Bioresour Technol 100: 387-399

Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J Ind Microbiol Biotechnol 35: 377-391

Langan P, Gnanakaran S, Rector KD, Pawley N, Fox DT, Cho DW, Hammel KE (2011) Exploring new strategies for cellulosic biofuels production. Energy & Environmental Science 4: 3820-3833

Larsson P, Wickholm K, Iversen T (1997) A CP/MAS C-13 NMR investigation of molecular ordering in celluloses. Carbohydr Res 302: 19-25

Lee I, Evans B, Woodward J (2000) The mechanism of cellulase action on cotton fibers: evidence from atomic force microscopy. Ultramicroscopy 82: 213-221

Leibovitz E, Beguin P (1996) A new type of cohesin domain that specifically binds the dockerin domain of the Clostridium thermocellum cellulosome-integrating protein CipA. J Bacteriol 178: 3077-3084

Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnology for Biofuels 6: 41

Levine SE, Fox JM, Blanch HW, Clark DS (2010) A Mechanistic Model of the Enzymatic Hydrolysis of Cellulose. Biotechnol Bioeng 107: 37-51

Li D, Li A, Papageorgiou AC (2011) Cellulases from thermophilic fungi: recent insights and biotechnological potential. Enzyme research 2011: 308730-308730

Li X, Beeson IV W T, Phillips C M, Marletta M A & Cate J H (2012). Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. Structure 20(6): 1051-1061.

Lin Z, Zhang H, Ji X, Chen J, Huang H (2011) Hydrolytic Enzyme of Cellulose for Complex Formulation Applied Research. Appl Biochem Biotechnol 164: 23-33

Lynd L, Weimer P, van Zyl W, Pretorius I (2002) Microbial cellulose utilization: Fundamentals and biotechnology. Microbiology and Molecular Biology Reviews 66: 506-+

Mansfield SD, Mooney C, Saddler JN (1999) Substrate and enzyme characteristics that limit cellulose hydrolysis. Biotechnol Prog 15: 804-816

Markus P, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant Journal 54: 559-568

Martinez A, Speranza M, Ruiz-Duenas F, Ferreira P, Camarero S, Guillen F, Martinez M, Gutierrez A, del Rio J (2005) Biodegradation of lignocellulosics: microbial chemical, and enzymatic aspects of the fungal attack of lignin. International Microbiology 8: 195-204

Master ER, Zheng Y, Storms R, Tsang A, Powlowski J (2008) A xyloglucan-specific family 12 glycosyl hydrolase from Aspergillus niger: recombinant expression, purification and characterization. Biochem J 411: 161-170

Meinke A, Damude HG, Tomme P, Kwan E, Kilburn DG, Miller RC, Warren RAJ, Gilkes NR (1995) Enhancement of the Endo-Beta-1,4-Glucanase Activity of an Exocellobiohydrolase by Deletion of a Surface Loop. J Biol Chem 270: 4383-4386

Moser F, Irwin D, Chen S, Wilson DB (2008) Regulation and characterization of Thermobifida fusca carbohydrate-binding module proteins E7 and E8. Biotechnol Bioeng 100: 1066-1077

Mudge SM (2008) Is the use of biofuels environmentally sound or ethical? Journal of Environmental Monitoring 10: 701-702

Nakazawa H, Kawai T, Ida N, Shida Y, Kobayashi Y, Okada H, Tani S, Sumitani J, Kawaguchi T, Morikawa Y, Ogasawara W (2012) Construction of a recombinant Trichoderma reesei strain expressing Aspergillus aculeatus ss-glucosidase 1 for efficient biomass conversion. Biotechnol Bioeng 109: 92-99

Ng I, Li C, Chan S, Chir J, Chen PT, Tong C, Yu S, Ho TD (2010) High-level production of a thermoacidophilic beta-glucosidase from Penicillium citrinum YS40-5 by solid-state fermentation with rice bran. Bioresour Technol 101: 1310-1317

Nidetzky B, Hayn M, Macarron R, Steiner W (1993) Synergism of Trichoderma-Reesei Cellulases while Degrading Different Celluloses. Biotechnol Lett 15: 71-76

Nidetzky B, Steiner W, Hayn M, Claeyssens M (1994) Cellulose Hydrolysis by the Cellulases from Trichoderma-Reesei - a New Model for Synergistic Interaction. Biochem J 298: 705-710

Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc 124: 9074-9082

Nishiyama Y, Sugiyama J, Chanzy H, Langan P (2003) Crystal structure and hydrogen bonding system in cellulose 1(alpha), from synchrotron X-ray and neutron fiber diffraction. J Am Chem Soc 125: 14300-14306

Ouyang J, Yan M, Kong D and Xu L (2006) A complete protein pattern of cellulase and hemicellulase genes in the filamentous fungus *Trichoderma reesei*. Biotechnology Journal 1: 1266–1274.

Pages S, Belaich A, Belaich J, Morag E, Lamed R, Shoham Y, Bayer E (1997) Species-specificity of the cohesin-dockerin interaction between Clostridium thermocellum and Clostridium cellulolyticum: Prediction of specificity determinants of the dockerin domain. Proteins-Structure Function and Genetics 29: 517-527

Pandey MP, Kim CS (2011) Lignin Depolymerization and Conversion: A Review of Thermochemical Methods. Chem Eng Technol 34: 29-41

Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK (2010) Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. Biotechnology for Biofuels 3: 10

Parsiegla G, Juy M, Reverbel-Leroy C, Tardif C, Belaich J, Driguez H, Haser R (1998) The crystal structure of the processive endocellulase CelF of Clostridium cellulolyticum in complex with a thiooligosaccharide inhibitor at 2.0 angstrom resolution. EMBO J 17: 5551-5562

Parthasarathi R, Bellesia G, Chundawat SPS, Dale BE, Langan P, Gnanakaran S (2011) Insights into Hydrogen Bonding and Stacking Interactions in Cellulose. Journal of Physical Chemistry a 115: 14191-14202

