

BIOPROSPECTING FOR FUNGAL CELLULOLYTIC ENZYMES FOR BIOMASS
CONVERSION

by

Bingyao Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Biology—Doctor of Philosophy

2018

ABSTRACT

BIOPROSPECTING FOR FUNGAL CELLULOLYTIC ENZYMES FOR BIOMASS CONVERSION

by

Bingyao Li

High enzyme cost is a major bottleneck for the economic production of bioethanol and other sugar-derived chemical products from lignocellulosic biomass. To reduce the cost, it is necessary to find novel enzymes of higher efficiency and to understand more fully which enzymes are important and why.

To find more efficient endo- β 1,4-glucanases (EGs), a key enzyme in cellulose depolymerization, I compared the activity of six homologous fungal EGs that come from two distinct sub-families of GH5. When included into an enzyme mixture, two EGs with additional mannanase activity gave higher glucose yields from pretreated corn stover when mannan, a known cellulase inhibitor, was present. Therefore, it is possible that endoglucanases that have activity on mannan as well as cellulose protect other cellulase mixtures from inhibition by mannan inhibitors by degrading them, thereby increasing overall biomass conversion efficiency.

Lignin in lignocellulose is known to strongly inhibit the activity of cellulolytic enzymes, but such inhibition is largely reduced above pH 6. Unfortunately, current available *Trichoderma reesei*-based commercial enzymes lose much of their activity above pH 6. To find enzymes that can efficiently degrade cellulose at high pH, I purified and characterized an EG and a BG from the alkaliphilic fungus *Cladorrhinum bulbillosum*. Two *C. bulbillosum* cellulases showed higher activity than their *T. reesei* homologs between pH 6-8. Therefore, they may be useful as the basis of cellulase cocktails with better activity at higher pH.

Lytic polysaccharide mono-oxygenases (LPMOs) significantly enhance enzyme efficiency in cellulose degradation, but the mechanism is not well understood. To understand how LPMO enhances cellulose degradation, in Chapter 4 I studied the molecular interaction of an LPMO (TrAA9A) and a cellobiohydrolase (TlCel7A) with bacterial microcrystalline cellulose (BMCC) in collaboration with Dr. Bo Song. Cellulose conversion by TlCel7A was enhanced 8% by TrAA9A. Atomic force microscope observation revealed that BMCC ribbons were split into fibrils with a smaller diameter after TrAA9A treatment. The dividing of the cellulose microfibrils occurred more rapidly when TrAA9A and TlCel7A were added together compared to TrAA9A alone, while TlCel7A alone caused no separation. Therefore, TrAA9A may increase the accessible surface area of BMCC by separating large cellulose ribbons, and thereby enhance cellulose hydrolysis yield.

Copyright by
BINGYAO LI
2018

ACKNOWLEDGMENTS

First and foremost, I would like to thank my dissertation advisor Dr. Jonathan Walton. He gave me the opportunity to work on this research project in his lab and has been supportive since day one. Dr. Walton not only provided me a research assistantship during my five and half year Ph.D. study, but also guided me through my Ph.D. research. Without his support and help, I would not be able to pursue my Ph.D. at Michigan State University and complete my dissertation in a timely manner.

I would like to thank the members of my committee, Dr. Shi-You Ding, Dr. R. Michael Garavito, Dr. Claire Vieille and Dr. Curtis Wilkerson, and the past and present members of the Walton lab, for their help during my Ph.D. study.

I would like to thank Dr. Bo Song for his contribution to figures 4.4-4.7 and figure 4.9, and to the writing of the manuscript (Chapter 4) entitled “*Real-time imaging reveals that lytic polysaccharide monooxygenase promotes cellulase activity by increasing cellulose accessibility*”. I would also like to thank Dr. Shi-You Ding and Dr. Jonathan Walton for their contribution to the editing of the manuscript.

I would like to thank Dr. Dina Jabbour for the screening, isolation and identification of the fungus *Cladorrhinum bulbillosum*, and the contribution of Figure 3.2 in Chapter 3. I would like to thank Dr. Daniel Jones, Dr. Tony Schillmiller and Mr. Douglas Whitten, MSU Mass Spectrometry Facility, for their help with the mass spectrometry experimental design and operation.

Finally, I would like to thank my family and friends. Thank you to my mother for her encouragement and financial assistance to enable me to pursue my Ph.D. in the United States.

Thank you to my husband for his support, advice and encouragement to overcome hurdles to complete my Ph.D. study. Special thanks to my friends Daniel Brickley and K.C. Steckelberg for helping me adjust to life in the United States.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1: LITERATURE REVIEW	1
The current situation of petroleum-based transportation fuels	1
Bioethanol as alternative transportation fuel	2
Bioethanol from sugar and starch	3
Lignocellulosic biomass for bioethanol production.....	3
Conversion of lignocellulosic biomass to bioethanol and other chemicals	4
Compositional and structural organization of lignocellulosic biomass	5
Cellulose	5
Hemicelluloses	7
Lignin.....	8
Structural organization of the major components in lignocellulose	9
Enzymes for biomass saccharification.....	11
Classification of cellulose and hemicellulose degrading enzymes	11
Development of <i>Trichoderma reesei</i> -based enzyme mixtures	13
Structural characteristics of <i>T. reesei</i> cellulolytic enzymes.....	15
Biomass recalcitrance to enzymatic digestions.....	15
Effects of lignin and hemicellulose.....	16
Effects of cellulose structure.....	16
Strategies to improve lignocellulose conversion	17
Pretreatment	17
Improving enzyme efficiency in biomass conversion	18
CHAPTER 2: FUNCTIONAL DIVERSITY FOR BIOMASS DECONSTRUCTION IN	
FAMILY 5 SUBFAMILY 5 (GH5_5) OF FUNGAL ENDO-β1,4-GLUCANASE	20
Abstract	21
Introduction.....	22
Materials and methods	23
Substrates	23
Enzymes	24
Expression of GH5_5 endo- β 1,4-glucanases.....	24
Enzyme production and purification.....	27
Protein analysis	27
Protein deglycosylation.....	27
Preparation of lignin-rich residues.....	28
Enzyme assays	29
Results.....	30
Expression of six GH5_5 enzymes in <i>T. reesei</i> Δ <i>xyl1</i>	30
Enzyme activities on pure cellulose.....	34

Hydrolysis of pretreated corn stover.....	36
Substrate specificities of six endo-glucanases	37
Effect of lignin on GH5_5 enzyme activity	39
Influence of mannan on enzymatic hydrolysis of corn stover	40
Discussion.....	43
CHAPTER 3: <i>CLADORRHINUM BULBILLOSUM</i> ALKALIPHILIC CELLULASES FOR BIOMASS DECONSTRUCTION.....	46
Abstract.....	47
Introduction.....	47
Materials and methods	49
Prediction of carbohydrate-active enzymes in the <i>C. bulbillosum</i> , <i>T. reesei</i> and <i>P. anserina</i> genomes	49
Fungus growth	50
Enzymes, substrates and buffer system	50
Expression and purification of <i>C. bulbillosum</i> cellulases.....	51
Proteomic analysis (as provided by MSU RTSF).....	52
Enzyme assays	53
Results.....	54
CAZymes in the <i>C. bulbillosum</i> , <i>T. reesei</i> and <i>P. anserina</i> genomes	56
Purification of <i>C. bulbillosum</i> cellulases	59
Proteomic analysis of the purified <i>C. bulbillosum</i> cellulases	61
Comparing pH profiles of <i>C. bulbillosum</i> and <i>T. reesei</i> cellulases	63
Comparison of kinetic parameters of <i>C. bulbillosum</i> and <i>T. reesei</i> cellulases	67
Discussion.....	69
CHAPTER 4: REAL-TIME IMAGING REVEALS LYTIC POLYSACCHARIDE MONOOXYGENASE PROMOTES CELLULASE ACTIVITY BY INCREASING CELLULOSE ACTIVITY	71
Abstract.....	72
Introduction.....	73
Materials and methods	74
General chemicals.....	74
Bacteria strain and medium	75
Enzymes.....	75
Production of BMCC	75
Preparation of BMCC samples for AFM imaging.....	76
AFM operation.....	76
AFM image processing.....	77
Preparation of TrAA9A-treated BMCC.....	77
TrAA9A product analysis by mass spectrometry (method details provided by MSU RTSF)	78
Enzymatic degradation of BMCC.....	79
Results.....	79
Enhancement of TICe17A hydrolysis of BMCC by TrAA9A	79
MS analysis of TrAA9A products	80