Pasqualino JC, Montane D, Salvado J (2006) Synergic effects of biodiesel in the biodegradability of fossil-derived fuels. Biomass & Bioenergy 30: 874-879

Poole DM, Morag E, Lamed R, Bayer EA, Hazlewood GP, Gilbert HJ (1992) Identification of the Cellulose-Binding Domain of the Cellulosome Subunit-S1 from Clostridium-Thermocellum Ys. FEMS Microbiol Lett 99: 181-186

Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JN, Johansen KS, Krogh KBRM, Jorgensen CI, Tovborg M, Anthonsen A, Tryfona T, Walter CP, Dupree P, Xu F, Davies GJ, Walton PH (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proc Natl Acad Sci U S A 108: 15079-15084

Raghothama S, Simpson PJ, Szabo L, Nagy T, Gilbert HJ, Williamson MP (2000) Solution structure of the CBM10 cellulose binding module from Pseudomonas xylanase A. Biochemistry (N Y) 39: 978-984

Raghu S, Anderson RC, Daehler CC, Davis AS, Wiedenmann RN, Simberloff D, Mack RN (2006) Adding biofuels to the invasive species fire? Science 313: 1742-1742

Reese TE (1975) Enzyme systems for cellulose. Biotechnol Bioeng Symp 5:77-80

Rincon M, Ding S, McCrae S, Martin J, Aurilia V, Lamed R, Shoham Y, Bayer E, Flint H (2003) Novel organization and divergent dockerin specificities in the cellulosome system of Ruminococcus flavefaciens. J Bacteriol 185: 703-713

Rincon M, Martin J, Aurilia V, McCrae S, Rucklidge G, Reid M, Bayer E, Lamed R, Flint H (2004) ScaC, an adaptor protein carrying a novel cohesin that expands the dockerin-binding repertoire of the Ruminococcus flavefaciens 17 cellulosome. J Bacteriol 186: 2576-2585

Rouvinen J, Bergfors T, Teeri T, Knowles JKC, Jones TA (1990) 3-Dimensional Structure of Cellobiohydrolase-Ii from Trichoderma-Reesei. Science 249: 380-386

Sakamoto R, Arai M, Murao S (1985) Enzymic Properties of 3 Beta-Glucosidases from Aspergillus-Aculeatus No-F-50. Agric Biol Chem 49: 1283-1290

Sakon J, Irwin D, Wilson D, Karplus P (1997) Structure and mechanism of endo/exocellulase E4 from Thermomonospora fusca. Nat Struct Biol 4: 810-818

Salamitou S, Tokatlidis K, Beguin P, Aubert JP (1992) Involvement of Separate Domains of the Cellulosomal Protein-S1 of Clostridium-Thermocellum in Binding to Cellulose and in Anchoring of Catalytic Subunits to the Cellulosome. FEBS Lett 304: 89-92

Sammond DW, Payne CM, Brunecky R, Himmel ME, Crowley MF, Beckham GT (2012) Cellulase Linkers Are Optimized Based on Domain Type and Function: Insights from Sequence Analysis, Biophysical Measurements, and Molecular Simulation. Plos One 7: e48615

Sanchez B, Gonzalez-Tejedo C, Ruas-Madiedo P, Urdaci MC, Margolles A (2011) Lactobacillus plantarum Extracellular Chitin-Binding Protein and Its Role in the Interaction between Chitin, Caco-2 Cells, and Mucin. Appl Environ Microbiol 77: 1123-1126

Schulein M (2000) Protein engineering of cellulases. Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology 1543: 239-252

Shen H, Schmuck M, Pilz I, Gilkes N, Kilburn D, Miller R, Warren R (1991) Deletion of the Linker Connecting the Catalytic and Cellulose-Binding Domains of Endoglucanase-a (Cena) of Cellulomonas-Fimi Alters its Conformation and Catalytic Activity. J Biol Chem 266: 11335-11340

Sierra R, Smith A, Granda C, Holtzapple MT (2008) Producing fuels and chemicals from lignocellulosic biomass. Chem Eng Prog 104: S10-S18

Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, Vorwerk S, Youngs H (2004) Toward a systems approach to understanding plant-cell walls. Science 306: 2206-2211

Srisodsuk M, Reinikainen T, Penttila M, Teeri T (1993) Role of the Interdomain Linker Peptide of Trichoderma-Reesei Cellobiohydrolase-i in its Interaction with Crystalline Cellulose. J Biol Chem 268: 20756-20761

Stahlberg J, Divne C, Koivula A, Piens K, Claeyssens M, Teeri T, Jones T (1996) Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from Trichoderma reesei. J Mol Biol 264: 337-349

Stalbrand H, Mansfield S, Saddler J, Kilburn D, Warren R, Gilkes N (1998) Analysis of molecular size distributions of cellulose molecules during hydrolysis of cellulose by recombinant Cellulomonas fimi beta-1,4-glucanases. Appl Environ Microbiol 64: 2374-2379

Storms R, Zheng Y, Li H, Sillaots S, Martinez-Perez A, Tsang A (2005) Plasmid vectors for protein production, gene expression and molecular manipulations in Aspergillus niger. Plasmid. 53: 191-204.

Sturcova A, Davies GR, Eichhorn SJ (2005) Elastic modulus and stress-transfer properties of tunicate cellulose whiskers. Biomacromolecules 6: 1055-1061

Sugiyama J, Okano T, Yamamoto H, Horii F (1990) Transformation of Valonia Cellulose Crystals by an Alkaline Hydrothermal Treatment. Macromolecules 23: 3196-3198

Sun X, Sun R, Fowler P, Baird M (2005) Extraction and characterization of original lignin and hemicelluloses from wheat straw. J Agric Food Chem 53: 860-870

Takahashi N, and Koshijima T (1988) Ester Linkages between Lignin and Glucuronoxylan in a Lignin-Carbohydrate Complex from Beech (Fagus crenata) Woods. Wood Sci. Technol 22:231.

Takashima S, Nakamura A, Hidaka M, Masaki H, Uozumi T (1999) Molecular cloning and expression of the novel fungal beta-glucosidase genes from Humicola grisea and Trichoderma reesei. J Biochem 125: 728-736

Tambor JH, Ren H, Ushinsky S, Zheng Y, Riemens A, St-Francois C, Tsang A, Powlowski J, Storms R (2012) Recombinant expression, activity screening and functional characterization identifies three novel endo-1,4-beta-glucanases that efficiently hydrolyse cellulosic substrates. Appl Microbiol Biotechnol 93: 203-214

Teeri T (1997) Crystalline cellulose degradation: New insight into the function of cellobiohydrolases. Trends Biotechnol 15: 160-167

Tenenbaum DJ (2008) Food vs. fuel: Diversion of crops could cause more hunger. Environ Health Perspect 116: A254-A257

Terwisscha Van Scheltinga AC, Armand S, Kalk KH, Isogai A, Henrissat B, Dijkstra BW (1995) Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and x-ray structure of a complex with allosamidin: Evidence for substrate assisted catalysis. Biochemistry (N Y) 34: 15619-15623

Tews Vvo AP, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nat Struct Biol 3: 638-648

Tokatlidis K, Salamitou S, Beguin P, Dhurjati P, Aubert JP (1991) Interaction of the Duplicated Segment Carried by Clostridium-Thermocellum Cellulases with Cellulosome Components. FEBS Lett 291: 185-188

Tomme P, Vantilbeurgh H, Pettersson G, Vandamme J, Vandekerckhove J, Knowles J, Teeri T, Claeyssens M (1988) Studies of the Cellulolytic System of Trichoderma-Reesei Qm-9414 - Analysis of Domain Function in 2 Cellobiohydrolases by Limited Proteolysis. European Journal of Biochemistry 170: 575-581

Vaaje-Kolstad G, Horn S, van Aalten D, Synstad B, Eijsink V (2005) The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. J Biol Chem 280: 28492-28497

Vaaje-Kolstad G, Houston D, Riemen A, Eijsink V, van Aalten D (2005) Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. J Biol Chem 280: 11313-11319

Vaaje-Kolstad G, Westereng B, Horn S J, Liu, Z, Zhai H, Sørlie, M, & Eijsink V G (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science 330 (6001): 219-2

Vaaje-Kolstad G, Bøhle L A, Gåseidnes S, Dalhus B, Bjørås M, Mathiesen G & Eijsink V G (2012) Characterization of the Chitinolytic Machinery of< i> Enterococcus faecalis</i> V583 and High-Resolution Structure of Its Oxidative CBM33 Enzyme. Journal of molecular biology 416(2): 239-254.

Valjamae P, Sild V, Nutt A, Pettersson G, Johansson G (1999) Acid hydrolosis of bacterial cellulose reveals different modes of synergistic action between cellobiohydrolase I and endoglucanase I. European Journal of Biochemistry 266: 327-334

van Zyl WH, Lynd LR, den Haan R, McBride JE (2007) Consolidated bioprocessing for bioethanol production using Saccharomyces cereviside. Biofuels 108: 205-235

Varnai A, Siika-aho M, Viikari L (2013) Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. Biotechnology for Biofuels 6: 30

Varrot A, Schulein M, Davies G (1999) Structural changes of the active site tunnel of Humicola insolens cellobiohydrolase, Cel6A, upon oligosaccharide binding. Biochemistry (N Y) 38: 8884-8891

Wada M, Chanzy H, Nishiyama Y, Langan P (2004) Cellulose IIII crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. Macromolecules 37: 8548-8555

Wang Y, Zhao Y, & Deng Y (2008). Effect of enzymatic treatment on cotton fiber dissolution in NaOH/urea solution at cold temperature. Carbohydrate polymers, 72(1): 178-184.

Watts A, Damager I, Amaya M, Buschiazzo A, Alzari P, Frasch A, Withers S (2003) Trypanosoma cruzi trans-sialidase operates through a covalent sialyl-enzyme intermediate: Tyrosine is the catalytic nucleophile. J Am Chem Soc 125: 7532-7533

Wilson DB (2004) Studies of Thermobifida fusca plant cell wall degrading enzymes. Chemical Record 4: 72-82

Withers S (2001) Mechanisms of glycosyl transferases and hydrolases. Carbohydr Polym 44: 325-337

Withers S (1999) 1998 Hoffmann la Roche Award Lecture - Understanding and exploiting glycosidases. Canadian Journal of Chemistry-Revue Canadienne De Chimie 77: 1-11

Woodward J (1991) Synergism in Cellulase Systems. Bioresour Technol 36: 67-75

Zhang YHP, Lynd LR (2005) Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. Biomacromolecules 6: 1510-1515

Zhang YHP, Lynd LR (2004) Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. Biotechnol Bioeng 88: 797-824

Zhang YHP, Himmel ME, Mielenz JR (2006) Outlook for cellulase improvement: Screening and selection strategies. Biotechnol Adv 24: 452-481

Zhang Y, Cui J, Lynd L, Kuang L (2006) A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7: 644-648

Zhang YP, Ding S, Mielenz JR, Cui J, Elander RT, Laser M, Himmel ME, McMillan JR, Lynd LR (2007) Fractionating recalcitrant lignocellulose at modest reaction conditions. Biotechnol Bioeng 97: 214-223

Zou J, Kleywegt G, Stahlberg J, Driguez H, Nerinckx W, Claeyssens M, Koivula A, Teerii T, Jones T (1999) Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Cel6A from Trichoderma reesei. Structure Fold Des 7: 1035-1045

Zuber M, Zia KM, Bhatti IA, Ali Z, Arshad MU, Saif MJ (2012) Modification of cellulosic fibers by UV-irradiation. Part II: After treatments effects. Int J Biol Macromol 51: 743-748