Molecular action of TrAA9A on BMCC	85
Disassembly of BMCC ribbon into smaller fibrils by TrAA9A.....	91
Synergism between TrAA9A and TICel7A during hydrolysis of BMCC	93
Discussion	96
CHAPTER 5: FUTURE DIRECTIONS	98
REFERENCES	101

LIST OF TABLES

Table 2.1 Structural characteristics of the six GH5_5 enzymes	33
Table 3.1 All proteins identified in purified BG sample at >95% confidence (Scaffold)	62
Table 3.2 All proteins identified in partially purified CBH1 sample at >95% confidence (Scaffold)	62
Table 3.3 Kinetic parameters of four <i>C. bulbilosum</i> and <i>T. reesei</i> cellulases	68

LIST OF FIGURES

Figure 1.1 The structure of cellulose	6
Figure 1.2 Structural organization of the major components in lignocellulose	10
Figure 1.3 Current view on cellulose degradation by noncomplexed cellulolytic enzymes	12
Figure 2.1 Yields of Glc and Xyl from corn stover as a function of time and source of Cel5A for 8-component synthetic enzyme mixture plus either StCel5A or TrCel5A	31
Figure 2.2 SDS-PAGE of six recombinant endo- β 1,4-glucanases before (lanes 1, 3, 5, 7, 9, and 11) and after (lanes 2, 4, 6, 8, 10, and 12) chemical deglycosylation	33
Figure 2.3 Activities of six endoglucanases on cellulosic substrates	35
Figure 2.4 Yield of Glc from pretreated corn stover treated with a seven-component synthetic mixture plus different GH5_5 endo-glucanases at equal concentration	36
Figure 2.5 Activities of pure TrCel5A and StCel5A on defined substrates.....	38
Figure 2.6 Activity of six GH5_5 enzymes (reducing sugar assay) by themselves on different polysaccharides (logarithmic scale).....	39
Figure 2.7 Specific activities (reducing sugar assay) of six GH5_5 enzymes on CMC with and without addition of lignin.....	40
Figure 2.8 Inhibition by galactomannan of Glc release from pretreated corn stover by a seven-component synthetic enzyme mixture plus different GH5_5 endo-glucanases.....	42
Figure 3.1 <i>C. bulbillosum</i> growing on V8 juice agar, 10 days old	55
Figure 3.2 Comparison of relative activities of <i>C. bulbillosum</i> crude culture filtrate and commercial enzyme mixture CTec2/HTec2	56
Figure 3.3 Comparison of number of Carbohydrate-Active enZymes (CAZymes) in the <i>C. bulbillosum</i> , <i>T. reesei</i> and <i>P. anserina</i> genomes	58
Figure 3.4 Comparison of number of enzymes in eight CAZyme families that contain many of the most important enzymes and CBMs known to be necessary for efficient biomass conversion	58
Figure 3.5 First chromatographic fraction (anion exchange) of <i>C. bulbillosum</i> culture filtrate....	59

Figure 3.6 Cation exchange chromatography (CEX) of AEX fractions 5-7 (see Figure 3.5) to separate <i>C. bulbillosum</i> BG(s) and EG(s).....	60
Figure 3.7 Hydrophobic interaction chromatography of AEX fraction 23 (see Figure 3.5) to separate <i>C. bulbillosum</i> CBH1(s) and EG(s)	60
Figure 3.8 SDS-PAGE of purified <i>C. bulbillosum</i> and <i>T. reesei</i> cellulases	64
Figure 3.9 pH profiles of CbCel7E and TrCel7B	65
Figure 3.10 pH profiles of CbCel3A and TrCel3A.....	66
Figure 3.11 Initial reaction velocities of CbCel7E and TrCel7B at different concentrations of 4-nitrophenyl cellobioside.....	67
Figure 3.12 Initial reaction velocities of CbCel3A and TrCel3A at different substrate concentration.....	68
Figure 4.1 Cellulose conversion by TrAA9A and TICel7A	80
Figure 4.2 Regioselective hydroxylation of cellulose by lytic polysaccharide monooxygenases	81
Figure 4.3 ESI-MS analysis of oxidized cellobiose and cellotriose released from BMCC by TrAA9A	83
Figure 4.4 TrAA9A molecules move and diffuse randomly on BMCC.....	86
Figure 4.5 TrAA9A penetrating and moving inside a BMCC ribbon	88
Figure 4.6 Changes in morphology of BMCC ribbon when incubated with TrAA9A (indicated by cyan arrows) during 2 h continuous AFM observation	90
Figure 4.7 Disassembly of BMCC ribbon into smaller fibrils by TrAA9A	92
Figure 4.8 Morphology change of BMCC after 72 h incubation with or without enzymes	94
Figure 4.9 Synergism between TrAA9A and TICel7A during hydrolysis of BMCC	95
Figure 4.10 Comparison of cellulase hydrolysis of untreated and TrAA9A-treated BMCC	96

CHAPTER 1: LITERATURE REVIEW

The current situation of petroleum-based transportation fuels

Petroleum is formed by exposure of organismal remains to the heat and pressure of the earth's crust over millions of years. Its main components are hydrocarbons. In 1847, Scottish chemist James Young discovered that a light thin oil distilled from a natural petroleum seepage could be used as lamp oil. Since then, petroleum-derived fuels including gasoline and kerosene have been used worldwide and form the foundation of modern industrial economies. In particular, petroleum is the primary energy source for most vehicles that use the internal combustion engine. According to U.S. Energy Information Administration, approximately 88% of the U.S. transportation fuel used in 2016 comes from petroleum (<https://www.eia.gov/totalenergy/data/monthly/archive/00351704.pdf>).

Global economic development requires high petroleum input. In 2015, world liquid fuel consumption was about 91.1 million barrels (mb) per day and increased slightly to 98.3 mb per day in 2017 (U.S. Energy Information Administration, <https://www.eia.gov/outlooks/steo/data.cfm>). However, petroleum is considered to be a nonrenewable resource since its formation takes millions of years. Depletion of the finite hydrocarbon stock is inevitable due to current and increasing demand. It is predicted that the global conventional oil reserve will be depleted in the next 40 to 50 years (Sorrel et al. 2010; Vohra et al. 2014). Consumption of petroleum also causes negative impacts on the environment. For instance, transportation accounts for 26% of global emission of CO₂, a greenhouse gas that contributes to global warming (Chapman 2007). Excessive CO₂ emissions can also lead to ocean

acidification (Feely et al. 2009). Therefore, it is necessary to develop alternative transportation fuels that are renewable and environmentally friendly.