Appendix 1A

Protein sequences of recombinant cellobiohydrolases

>StheSEQ437

```
MYAKFATLAALVAGAAAQNACTLTAENHPSLTWSKCTSGGSCTSVQGSITIDANWRWTHR
TDSATNCYEGNKWDTSYCSDGPSCASKCCIDGADYSSTYGITTSGNSLNLKFVTKGQYST
NIGSRTYLMESDTKYQMFQLLGNEFTFDVDVSNLGCGLNGALYFVSMDADGGMSKYSGNK
AGAKYGTGYCDSQCPRDLKFINGEANVENWQSSTNDANAGTGKYGSCCSEMDVWEANNMA
AAFTPHPCTVIGQSRCEGDSCGGTYSTDRYAGICDPDGCDFNSYRQGNKTFYGKGMTVDT
TKKITVVTQFLKNSAGELSEIKRFYVQNGKVIPNSESTIPGVEGNSITQDWCDRQKAAFG
DVTDFQDKGGMVQMGKALAGPMVLVMSIWDDHAVNMLWLDSTWPIDGAGKPGAERGACPT
TSGVPAEVEAEAPNSNVIFSNIRFGPIGSTVSGLPDGGSGNPNPPVSSSTPVPSSSTTSS
GSSGPTGGTGVAKHYEQCGGIGFTGPTQCESPYTCTKLNDWYSQCL
```

>Afu6g11610

```
MLASTFSYRMYKTALILAALLGSGQAQQVGTSQAEVHPSMTWQSCTAGGSCTTNNGKVVI
DANWRWVHKVGDYTNCYTGNTWDTTICPDDATCASNCALEGANYESTYGVTASGNSLRLN
FVTTSQQKNIGSRLYMMKDDSTYEMFKLLNQEFTFDVDVSNLPCGLNGALYFVAMDADGG
MSKYPTNKAGAKYGTGYCDSQCPRDLKFINGQANVEGWQPSSNDANAGTGNHGSCCAEMD
IWEANSISTAFTPHPCDTPGQVMCTGDACGGTYSSDRYGGTCDPDGCDFNSFRQGNKTFY
GPGMTVDTKSKFTVVTQFITDDGTSSGTLKEIKRFYVQNGKVIPNSESTWTGVSGNSITT
EYCTAQKSLFQDQNVFEKHGGLEGMGAALAQGMVLVMSLWDDHSANMLWLDSNYPTTASS
TTPGVARGTCDISSGVPADVEANHPDAYVVYSNIKVGPIGSTFNSGGSNPGGGTTTTTTT
QPTTTTTTAGNPGGTGVAQHYGQCGGIGWTGPTTCASPYTCQKLNDYYSQCL
```

>389-Seq274NdisCbhGH6A

```
MTGSLAAQASAVAKVPTFQWLDRNVTVDTLMKSTLEEIRAANKAGANPPYAAHFVVYDLP
DRDCAAAASNGEFSIANNGVANYKTYINAIRKLLVQYSDIRTILVIEPDSLANLVTNTNV
AKCANAASAYKECTNYAITQLDLPHVAQYLDAGHGGWLGWPANIQPAATLFADIYKAAGK
PKSVRGLVTNVSNYNGWSLSSAPSYTTPNPNYDEKKYIEAFSPLLNAAGFPAQFIVDTGR
SGKQPTGQIEQGDWCNAIGTGFGVRPTTDTGSSLADAFVWVKPGGESDGTSDTSATRYDY
HCGLSDALKPAPEAGQWFQAYFEQLLKNANPAF
```

>393-Seq285AoryCbhGH7A

```
MTWQTCGNGGSCSTNQGSVVLDANWRWVHQTGSSSNCYTGNKWDTSYCSTNDACAQKCAL
DGADYSNTYGITTSGSEVRLNFVTSNSNGKNVGSRVYMMADDTHYEVYKLLNQEFTFDVD
VSKLPCGLNGALYFVVMDADGGVSKYPNNKAGAKYGTGYCDSQCPRDLKFIQGQANVEGW
VSSTNNANTGTGNHGSCCAELDIWESNSISQALTPHPCDTPTNTLCTGDACGGTYSSDRY
SGTCDPDGCDFNPYRVGNTTFYGPGKTIDTNKPITVVTQFITDDGTSSGTLSEIKRFYVQ
DGVTYPQPSADVSGLSGNTINSEYCTAENTLFEGSGSFAKHGGLAGMGEAMSTGMVLVMS
LWDDYYANMLWLDSNYPTNESTSKPGVARGTCSTSSGVPSEVEASNPSAYVAYSNIKVGP
IGSTFKS
```

>396-Seq270NdisCbhGH7A

```
MCEGDSCGGTYSDDRYGGTCDADGCDFNSYRMGNTTFYGEGKTVDTSSKFTVVTQFIKDS
AGDLAEIKRFYVQNGKVIENSQSNIDGVSGNSITQSFCNAQKTAFGDIDDFSKKGGLKQM
GKALAKPMVLVMSIWDDHAANMLWLDSTYPVEGGNPGAYRGECPTTSGVPADVEANAPNS
KVVFSNIKFGPIGSTFSGGGSGTTPSNPSSSVKPVTSTAKPSSTSTASNPSGTGAAHWAQ
CGGIGFSGPTTCQSPYTCQKINEYYSQCV
```

>400-Seq275TterCbhGH7A

```
MAPKSTVLAAWLLSSLAAAQQIGKAVPEVHPKLTTQKCTLRGGCKPVRTSVVLDSSARSL
HKVGDPNTSCSVGGDLCSDAKSCGKNCALEGVDYAAHGVATKGDALTLHQWLKGADGTYR
TVSPRVYLLGEDGKNYEDFKLLNAELSFDVDVSQLVCGMNGALYFSEMEMDGGRSPLNPA
GATYGTGYCDAQCPKLDFINGEVFLLSSVFLFHRFF
```

> Outergroup Bacterial protein `Endoglucanase F [Ruminiclostridium thermocellum BC1]`