Bioethanol as alternative transportation fuel

Bioethanol was first suggested as an automotive fuel by Henry Ford in 1925, well before the widespread adoption of petroleum. He predicted that “The fuel of the future is going to come from fruit like that sumac out by the road, or from apples, weeds, sawdust – almost anything” (“Ford predicts fuel from vegetation”, 1925). There are several advantages to use bioethanol as transportation fuel. First of all, bioethanol is compatible with existing vehicle combustion engines and has a higher octane rating than gasoline, resulting in a lower engine knock (Agarwal 2007; Anderson et al. 2012). In addition, it contains 34.7% oxygen comparing to 0% in gasoline, which leads to approximately 15% higher combustion efficiency (Kar and Deveci 2006). Burning bioethanol also reduces emissions of the greenhouse gas CO₂ and air pollutants such as CO and benzene. Last but not least, bioethanol can be produced from renewable resources (Niven 2005; Nigam and Singh 2011).

By far, bioethanol is the most widely used alternative fuel for motor vehicles, with a global production reaching 25.7 billion gallons in 2015 (Gray et al. 2006; Zabed et al. 2017). It is produced from plant biomass that is synthesized through photosynthesis using only water and carbon dioxide (Chen 2005) and can be classified based on the type of feedstock from which it is made. First-generation bioethanol is produced from sugar- and starch-rich biomass, and second-generation bioethanol is made from lignocellulosic biomass. Third-generation bioethanol is designed to be produced by microbes including algae (Zabed et al. 2017).

Bioethanol from sugar and starch

Currently, bioethanol is mainly produced from sugar and starch feedstocks, accounting for 90% of the world total ethanol production (Kumar and Bhattacharya 2008). Residues of fruits and sugar crops contain high amount of soluble sugars including sucrose, glucose, and fructose, which can be directly fermented to ethanol by yeasts and other micro-organisms (Vohra et al. 2014; Zabed et al. 2017). Potential sugar feedstocks include grape pomace, apple pomace, date industry waste, sweet sorghum stalk, sugar beet root, and sugarcane stalk (Vohra et al. 2014; Zabed et al. 2017). Sugarcane juice is the main feedstock in Brazil, accounting for approximately 79% of that country's total bioethanol production. The cost of converting sugar feedstocks to ethanol is relatively low compared to other sources. Limitations of sugar as a feedstock is that the availability of sugar is limited by season, climate, and demand by other applications (Vohra et al. 2014).

Many crops store photosynthetic energy in the form of starch, which chemically is a branched chain of glucose connected together mainly by α -1,4-glycosidic bonds. Due to its worldwide availability and ease of conversion, starch is widely used for bioethanol production. In the United States, corn is the major starch source for bioethanol, 36% of which was used for bioethanol production in 2016 (U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy, <https://www.afdc.energy.gov/data/>). Other major starch crops such as potato, wheat, cassava and sweet potatoes are potential feedstocks for ethanol production (Zabed et al. 2017).

Lignocellulosic biomass for bioethanol production

Although fruits, sugar and starch crops can be used for bioethanol production, they are also important sources of food and feed. To avoid competition with food and feed,

lignocellulosic biomass is a more suitable feedstock for bioethanol production. Lignocellulosic biomass is the plant-cell-wall-based biomass that is composed of up to 75% polysaccharides (Jørgensen et al. 2007). The available lignocellulose feedstocks can be divided into three categories: wastes, forest products and energy crops. Biomass wastes include agricultural residues, waste crops, mill wood wastes, and municipal solid wastes. Forest products comprise wood scraps, sawdust, and logging residues (slash). Herbaceous crops, short rotation woody crops, and grasses such as *Miscanthus* (*Miscanthus giganteus*) and switchgrass (*Panicum virgatum*), are potential energy crops (Demirbaş 2001; Gomez et al. 2008).

Corn stover, which is composed of the leaves, stalks and husks, is the most abundant crop residue in the United States. The annual US yield of corn stover is estimated to be 82 million dry tonnes, sufficient to produce 4-5 billion gallons of ethanol (Kim and Dale 2004). Other agricultural residues available in large quantities worldwide include sugarcane bagasse, rice straw, and wheat straw. It has been estimated that total agricultural residues and crop waste could produce 129 billion gallons of ethanol, which is five times the global ethanol production in 2015 (Kim and Dale 2004; <http://www.ethanolrfa.org/>).

The energy that plants store in biomass every year is 10 times what is consumed (Demirbaş 2001). The worldwide biomass production by land plants is estimated at $1.7 - 2 \times 10^{11}$ tonnes per year, 70% of which represents plant cell walls (Duchesne and Larson 1989; Lieth 1975). Therefore, lignocellulosic biomass is a potentially sustainable resource for biofuel production.

Conversion of lignocellulosic biomass to bioethanol and other chemicals

Besides ethanol, the sugar monomers derived from the lignocellulosic biomass can also serve as platforms for the production of other valuable chemical products. For instance, glutamic

acid, glycerol, and sorbitol converted from sugar possess the potential to be transformed into new families of useful molecules (Werpy et al. 2004). In the future, many of the chemicals currently produced from petroleum could be instead made from biomass in “biorefineries” (Cherubini 2010; Werpy et al. 2004).

The major steps involved in the course of converting lignocellulose into ethanol and chemicals are: pretreatment, enzyme production, enzymatic hydrolysis of lignocellulose, and sugar fermentation. The process where all three steps are conducted separately is the most commonly used by far, whereas consolidated bioprocessing, in which biomass is converted to fuel and chemicals in a single step, is more cost-effective and considered the main biomass processing strategy of the future (Lynd et al. 2002; Wilson 2009; Xu et al. 2009).

Compositional and structural organization of lignocellulosic biomass

As a plant-cell-wall-based biomass, lignocellulose is mainly composed of cellulose, hemicellulose and lignin, which account for 20-50%, 15-35% and 10-30% of the dry weight of the wall material, respectively (Pauly and Keegstra 2008). It also contains a small amount of soluble sugar, pectin, protein, lipid, and minerals (Cheng 2005; Chundawat et al. 2011).

Cellulose

Cellulose is the most abundant organic polymer on earth, accounting for up to 51% of plant cell walls (Klemm et al. 2005; Paul et al. 2008). It is produced mainly by photosynthetic plants as well as some nonphotosynthetic bacteria, fungi and other organisms (Zhang and Lynd 2004). Chemically, cellulose is a linear chain of 100 to 15,000 glucose residues joined together by β 1,4-glucosidic bonds and is water-insoluble (Agbor et al. 2011). Cellobiose is the repeating

unit of cellulose, where two adjacent glucose residues are linked by a β 1,4 glucosidic bond, which results in the two glucose units being rotated 180° in respect to each other (Figure 1.1a). The configuration of the cellobiose unit results in symmetry of the cellulose chain, in which equal proportions of the hydroxyl groups are distributed on each side of the molecule. Cellulose molecules adhere to each other along their lengths via Van de Waals forces and inter- and intra-chain hydrogen bonds. A group of cellulose chains (the number is in dispute) held together by these forces thus form a rigid macromolecule denoted the cellulose microfibril. Cellulose microfibrils further aggregate by Van de Waals force and hydrogen bonds to form microfibril bundles (Ding and Himmel 2006; Zhang and Lynd 2004).

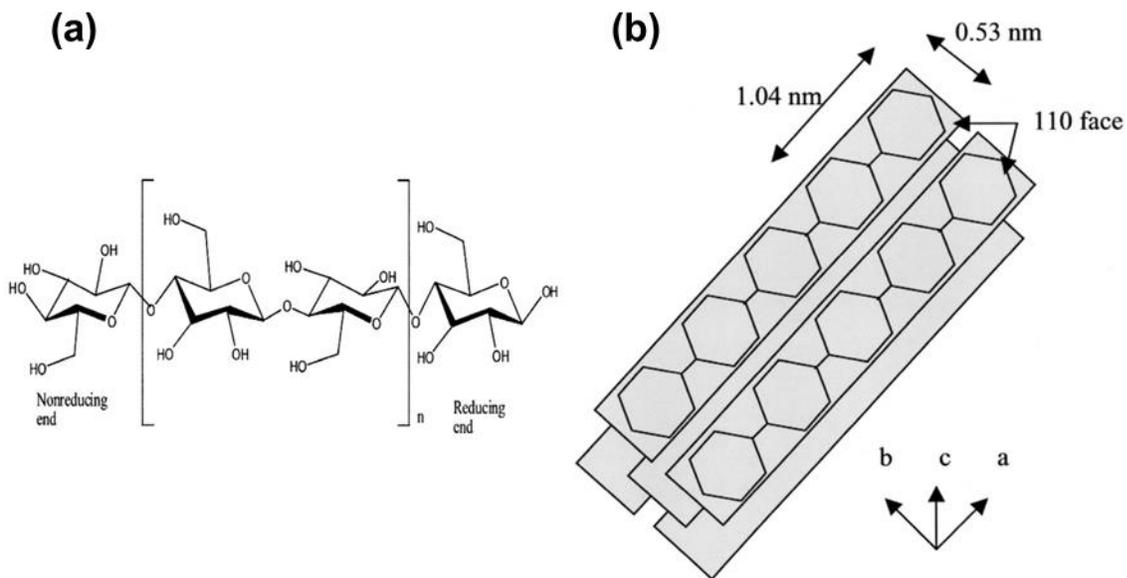


Figure 1.1 The structure of cellulose. (a) Cellulose chain featuring repeating β 1,4-linked cellobiose unit. (b) Structure of a cellulose crystal. The axes of the repeating unit (cellobiose) are: $a = 0.817$ nm, $b = 1.04$ nm, and $c = 0.786$ nm. The faces of the glucopyranose rings are parallel to the ab plane of the crystal. Reprinted from Zhang and Lynd (2004) with permission.

Cellulose microfibril bundles are composed of crystalline regions where cellulose chains are aligned in parallel, amorphous regions in which the molecules are in poor alignment, and

transitional paracrystalline regions (Martínez-Sanz et al. 2016). In crystalline regions, cellulose chains are stacked in sheets in which glucopyranose rings in a single molecule lie in the same plane (Figure 1.1b). The six faces of microfibril bundles show distinct properties when the cellulose fibrils interact with the aqueous environment or with enzymes. As shown in Figure 1.1b, the two ab planes that contain all the carbon atoms and β 1,4-glucosidic bonds are hydrophobic, while the two ac planes containing all of the hydroxyl groups are hydrophilic. The remaining two bc planes represent reducing and nonreducing ends of the microfibrils (Zhang and Lynd 2004). Due to its homogeneity and great abundance, cellulose is of major interest for the production of glucose by biorefineries (Himmel et al. 2007; Ragauskas et al. 2006).

Hemicelluloses

Hemicellulose is the collective name for a heterogeneous group of branched polysaccharides with 500-3000 sugar units (Gibson 2012). They all contain a β 1,4-linked sugar backbone of glucose, mannose, xylose, or galactose in an equatorial configuration. Hemicelluloses are also an important source for monomeric sugars. Based on the backbone sugars, hemicelluloses can be classified as xyloglucans, xylans, mannans, glucomannans, or mixed linkage glucan (Scheller and Ulvskov 2010). Xyloglucan contains a β 1,4-linked glucose backbone substituted with α 1,4-xylose residues that can be further substituted with galactose or fucose. Xylans have a β 1,4-linked xylose backbone that can be branched with glucuronosyl and arabinose residues. Interestingly, arabinose branches can form ester bonds with ferulic acid that can cross-link with lignin (Grabber 2005; Hatfield et al. 1999). The backbone of mannan consists entirely of mannose while that of glucomannan is composed of both mannose and glucose. Both polysaccharides can be branched with galactose. Different from other hemicelluloses, mixed-

linkage glucan is not branched but its glucose residues are joined by alternating β 1,3 and β 1,4 linkages in a 1:3 or 1:4 ratio (Scheller and Ulvskov 2010).

The compositions of hemicelluloses vary from species to species (Scheller and Ulvskov 2010). In general, mixed-linkage glucan is present in grasses but absent in other seed plants. Xylans are abundant in the primary and secondary wall of grasses and present in the secondary wall of dicots and conifers. Xyloglucan is the most abundant hemicellulose in the primary wall of seed plants except for grasses. Both mannan and glucomannan are common wall polysaccharides in green algae while glucomannan is a main wall component in gymnosperms (Popper and Fry 2003; Popper 2008; Scheller and Ulvskov 2010). Some seeds accumulate hemicelluloses such as xyloglucan and galactomannans as energy reserves or mucilages to hold water (Kumar and Bhattacharya 2008).

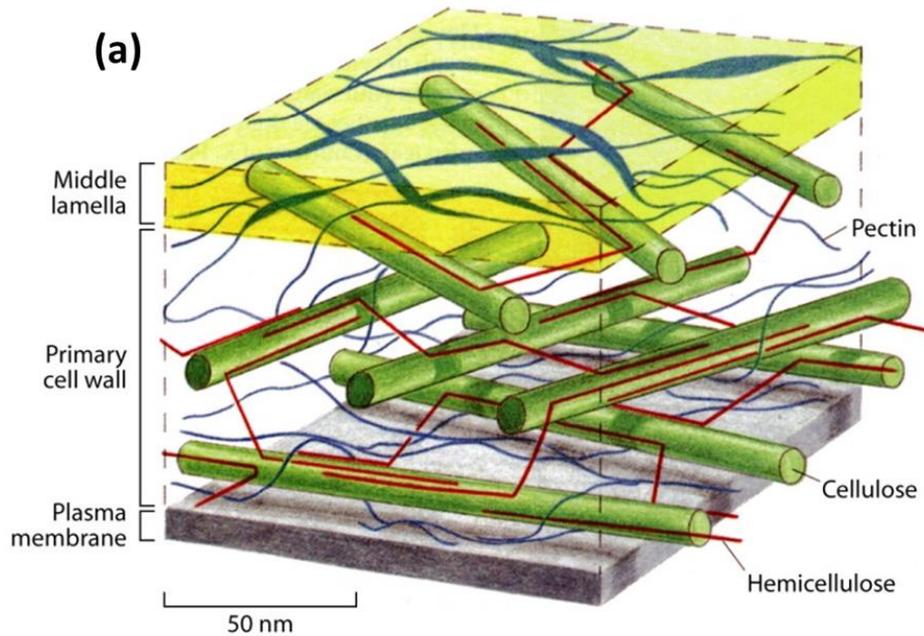
Lignin

Lignin is the second most abundant organic polymer on earth. Chemically, it is a group of heterogeneous 4-hydroxyphenylpropanoid polymers that are mainly composed of three monolignol units: guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units. The composition of monolignols in lignin varies among species and cell types. In general, G units are dominant in softwood lignin, while hardwood lignin consists mainly of G and S units and trace amount of H units. Monocots contain equal amounts of G and S units in lignin and more H units than dicots. Lignin is deposited into the secondary cell wall after the plant cell stops expanding, making the walls rigid and impermeable to water. The incorporation of lignin strengthens the plant stem and was crucial for the evolution and adaptation of plants to a terrestrial habitat. In addition, the water-impermeable property of lignin enables the transport of water and solutes in the vascular

tissue. Lignin also protects the plant against pathogens and insects (Boerjan et al. 2003; Campbell and Sederoff 1996; Vanholme et al. 2010).

Structural organization of the major components in lignocellulose

Cellulose, hemicellulose and lignin interconnect to each other, resulting in a polymer network where cellulose microfibrils are surrounded by a matrix of hemicellulose and lignin (Figure 1.2). Cellulose microfibrils aggregate into bundles and form the skeleton of the matrix. Hemicellulose fills the space between microfibril bundles and connects to the cellulose surface via hydrogen bonds, while it forms cross-links to itself and to exterior lignin (Chundawat et al. 2011; Scheller and Ulvskov 2010; Somerville et al. 2004; Zhang and Lynd 2004; Zhao et al. 2012).