KFVLTVLLFFLTAVLLPSSSFESKVQAASSPRYGGAYYNYGEALQKAILFYKANRLGDL
PDDYILPYRADAAMTDGQDVGLDLTGGWADAGDGIKFTHPMSYAAGQLGWAVYEYRQAFEKAGLLDDILD
EIKWATDFFIKAHPEPNVLYYMCGYNDSDHSVWVPHELLDYVTDRKSFVLNPSTPGSDVAGQTAACLAIA
SIIFEPTDPEYAETCLTHAKQIFEFGDKYRGKNPLDVLYPSGGYLDDLAWGAVWLYIKTGDSTYLEKAKS
FLPVTSLGGGHTHCWDDVSYGAALKIAQLTHDEGYAAMVEKNLDFWLPGGGITYTPGGLAWLSPWGSLRY
ASTAAFLAFVWADDPTVGTPSKKETYRAFAEKQINYILGDNPRKGSYVVGFGENSPKHPHHRTAHGSWVS
MLEVPSFHRHILYGALVGGPSSDDSWEDDISDYTRNEVATDYNAGFVGALAKMYDMYGGEPLENWPQPED
FRAPEDDIVEYFCRGWIIYEGYGTLNLLLQVNNRSGWPPTMKDKLSVRYFMDLTEVFESGGTVDDVQISL
GQNEGAKLIGLKHYRDNIYYFTVDFTGTMIMPAEWEMCEKDAHVTIKYRDGITGSNENDWSYQNLRKDPD
YDATSFAGLTPYIPVYDNGVLLWGEEPPAGGDDPGSSPPPTPTEPVIVYGDLNGDGNINSTDFTMLKRAI
LGNPAPGTNLAAGDLNRDGNT NSTDLMILRRYLLKLIGSLPI

Appendix 1B

Gene and primer sequences of recombinant cellobiohydrolases

>StheSEQ437 orf

ATGTACGCCAAGTTCGCGACCCTCGCCGCCCTTGTGGCTGGCGCCGCTGCTCAGAACGCCTGCACTCTGACCGC
TGAGAACCACCCCTCGCTGACGTGGTCCAAGTGCACGTCTGGCGGCAGCTGCACCAGCGTCCAGGGTTCCATCA
CCATCGACGCCAACTGGCGGTGGACTCACCGGACCGATAGCGCCACCAACTGCTACGAGGGCAACAAGTGGGAT
ACTTCGTACTGCAGCGATGGTCCTTCTTGCGCCTCCAAGTGCTGCATCGACGGCGCTGACTACTCGAGCACCTA
TGGCATCACCACGAGCGGTAACTCCCTGAACCTCAAGTTCGTCACCAAGGGCCAGTACTCGACCAACATCGGCT
CGCGTACCTACCTGATGGAGAGCGACACCAAGTACCAGATGTTCCAGCTCCTCGGCAACGAGTTCACCTTCGAT
GTCGACGTCTCCAACCTCGGCTGCGGCCTCAATGGCGCCCTCTACTTCGTGTCCATGGATGCCGATGGTGGCAT
GTCCAAGTACTCGGGCAACAAGGCAGGTGCCAAGTACGGTACCGGCTACTGTGATTCTCAGTGCCCCCGCGACC
TCAAGTTCATCAACGGCGAGGCCAACGTAGAGAACTGGCAGAGCTCGACCAACGATGCCAACGCCGGCACGGGC
AAGTACGGCAGCTGCTGCTCCGAGATGGACGTCTGGGAGGCCAACAACATGGCCGCCGCCTTCACTCCCCACCC
TTGCACCGTGATCGGCCAGTCGCGCTGCGAGGGCGACTCGTGCGGCGGTACCTACAGCACCGACCGCTATGCCG
GCATCTGCGACCCCGACGGATGCGACTTCAACTCGTACCGCCAGGGCAACAAGACCTTCTACGGCAAGGGCATG
ACGGTCGACACGACCAAGAAGATCACGGTCGTCACCCAGTTCCTCAAGAACTCGGCCGGCGAGCTCTCCGAGAT
CAAGCGGTTCTACGTCCAGAACGGCAAGGTCATCCCCAACTCCGAGTCCACCATCCCGGGCGTCGAGGGCAACT
CCATCACCCAGGACTGGTGCGACCGCCAGAAGGCCGCCTTCGGCGACGTGACCGACTTCCAGGACAAGGGCGGC
ATGGTCCAGATGGGCAAGGCCCTCGCGGGGCCCATGGTCCTCGTCATGTCCATCTGGGACGACCACGCCGTCAA
CATGCTCTGGCTCGACTCCACCTGGCCCATCGACGGCGCCGGCAAGCCGGGCGCCGAGCGCGGTGCCTGCCCCA
CCACCTCGGGCGTCCCCGCTGAGGTCGAGGCCGAGGCCCCCAACTCCAACGTCATCTTCTCCAACATCCGCTTC
GGCCCCATCGGCTCCACCGTCTCCGGCCTGCCCGACGGCGGCAGCGGCAACCCCAACCCGCCCGTCAGCTCGTC
CACCCCGGTCCCCTCCTCGTCCACCACATCCTCCGGTTCCTCCGGCCCGACTGGCGGCACGGGTGTCGCTAAGC
ACTATGAGCAATGCGGAGGAATCGGGTTCACTGGCCCTACCCAGTGCGAGAGCCCCTACACTTGCACCAAGCTG
AATGACTGGTACTCGCAGTGCCTGTAA

Cloned using Gateway system

>Afu6g11610

Fprimer: Afu6g11610:cbh1: F: NheI (29 ntp, Tm 55, 8):
AGAATGCTAGCATAATGCTGGCCTCCACC

Rprimer: Afu6g11610:cbh1: R: FseI (38 ntp, Tm 54, 0):
AGACTAGGCCGGCCCTACAGGCACTGAGAGTAATAATC

Gene sequence:

ATGCTGGCCTCCACCTTCTCCTACCGCATGTACAAGACCGCGCTCATCCTGGCCGCCCTTCTGGGCTCTGGCCA
GGCTCAGCAGGTCGGTACTTCCCAGGCGGAAGTGCATCCGTCCATGACCTGGCAGAGCTGCACGGCTGGCGGCA
GCTGCACCACCAACAACGGCAAGGTGGTCATCGACGCGAACTGGCGTTGGGTGCACAAAGTCGGCGACTACACC
AACTGCTACACCGGCAACACCTGGGACACGACTATCTGCCCTGACGATGCGACCTGCGCATCCAACTGCGCCCT
TGAGGGTGCCAACTACGAATCCACCTATGGTGTGACCGCCAGCGGCAATTCCCTCCGCCTCAACTTCGTCACCA
CCAGCCAGCAGAAGAACATTGGCTCGCGTCTGTACATGATGAAGGACGACTCGACCTACGAGATGTTTAAGCTG
CTGAACCAGGAGTTCACCTTCGATGTCGATGTCTCCAACCTCCCCTGCGGTCTCAACGGTGCTCTGTACTTTGT
CGCCATGGACGCCGACGGTGGCATGTCCAAGTACCCAACCAACAAGGCCGGTGCCAAGTACGGTACTGGATACT
GTGACTCGCAGTGCCCTCGCGACCTCAAGTTCATCAACGGTCAGGCCAACGTTGAAGGGTGGCAGCCCTCCTCC
AACGATGCCAATGCGGGTACCGGCAACCACGGGTCCTGCTGCGCGGAGATGGATATCTGGGAGGCCAACAGCAT
CTCCACGGCCTTCACCCCCCATCCGTGCGACACGCCCGGCCAGGTGATGTGCACCGGTGATGCCTGCGGTGGCA
CCTACAGCTCCGACCGCTACGGCGGCACCTGCGACCCCGACGGATGTGATTTCAACTCCTTCCGCCAGGGCAAC
AAGACCTTCTACGGCCCTGGCATGACCGTCGACACCAAGAGCAAGTTTACCGTCGTCACCCAGTTCATCACCGA
CGACGGCACCTCCAGCGGCACCCTCAAGGAGATCAAGCGCTTCTACGTGCAGAACGGCAAGGTGATCCCCAACT
CGGAGTCGACCTGGACCGGCGTCAGCGGCAACTCCATCACCACCGAGTACTGCACCGCCCAGAAAAGCCTGTTC
CAGGACCAGAACGTCTTCGAAAAGCACGGCGGCCTCGAGGGCATGGGTGCTGCCCTCGCCCAGGGCATGGTTCT
CGTCATGTCCCTGTGGGATGATCACTCGGCCAACATGCTCTGGCTCGACAGCAACTACCCGACCACTGCCTCTT
CCACCACTCCCGGCGTCGCCCGTGGTACCTGCGACATCTCCTCCGGCGTCCCTGCGGATGTCGAGGCGAACCAC
CCCGACGCCTACGTCGTCTACTCCAACATCAAGGTCGGCCCCATCGGCTCGACCTTCAACAGCGGTGGCTCGAA
CCCCGGTGGCGGAACCACCACGACAACTACCACCCAGCCTACTACCACCACGACCACGGCTGGAAACCCTGGCG
GCACCGGAGTCGCACAGCACTATGGCCAGTGTGGTGGAATCGGATGGACCGGACCCACAACCTGTGCCAGCCCT
TATACCTGCCAGAAGCTGAATGATTATTACTCTCAGTGCCTGTAG

>389- Seq- 274 NdisCbhGH6A-1

Primers:

274 NdisCbhGH6A: F: SpeI AGAATACTAGTATGGC TGCCAAGAAGCTC

274 NdisCbhGH6A: R: FseI AGACTAGGCCGGCCTTAGAAAGCGGGGTTGG

Gene sequence:

ATGGCTGCCAAGAAGCTCCTTCTCGCTGCCGCCTTGACGGCCTCTGCCCTCGCCGCTCCCGTTCTTGAGGATCG
TCAGAACTGCGGCTCTACTTGGTATGATCACCATCACCCTCAAAACTTGCTTGTTCACTAATTCGAACAAATAA
TATAGGACCCAATGCGGTGGTATCGGCTGGTCTGGCGCGACTTGCTGCTCATCCGGCAACTCCTGCGTTGAGAT
TAACTCCTACTACTCCCAGTGCTTGCCCGGCTCTCAAGTTACCACCACTGCTGCGGCTTCTTCCACCAGCCCTA
CCAGCGCCAGCAAGGTCTCCAGCACCACCAGCAGGGCTACGAGCAGCGGCGTCGCCCAGCCCATCACCACTACT
ACCGCTCCTTCGGTGCCCACCACCACCATTGCTGGCGGTGCTTCCACCACTGCTAGCTTCACTGGCAACCCCTT
CCTGGGTGTTCAGGGCTGGGCCAACAGCTACTACTCGTCCGAGATCTACAACCATGCCATTCCTTCCATGACTG
GCAGCTTGGCTGCTCAGGCGTCTGCCGTTGCCAAGGTTCCCACCTTCCAGTGGCTCGACCGCAACGTCACCGTA
GACACTCTGATGAAGAGCACCCTCGAGGAGATCCGCGCGGCCAACAAGGCCGGTGCCAACCCTCCCTACGCCGC
TCACTTTGTCGTCTACGATCTCCCCGACCGTGACTGCGCTGCTGCTGCCTCCAACGGCGAGTTCTCCATCGCCA
ACAACGGCGTTGCCAACTACAAGACCTACATCAACGCCATCCGCAAGCTCCTGGTCCAGTACTCTGACATCCGT
ACTATCCTCGTCATTGAACCCGACTCGCTTGCCAACCTCGTCACCAACACCAACGTTGCGAAGTGTGCCAACGC
TGCCTCCGCCTACAAGGAGTGCACCAACTACGCCATCACCCAGCTCGACCTTCCCCACGTCGCCCAGTACCTCG
ATGCTGGTCACGGCGGCTGGCTTGGCTGGCCCGCTAACATCCAGCCCGCTGCCACCCTCTTCGCAGACATCTAC
AAGGCCGCCGGCAAGCCCAAATCCGTCCGTGGTCTCGTCACGAACGTCTCCAACTACAACGGCTGGTCGCTCTC
CTCCGCTCCCTCGTACACCACCCCCAACCCCAACTATGACGAGAAGAAATACATTGAGGCTTTCTCCCCTCTTC
TCAACGCTGCCGGCTTCCCCGCACAGTTCATCGTCGATACCGGCCGTTCCGGCAAGCAGCCCACCGGCCAGATC
GAGCAGGGTGACTGGTGCAACGCCATCGGCACCGGTTTTGGTGTCCGCCCAACCACCGACACTGGCTCCTCCTT
GGCTGATGCCTTCGTCTGGGTCAAGCCCGGTGGTGAGTCCGATGGTACTAGCGACACCTCTGCTACCCGTTATG
ACTACCACTGCGGTCTCTCGGATGCCCTCAAGCCTGCTCCTGAGGCTGGCCAGTGGTTCCAGGCTTACTTTGAG
CAGCTGCTCAAGAACGCCAACCCCGCTTTCTAA