(b)

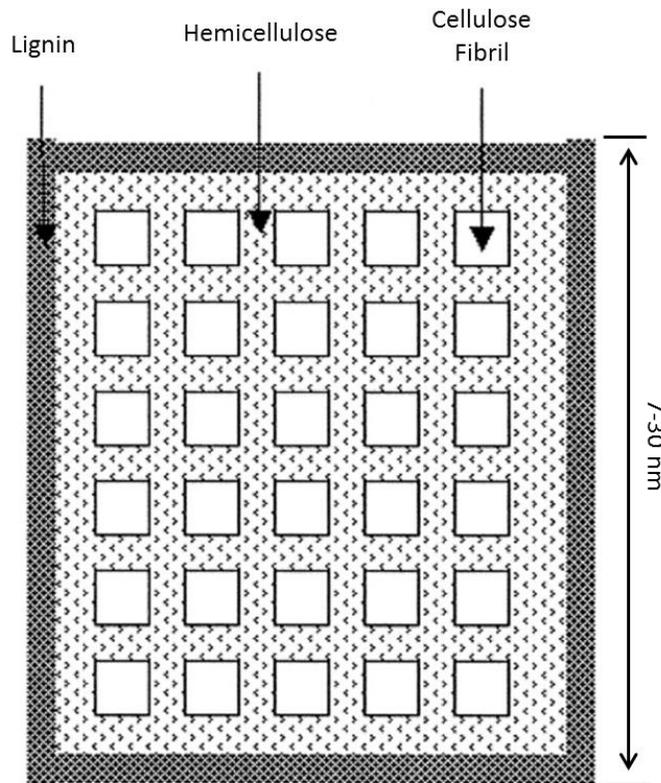


Figure 1.2 Structural organization of the major components in lignocellulose. (a) Simplified primary cell wall model showing the interconnection between cellulose and other wall polysaccharides. Reprinted from Scheller and Ulvskov (2010) with permission. (b) Organization of lignocellulose. Reprinted from Zhang and Lynd (2004) with permission.

Enzymes for biomass saccharification

Biomass degrading enzymes are commonly found in bacteria and fungi (Kumar and Bhattacharya 2008; Lynd et al. 2002). For thorough cell wall deconstruction, microbes produce multiple extracellular enzymes (Lynd et al. 2002). Cell wall degrading enzymes are classified into two main types: complexed and noncomplexed (Bayer et al. 2004; Zhang and Lynd 2004). The enzyme components in complexed enzymes form a stable high-molecular weight complex called the cellulosome, which is commonly found in anaerobic bacteria such as *Clostridium thermocellum* (Lamed and Bayer 1988). Noncomplexed enzymes exist as single proteins and are widely seen in aerobic bacteria and fungi such as *Streptomyces lividans* and *Trichoderma reesei* (Kluepfel et al. 1986; Schülein 1988). Enzymes in both systems are classified by their substrate specificity, molecular mechanism, and amino acid sequence similarity (Lynd et al. 2002; Lombard et al. 2013).

Classification of cellulose and hemicellulose degrading enzymes

Traditional cellulose-degrading enzymes (cellulolytic enzymes or cellulases) are classified into endo- β 1,4-glucanases, cellobiohydrolases, and β -glucosidases (Figure 1.3). Endo- β 1,4-glucanases (EGs) randomly cleave the internal glucosidic bonds of cellulose molecules at amorphous regions and generate cello-oligosaccharides. Cellobiohydrolases (CBHs) adsorb to crystalline regions, move processively from one end of a cellulose chain to the other, and liberate cellobiose from the substrate. β -Glucosidases (BGs) convert cellobiose to glucose (Lynd et al. 2002). Lytic polysaccharide monooxygenase lyases (LPMOs) differ from cellulases that break β 1,4-linkage via hydrolysis in that they cleave cellulose chains by oxidation of the C1 and/or C4 carbon of the glucose units and generate oxidized cello-oligomers (Beeson et al. 2015;

Hemsworth et al. 2015; Horn et al. 2012; Vaaje-Kolstad et al. 2010; Figure 1.3). Similar to cellulases, hemicellulose-degrading enzymes (collectively known as hemicellulases) are classified based on their substrates and mode of catalytic action. For instance, endo- β 1,4-xylanase, which cleaves xylan chains in an endo- manner, and β -xylosidase, which cleaves individual xylose residues in an exo- manner, together hydrolyze xylan to xylose. Likewise, endo- β 1,4-mannanase and β -mannosidase convert mannan to mannose (de Vries and Visser 2001).

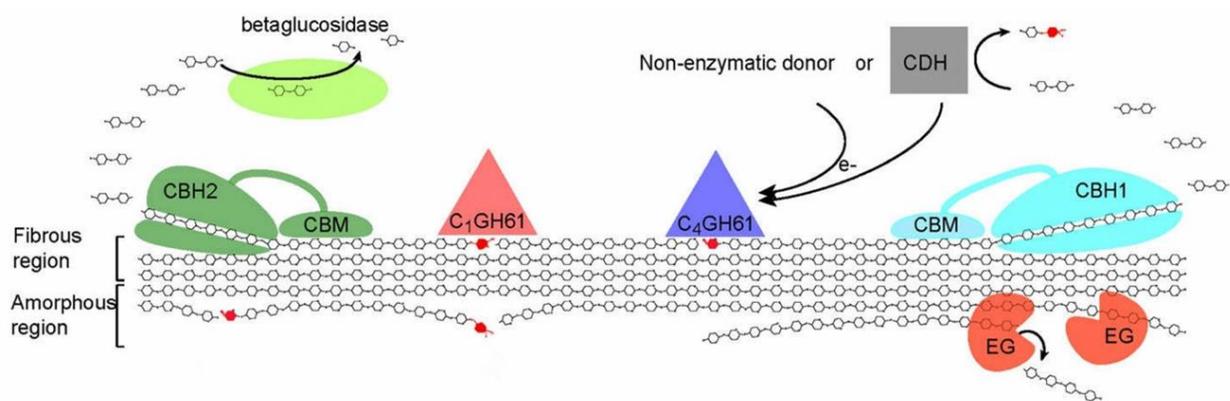


Figure 1.3 Current view of cellulose degradation by noncomplexed cellulolytic enzymes. Abbreviations: EG, endoglucanase; CBH, cellobiohydrolase; CDH, cellobiose dehydrogenase; CBM, carbohydrate-binding module. The picture shows a C1 and a C4 oxidizing GH61 (now classified as AA9) which generate non-oxidized ends for CBH2 and CBH1, respectively, to act on (oxidized sugars are colored red). CDH may provide GH61 with electrons. Other non-enzymatic reductants (electron donors) have been demonstrated to induce oxidative activity (e.g. reduced glutathione, ascorbic acid, and gallic acid). Reprinted from Horn et al. (2012) with permission.

Cellulose and hemicellulose degrading enzymes are grouped into families based on the similarity of their amino acid sequences, where enzymes in the same family share similar sequence and structure (Henrissat and Davies 1997). The classification is updated and maintained by the Carbohydrate-Active enZymes (CAZy) database (www.cazy.org; Lombard et

al. 2013). Many CBHs belong to glycoside hydrolase (GH) families 6 and 7, several EGs belong to GH5, 6 and 7, BGs to GH1 and GH3, endo- β 1,4-xylanases to GH11 and GH12, and β -xylosidases to GH3 and GH43 (Henrissat and Bairoch 1993). LPMOs were previously classified in GH61 but are now classified in auxiliary activities family AA9, AA10, AA11, and AA13 (Hemsworth et al. 2015).