>393-Seq- 285 AoryCbhGH7A

Primers:

285 AoryCbhGH7A: F: NheI AGAATGCTAGCATAATGGCTTCCCTTTCCCTCTC

285 AoryCbhGH7A: R: FseI AGACTAGGCCGGCCTCAGCTCTTGAAGGTGGAGC

Gene sequence:

ATGGCTTCCCTTTCCCTCTCCAAGATCTGCCGCAATGCCCTCATCTTGTCCTCAGTTCTGTCAACTGCCCAGGG
TCAGCAGGTTGGGACTTACCAGACCGAACCCATCCCTCGATGACCTGGCAGACATGCGGTAACGGCGGTAGTTG
CAGCACCAACCAAGGCTCCGTTGTCCTCGATGCCAACTGGCGTTGGGTCCACCAAACTGGAAGCTCTAGCAATT
GTTACACCGGCAACAAGTGGGATACTTCCTACTGCAGTACCAATGACGCATGTGCCCAGAAATGTGCCCTGGAT
GGTGCCGACTACTCAAACACTTACGGCATCACCACCAGCGGCAGTGAGGTCCGTCTCAACTTTGTCACCAGCAA
CTCCAACGGAAAGAACGTCGGTTCCCGTGTCTACATGATGGCCGACGATACCCATTACGAGGTTTACAAGCTGC
TGAACCAAGAGTTTACCTTTGATGTGGACGTCTCCAAGCTCCCATGTGGTCTTAATGGCGCTTTGTACTTCGTC
GTAATGGATGCCGATGGAGGGGTTTCCAAATATCCAAACAACAAAGCGGGTGCTAAGTACGGTACTGGTTACTG
TGACTCTCAATGCCCACGGGATCTCAAATTTATCCAGGGACAGGCCAATGTCGAAGGCTGGGTATCGTCAACCA
ACAATGCCAATACAGGCACTGGAAACCATGGCTCCTGCTGCGCGGAGCTGGACATATGGGAGTCCAACAGCATT
TCGCAGGCTCTCACTCCTCACCCATGCGATACGCCGACCAACACTCTATGCACTGGGGATGCCTGCGGTGGCAC
ATATAGTTCCGATCGCTACAGTGGCACGTGTGATCCTGACGGTTGCGACTTCAACCCATATCGTGTGGGCAACA
CCACCTTCTACGGCCCTGGCAAGACAATTGATACCAACAAACCGATCACAGTTGTGACCCAGTTCATCACCGAC
GACGGCACTTCCAGCGGCACCCTATCCGAGATTAAACGGTTCTACGTGCAAGATGGCGTCACGTACCCCCAGCC
CAGCGCAGACGTTAGCGGTCTCAGCGGCAACACCATCAACTCTGAGTATTGCACCGCGGAGAATACCCTGTTCG
AAGGCTCGGGCAGCTTCGCGAAACATGGTGGGCTCGCCGGTATGGGTGAAGCCATGTCGACTGGCATGGTGCTG
GTTATGAGCTTGTGGGATGATTACTATGCCAACATGCTCTGGCTTGACAGCAACTACCCCACCAACGAGTCCAC
GAGCAAGCCCGGCGTGGCCCGTGGAACTTGTTCCACGTCGTCTGGCGTTCCTAGCGAGGTCGAAGCCTCCAACC
CTAGCGCCTATGTGGCCTATTCCAACATTAAAGTTGGCCCTATCGGCTCCACCTTCAAGAGCTGA

396-Seq 270 NdisCbhGH7A

Primers:

270 NdisCbhGH7A: F: NheI AGAATGCTAGCATAATGCTCGCCAAGTTCGCTG

270 NdisCbhGH7A: R: FSEI AGACTAGGCCGGCCTTACACGCACTGGGAGTAGTATTCGTTG

Gene sequence:

ATGCTCGCCAAGTTCGCTGCCCTCGCGGCCCTTGTGGCCTCTGCCAACGCCCAGGCTGTTTGCTCTCTCACTAC
CGAGAGCCACCCCTCCCTCAACTGGTCCAAGTGCACTTCTTCCGGATGCACCAACGTCGCCGGATCCGTCACTA
TTGATGCCAATTGGCGCTGGACTCACATCCTTTCCGGCAGCACCAACTGCTACAGCGGCAACAAGTGGGATACC
TCTCTCTGCAGCACCAACACCGACTGCGCTACCAAGTGCTGCGTTGATGGTGCTGATTACACTTCTACCTATGG
TATTCAGGCCAGCGGCAACTCGCTCAGCCTTCAATTTGTCACGAAGGGCCAGTACTCAACCAACATTGGTTCCC
GTACTTACCTTATGAACGGTGCCGATGCCTACCAGGGTTCGTGCTCATCTTTCCCCATTCCTTGAAGACCTAAG
GGCTAACACACATTCCATCTAGGCTTCGAGCTCCTTGGCAACGAGTTCACATTCGATGTCGATGTGTCCGGCAC
TGGCTGTGGTCTCAACGGCGCCCTCTACTTCGTCTCCATGGATCTTGATGGTGGCAAAGCCAAGTACACCAACA
ACAAGGCTGGTGCTAAGTACGGCACCGGTTATTGCGACGCTCAGTGCCCCCGTGATCTCAAGTACATCAACGGT
ATTGCCAACGTTGAGGGCTGGACCCCCTCCACCAACGATGCTAACGCTGGTATTGGTGATCACGGTACTTGCTG
CTCTGAGATGGATATCTGGTTCGTTTACCCATCCCTCTCCCACCGTTAGCATCACAGGCAACTAACACCCACCC
AAGGGAAGCGAACAAAGTCTCTACAGCGTTCACCCCGCACCCCTGCACCACCATCGAACAGCACATGTGCGAGG
GTGACTCGTGCGGTGGTACCTACTCCGACGACCGCTATGGCGGTACCTGCGATGCCGATGGTTGCGACTTCAAC
AGCTACCGTATGGGCAACACTACCTTCTACGGCGAGGGCAAGACTGTCGACACCAGCTCCAAGTTCACCGTCGT
CACCCAGTTCATCAAGGACTCCGCCGGCGATCTTGCTGAGATCAAGCGCTTCTACGTCCAGAACGGCAAGGTCA
TTGAGAATTCTCAGTCCAACATTGATGGTGTCTCCGGCAACTCCATCACCCAGTCTTTCTGCAATGCTCAGAAG
ACTGCTTTCGGCGATATCGATGACTTTAGCAAGAAGGGTGGCCTGAAGCAGATGGGCAAGGCCCTTGCCAAGCC
CATGGTCCTCGTCATGTCCATCTGGGACGACCATGCCGCCAACATGCTCTGGCTCGACTCCACCTACCCTGTCG
AGGGCGGCAACCCCGGTGCTTACCGTGGCGAGTGCCCTACCACCTCGGGTGTCCCCGCCGATGTCGAGGCCAAC
GCTCCCAACTCCAAGGTCGTCTTCTCCAACATCAAGTTCGGCCCCATTGGCTCTACCTTTAGCGGCGGCGGCTC
CGGCACCACTCCGTCCAACCCTTCGAGCTCCGTCAAGCCCGTTACCTCCACGGCTAAGCCTTCTTCCACCTCGA
CTGCCTCCAACCCCAGCGGTACCGGTGCTGCTCACTGGGCCCAGTGCGGTGGCATCGGCTTCTCTGGCCCCACC
ACTTGCCAGAGCCCTTACACTTGCCAGAAGATCAACGAATACTACTCCCAGTGCGTGTAA

>400-Seq 275 TterCbhGH7A

Primers:

275 TvirCbhGH7A: F: NheI AGAATGCTAGCATAATGTATCAGAAATTGGCCGCCATC

275 TvirCbhGH7A: R: FseI AGACTAGGCCGGCCTTACAGACACTGAGAATAGTATGGGTTCAGAACC

Gene sequence:

ATGGCGCCCAAGTCTACAGTTCTGGCCGCCTGGCTGCTCTCCTCGCTGGCCGCGGCCCAGCAGATCGGCAAAGC
CGTGCCCGAGGTCCACCCCAAACTGACAACGCAGAAGTGCACTCTCCGCGGCGGGTGCAAGCCTGTCCGCACCT

CGGTCGTGCTCGACTCGTCCGCGCGCTCGCTGCACAAGGTCGGGGACCCCAACACCAGCTGCAGCGTCGGCGGC
GACCTGTGCTCGGACGCGAAGTCGTGCGGCAAGAACTGCGCGCTCGAGGGCGTCGACTACGCGGCCCACGGCGT
GGCGACCAAGGGCGACGCCCTCACGCTGCACCAGTGGCTCAAGGGGGCCGACGGCACCTACAGGACCGTCTCGC
CGCGCGTATACCTCCTGGGCGAGGACGGGAAGAACTACGAGGACTTCAAGCTGCTCAACGCCGAGCTCAGCTTC
GACGTCGACGTGTCCCAGCTCGTCTGCGGCATGAACGGCGCCCTGTACTTCTCCGAGATGGAGATGGACGGCGG
CCGCAGCCCGCTGAACCCGGCGGGCGCCACGTACGGCACGGGCTACTGCGACGCGCAGTGCCCCAAGTTGGACT
TTATCAACGGCGAGGTATTTCTTCTCTCTTCTGTTTTTCTTTTCCATCGCTTTTTCTGACCGGAATCCGCCCTC
TTAGCTCAACACCAACCACACGTACGGGGCGTGCTGCAACGAGATGGACATCTGGGAGGCCAACGCGCTGGCGC
AGGCGCTCACGCCGCACCCGTGCAACGCGACGCGGGTGTACAAGTGCGACACGGCGGACGAGTGCGGGCAGCCG
GTGGGCGTGTGCGACGAATGGGGGTGCTCGTACAACCCGTCCAACTTCGGGGTCAAGGACTACTACGGGCGCAA
CCTGACGGTGGACACGAACCGCAAGTTCACGGTGACGACGCAGTTCGTGACGTCCAACGGGCGGGCGGACGGCG
AGCTGACCGAGATCCGGCGGCTGTACGTGCAGGACGGCGTGGTGATCCAGAACCACGCGGTCACGGCGGGCGGG
GCGACGTACGACAGCATCACGGACGGCTTCTGCAACGCGACGGCCACCTGGACGCAGCAGCGGGGCGGGCTCGC
GCGCATGGGCGAGGCCATCGGCCGCGGCATGGTGCTCATCTTCAGCCTGTGGGTTGACAACGGCGGCTTCATGA
ACTGGCTCGACAGCGGCAACGCCGGGCCCTGCAACGCCACCGAGGGCGACCCGGCCCTGATCCTGCAGCAGCAC
CCGGACGCCAGCGTCACCTTCTCCAACATCCGATGGGGCGAGATCGGCAGCACGTACAAGAGCGAGTGCAGCCA
CTAG