Development of Trichoderma reesei-based enzyme mixtures

The fungus *Trichoderma reesei* (teleomorph *Hypocrea jecorina*) is a filamentous mesophilic ascomycete that was isolated from rotting US Army cotton tents on the Solomon Islands during World War II. Mary Mandels and Elwyn T. Reese discovered that *T. reesei* is highly active at degrading crystalline cellulose and showed great potential for producing extracellular cellulases (Allen et al. 2009; Mandels and Reese 1957; Mandels and Weber 1969). Since then, many studies have been conducted to explore the potential of *T. reesei* cellulases for economic saccharification of lignocellulosic biomass for bioethanol production (Bischof et al. 2016; Wilson 2009).

T. reesei produces at least two CBHs (CBH1 and CBH2), five EGs (EGI to EGV), and two BGs (BGI and BGII) (Chen et al. 1992; Nogawa et al. 2001). CBH activities are essential for the degradation of crystalline cellulose. CBH1 (also called TrCel7A) hydrolyzes the cellulose chain from the reducing end. In contrast, CBH2 (TrCel6A) liberates cellobiose from the nonreducing end. The two CBHs exhibit higher activity when added together than the sum of their individual enzyme activities, a phenomenon known as synergism (Teeri 1997). Synergism is also observed between other cellulases: CBH and EG, CBH and BG, and EG and BG (Henrissat et al. 1985; Woodward and Wiseman 1982). Moreover, cellulases show synergy with

hemicellulases and accessory enzymes such as xylanases and LPMO's when degrading lignocellulosic substrates (Banerjee et al. 2010b,c; Harris et al. 2010).

New *T. reesei* strains have been developed via mutagenesis to improve their capability to produce extracellular biomass degrading enzymes (Seidl and Seiboth 2010). For instance, strain RUT-30 produces 15-20-fold more secreted protein than the native strain QM6a (Peterson and Nevalainen 2012). Genes in *T. reesei* strains were also mutated to improve the stability of the secreted enzymes. Nagendran et al. (2009) showed that commercial cellulase mixtures are lacking the major proteases, which enhances the stability of the commercial enzymes. Improved *T. reesei* strains are used in industry for the development of commercial enzyme mixtures such as Accellerase, Spezyme CP and Cellic CTec3 for biomass saccharification (Bischof et al. 2016; Nagendran et al. 2009; Wilson 2009).

According to Nagendran et al. (2009), *T. reesei* CBH1 and CBH2 are the first and second most abundant proteins in the commercial enzyme mixture Spezyme CP, accounting for 24% and 14% of the total secreted proteins, respectively. Other major components include three EGs, one BG, one LPMO, one xylanase, one β -xylosidase, one xyloglucanase, and one glucuronoyl esterase. To understand which enzymes in the mixtures are the most important to biomass conversion and why, synthetic cocktails of purified biomass-degrading enzymes were constructed (Banerjee et al. 2010a,b,c,d). Banerjee et al. (2010b) showed that at least eight *T. reesei* enzymes are required for efficient lignocellulose degradation, including a CBH1, a CBH2, a EG, a BG, an LPMO, two xylanases and a β -xylosidase. These enzymes are among the most abundant in Spezyme CP, a commercial "cellulase" product (Nagendran et al. 2009).

Structural characteristics of T. reesei cellulolytic enzymes

Besides the catalytic domain, many cellulolytic enzymes from *T. reesei* have family 1 carbohydrate binding module(s) (CBM) and a heavily glycosylated linker peptide (Beckham et al. 2010; Linder and Teeri 1996; Nimlos et al. 2012; Srisodsuk et al. 1993). The CBM adsorbs to the cellulose surface and brings the catalytic domain into proximity with its substrate (Boraston et al. 2004). Structural studies suggest that three aromatic amino acid residues in the CBM are aligned to assist adsorption to the cellulose surface (Kraulis et al. 1986). The active sites in the catalytic domain of different types of *T. reesei* cellulolytic enzymes have different architectures. CBH1 and CBH2 have a tunnel-shape active site, which enables them to proceed from one end to the other end of the cellulose chain (Divne et al. 1993, 1998; Igarashi et al. 2009; Rouvinen et al. 1990). In contrast, EGI enzymes contain a groove-shape active site and BGI enzymes have a pocket-shape active site (Henriksson et al. 1996; Karkehabadi et al. 2014). Despite different active sites, cellulases catalyze the hydrolysis of β 1,4-glucosidic bonds using an acid hydrolysis mechanism utilizing an acid (proton donor) and a base (nucleophile) (Davies and Henrissat 1995). In contrast, the active site of LPMO TrAA9 is relatively flat with a copper coordination complex composed of a copper ion and three nitrogens contributed by two conserved histidine residues (Karkehabadi et al. 2008; Quinlan et al. 2011).

Biomass recalcitrance to enzymatic digestions

Although lignocellulosic biomass contains sugars required for fuel and chemical production, one challenge is that lignocellulose is very recalcitrant to enzymatic digestion. The

presence of lignin and hemicellulose significantly influences lignocellulose digestibility (Chundawat et al. 2011; Himmel et al. 2007; Van Dyk and Pletschke 2012; Zhao et al. 2012).

Effects of lignin and hemicellulose

Lignin is one of the most important factors that negatively affects enzymatic digestion of lignocellulosic biomass (Mansfield et al. 1999). Removal of lignin usually increases biomass saccharification yield because lignin irreversibly adsorbs and inhibits cellulases (Chang and Holtzapfle 2000; Draude et al. 2001; Yoshida et al. 2008; Berlin et al. 2006; Palonen et al. 2004; Ooshima et al. 1990; Rahikainen et al. 2013). Hemicelluloses also can inhibit cellulase activities. For instance, xylan and xylo-oligomers inhibit cellulase hydrolysis of lignocellulose and this inhibition is reduced when the reaction mixtures are supplemented with xylanase and β -xylosidase (Qing et al. 2010; Qing and Wyman 2011; Zhang et al. 2012). Mannan and mannan oligosaccharides are also strong cellulase inhibitors (Kumar and Wyman 2014). As the main constituents of the cell wall matrix, lignin and hemicelluloses also act as physical barriers to restrict the access of cellulases to cellulose (Liao et al. 2005; Mooney et al. 1998).

Effects of cellulose structure

Cellulose molecules form a rigid macromolecule via inter- and intra-chain hydrogen bonds (Ding and Himmel 2006). Many cellulose chains buried inside the cellulose macromolecule are therefore not exposed to the aqueous environment. Accessible surface area is the surface area on which cellulases can contact cellulose macromolecules. The accessible surface area can be divided into external and internal surfaces (Zhang and Lynd 2004). While the external surface area is affected mainly by the dimension of cellulose microfibril bundles, the

internal surface area is governed by the void volumes between the microfibril bundles (Moon et al. 2011; Rowland 1997; Zhao et al. 2012). Many studies show that cellulose digestibility is related to accessible surface area (Converse et al. 1990; Fan et al. 1981; Grethlein 1985; Yeh et al. 2010). Since the prerequisite for enzymatic degradation of cellulose is the adsorption of the enzyme onto the substrate, accessible surface area is considered a critical factor that limits the enzymatic degradation of cellulose (Arantes and Saddler 2011; Huang et al. 2010). Cellulose crystallinity and a high degree of polymerization also limit the enzymatic degradation of cellulose particles, although not as significantly as accessible surface area (Chundawat et al. 2011b; Koullas et al. 1990; Nazhad et al. 1995; Huang et al. 2010).

Strategies to improve lignocellulose conversion

Saccharification of lignocellulosic biomass can be improved in at least two ways: by reducing biomass recalcitrance or by increasing enzyme efficiency.

Pretreatment

Lignocellulosic feedstocks are usually pretreated before enzyme digestion. The goal of pretreatment is to make cellulose more accessible to enzyme degradation (Chundawat et al. 2011a; Kumar et al. 2009). Pretreatment methods can be divided into several categories, including physical, chemical, biological, organic solvent-fraction, and combinations (Alvira et al. 2010; Kumar et al. 2009; Sarkar et al. 2012). Although what physico-chemical properties of lignocellulose are altered by pretreatments are not fully understood, several studies reveal that chemical pretreatments result in: 1) the cleavage of cross-links between hemicellulose and lignin and the removal of hemicellulose; 2) modification and redistribution of lignin; 3) cellulose

decrystallization (Chundawat et al. 2011a). For instance, diferulate ester linkages between lignin and hemicellulose are hydrolyzed during alkali pretreatments such as ammonia fiber expansion (AFEX) and alkaline peroxide (AHP) (Banerjee et al. 2011; Chundawat et al. 2010). AFEX pretreatment also alters the distribution of lignin and hemicellulose while keeping the intactness of the two components as well as cellulose (Chundawat et al. 2010, 2011a). In contrast, acid pretreatment solubilizes hemicellulose to sugars and ionic liquid pretreatment reduces the crystallinity of cellulosic substrates (Arora et al. 2010; Kumar et al. 2009; Pingali et al. 2010; Wyman et al. 2005). Degradation of hemicelluloses, cleavage of cross-links, and redistribution of lignin removes lignin and hemicelluloses that act as both a physical barrier and as direct enzyme inhibitors. Such property changes also increase cellulose accessibility (Chundawat et al. 2011a).

Improving enzyme efficiency in biomass conversion

High enzyme cost is a bottleneck for economic bioethanol production (Banerjee et al. 2010a). The cost of biomass degrading enzymes is estimated to be between \$0.68/gal and \$1.47 per gallon bioethanol produced (Klein-Marcuschamer et al. 2012). The reason for the high cost is that compared to starch hydrolysis, 40- to 100- fold more enzyme loading is required to produce an equivalent amount of ethanol (Merino and Cherry 2007). Due to the difficulty of reducing the enzyme production cost, it is necessary to improve enzyme efficiency, which would lower the enzyme loading and hence the cost (Klein-Marcuschamer et al. 2012).

Many approaches could be taken to improve the efficiency of biomass degrading enzymes, including enzyme improvement by directed evolution and protein engineering, bioprospecting for superior versions of current enzymes, and screening for hemicellulases, accessory enzymes and other helper proteins that enhance the efficiency of existing enzyme

mixtures (Banerjee et al. 2010a; Chundawat et al. 2011; Duan and Feng 2010; Lin et al. 2011; Merino and Cherry 2007). In my dissertation, I took several approaches. One approach was to bioprospect for novel and better fungal cellulose-degrading enzymes. Chapters 2 and 3 of the dissertation focus on bioprospecting for superior versions of current enzymes. Superiority can be defined in several ways: by improved catalytic efficiency, improved synergism with other enzymes, or enhanced properties such as a different or broader pH optimum. In Chapter 4 of the dissertation, I contributed to our basic understanding about how one important “accessory” enzyme, lytic polysaccharide monooxygenase (LPMO), enhances biomass degradation in combination with conventional hydrolytic cellulases.

CHAPTER 2: FUNCTIONAL DIVERSITY FOR BIOMASS DECONSTRUCTION IN FAMILY 5 SUBFAMILY 5 (GH5_5) OF FUNGAL ENDO- β 1,4-GLUCANASE

The material in this chapter has been published: Li B, Walton JD (2017) Functional diversity for biomass deconstruction in family 5 subfamily 5 (GH5_5) of fungal endo- β 1,4-glucanases.

Applied Microbiology and Biotechnology, 101(10): 4093-4101

Abstract

Endo- β 1,4-glucanases in CAZy glycoside hydrolase family 5 (GH5) are ubiquitous enzymes among the multicellular fungi and are common components of enzyme cocktails for biomass conversion. We recently showed that an endo-glucanase of subfamily 5 of GH5 (GH5_5) from *Sporotrichum thermophile* (StCel5A) was more effective at releasing glucose from pretreated corn stover, when part of an eight-component synthetic enzyme mixture, compared to its closely related counterpart from *Trichoderma reesei*, TrCel5A. StCel5A and TrCel5A belong to different clades of GH5_5 (GH5_5_1 and GH5_5_2, respectively). To test whether the superior activity of StCel5A was a general property of all enzymes in the GH5_5_2 clade, StCel5A, TrCel5A, and two additional members of each subfamily were expressed in a common host that had been engineered to suppress its native cellulases (*T. reesei* Δ *xyl1*) and compared against each other alone on pure substrates, in synthetic mixtures on pure substrates, and against each other in synthetic mixtures on real biomass. The results indicated that superiority is a unique property of StCel5A and not of GH5_5_2 generally. The six Cel5A enzymes had significant differences in relative activities on different substrates, in specific activities, and in sensitivities to mannan inhibition. Importantly, the behavior of the six endo-glucanases on pure cellulose substrates did not predict their behavior in combination with other cellulolytic enzymes on a real lignocellulosic biomass substrate.

Introduction

Enzymatic deconstruction of lignocellulosic biomass to its component sugars is a promising technology for the production of biofuels and other chemical feedstocks. However, currently available enzyme preparations are inefficient and hence prohibitively expensive. Potential strategies to reduce the loading, and hence the cost, of cellulolytic enzymes have been proposed and in some cases experimentally explored, including optimization of the relative proportions of key enzymes, discovery or development of superior versions of existing enzymes, enzyme recycling, engineering of more easily digestible biomass feedstocks, and more effective pretreatments (Banerjee et al. 2010a; Klein-Marcuschamer et al. 2012). Probably ultimately a combination of some or all of these approaches will be necessary to achieve the necessary reduction in biomass enzyme cost.

Using synthetic, defined mixtures of cellulases and accessory enzymes, we have developed a platform suitable for the discovery of novel enzymes and superior versions of existing enzymes (Banerjee et al. 2010b,c). The GENPLAT system incorporates an array of pure enzymes, mixture optimization with design of experiment software, robotic pipetting of biomass and enzymes, and automated assay of released glucose (Glc) and xylose (Xyl) (Walton et al. 2011). Using this system, we identified and purified an alternative endo- β 1,4-glucanase (StCel5A) from *Sporotrichum thermophile* (also known as *Myceliophthora thermophila* or *Thermothelomyces thermophila*) that showed improved Glc release from pretreated corn stover compared to TrCel5A, its counterpart from *Trichoderma reesei* (synonym *Hypocrea jecorina*) (Ye et al. 2014). The two enzymes were shown to define two separate clades (GH5_5_1 and GH5_5_2, respectively) of subfamily 5 of glycoside hydrolase (GH) family 5 (Aspeborg et al.

2012; Ye et al. 2014). Factors that might account for the superiority of StCel5A were examined, including substrate range, differences in pH and temperature optima, phylogenetic relatedness among Family 5 glycoside hydrolases, and structural features (Ye et al. 2014). Significantly, the superiority of StCel5A was manifested only when the enzymes were tested on real biomass in combination with other cell wall degrading enzymes including cellobiohydrolases (CBH1 and CBH2) and lytic polysaccharide mono-oxygenases (AA9), and not when tested by themselves on pure cellulosic substrates (Ye et al. 2014). In order to further our understanding of the basis of superiority of StCel5A over TrCel5A, with the ultimate goal of developing superior enzyme cocktails for biomass conversion, we expressed and enzymatically compared four additional related fungal endo- β 1,4-glucanases from clades 1 and 2 to test if superiority were a general feature of clade 2.

Materials and methods

Substrates

Low-viscosity carboxymethyl cellulose (CMC, Catalog C5678), and locust bean gum galactomannan (G0753) were purchased from Sigma-Aldrich (St. Louis, MO), oat β -glucan (mixed-linkage glucan) and β 1,4-mannan from Megazyme, Ltd. (Bray, Ireland), and *Amorphophallus konjac* glucomannan from Now Foods (Bloomington, IL). Corn stover was ground in a Wiley mill to 0.5 mm particle size and pretreated with 12.5% alkaline hydrogen peroxide (AHP) (Banerjee et al. 2011). None of the stover was discarded during grinding. The biomass was neutralized but not washed after pretreatment. The weight of the added salt was considered in calculating the glucan content of the pretreated biomass.

Enzymes

Enzymes used in the synthetic enzyme mixtures were cellobiohydrolase 1 (CBH1), cellobiohydrolase 2 (CBH2; GenBank P07987), β -glucosidase (BG; AAA18473), endo- β 1,4-xylanase 2 (EX2; AAB29346), endo- β 1,4-xylanase 3 (EX3; BAA89465), β -xylosidase (BX; CAA93248) and lytic polysaccharide mono-oxygenase AA9 (formerly known as GH61A; CAA71999). All of the enzymes were *Trichoderma reesei* genes expressed in *Pichia pastoris* except CBH1, which was from *T. longibrachiatum* and purchased from Megazyme, Ltd., Bray, Ireland (Banerjee et al. 2010b,c). Proportions of enzymes in the four-component synthetic mixtures (4-CSM) (i.e., three fixed enzymes plus alternate GH5_5) were, by mass, 40% CBH1, 13% CBH2, and 7% BG, plus the alternate endo- β 1,4-glucanase at 40%. Proportions in the 8-component synthetic mixture (8-CSM) were 30% CBH1, 10% CBH2, 5% BG, 5% EX2, 5% EX3, 5% BX, and 10% AA9 plus 30% of the alternate GH5_5.

Driselase™ (Lot SLBH7888V) and lysing enzyme (from *Trichoderma harzianum*, Lot SLBG2959V) were obtained from Sigma-Aldrich. Commercial cellulases CTec2™ and HTec2™, which were used for the isolation of lignin (see below) were a generous gift of Novozymes, Inc. (Davis, CA). Pronase™ protease (from *Streptomyces griseus*) was purchased from Calbiochem (San Diego, CA).

Expression of GH5_5 endo- β 1,4-glucanases

The six expressed genes were ApCel5A from *Aureobasidium pullulans*, GenBank AEM23896; PaCel5A from *Polyporus arcularius*, GenBank BAF75943; TrCel5A from *T. reesei*, GenBank ABA64553; HgCel5A from *Humicola grisea*, GenBank BAA12676; StCel5A from *S. thermophile*, GenBank AEM23898; and VvCel5A from *Volvariella volvacea*, GenBank

ACE06751 and AAG59832 (Ding et al. 2001, 2002; Lee et al. 2011; Saloheimo et al. 1988; Tambor et al. 2012; Takashima et al. 1997; Zheng and Ding 2013). The genes for five heterologous GH5_5 enzymes were synthesized commercially with codon optimization based on the *T. reesei* (*Hypocrea jecorina*) codon usage table (Nakamura et al. 2000) and an insertion of a C-terminal 6x His tag and *Cla*I and *Sal*I restriction sites upstream and downstream, respectively. A *Not*I site was added between the stop codon and the *Sal*I site for terminator insertion. Accession numbers for the synthetic genes are KY458054-KY458058. A 491-bp fragment from the terminator of TrCel5A was amplified by PCR from genomic DNA using forward primer 5'-AACGCGGCCGCCACTCTGAGCTGAATGCACAA-3' and reverse primer 5'-GTAGTCGACCCAACAGATAGAAAGAAGAAAGGA-3'(the *Not*I and *Sal*I sites are underlined).

The six synthetic genes with the inserted terminators were cloned into the *Cla*I/*Sal*I site of the expression vector developed by Uzbas et al. (2012). This vector contains the *hph* gene conferring hygromycin B resistance for selection and the constitutive *cdna*I promoter driving expression. The genes were transformed into *T. reesei* Δ *xyl*I, which is derived from strain QM9414 (ATCC 26921) (Stricker et al. 2006; Uzbas et al. 2012). All strains were stored as spore suspensions in 20% glycerol at -80°C.

Protoplast preparation and transformation were performed essentially as described (Gruber et al. 2000). *T. reesei* Δ *xyl*I was grown on plates of potato dextrose agar (PDA, Becton Dickinson, New Jersey) for 1 week. Conidia were separated by filtration through glass wool. Flasks (1 L) containing 250 mL potato dextrose broth were inoculated with spores (10^6 per mL) and grown at 28°C and 250 rpm for 14 h. The mycelia were filtered through Miracloth (Calbiochem, San Diego, CA) and digested in 10 mL protoplasting solution, which was

composed of 1.2 M KCl, 20 mg/mL Driselase, and 20 mg/mL lysing enzyme. The digestion conditions were 30-32°C with shaking at 50 rpm for 5-7 h. Protoplasts were purified by filtering through cheesecloth and 0.45- μ m nylon mesh and pelleted by centrifugation for 7 min, 4400 x *g* at 4°C. Protoplasts were washed three times with stabilization buffer (10 mM Tris·HCl, 5 mM CaCl₂, 1.2 M sorbitol, pH 7.5) and resuspended at a concentration of $\sim 10^8$ protoplasts per mL. Transformed protoplasts were plated on regeneration medium consisting of PDA plus 20% sucrose. After overnight growth, potato dextrose broth containing 0.9% agar and hygromycin B (Calbiochem) was overlaid on the plate to a final concentration of 50 μ g/mL hygromycin B. After one week, colonies were transferred to PDA + 50 μ g/mL hygromycin + 0.1% Triton X-100 and single spores isolated. To check transformant stability, the resulting mycelia were transferred sequentially three times to PDA plates without hygromycin B and then transferred back to PDA plates containing hygromycin B.

Five stable transformants of each gene construct, in duplicate, were grown and compared for cellulase expression. The cellulase induction medium contained 10 g/L glucose (Glc), 2 g/L KH₂PO₄, 1.4 g/L (NH₄)₂SO₄, 1 g/L peptone, 0.4 g/L CaCl₂·2H₂O, 0.3 g/L MgSO₄·7H₂O, 5 mg/L FeSO₄·7H₂O, 2 mg/L CoCl₂·2H₂O, 1.6 mg/L MnSO₄·H₂O, 1.4 mg/L ZnSO₄·7H₂O, and 5 mM urea, pH 6.5 (Uzbas et al. 2012). Tween-20 was omitted because polyoxyethylenes strongly interfere with mass spectrometry-based proteomics. Flasks (250 mL) containing 50 mL medium were inoculated with 5×10^6 spores and incubated at 28°C, 250 rpm. Cultures were sampled for endo- β 1,4-glucanase activity using CMC as substrate every 4 hr from 20 to 44 hr and every 12 h from 44 to 96 hr. Glc assays of the medium indicated that under these growth conditions the free Glc was completely consumed by 20 hr. The strains showing consistently highest enzyme activity were selected for further analysis.

Enzyme production and purification

Spores (2.5×10^8) from plates of the *T. reesei* $\Delta xyr1$ transformants growing on PDA + hygromycin B were used to inoculate 1-L flasks, each of which contained 250 mL cellulase induction medium, and grown with shaking at 250 rpm for 24-26 hr at 28°C. The mycelial mass was removed by filtration through Miracloth and the culture filtrate concentrated to 15 mL by rotary evaporation under vacuum at 35°C. The concentrated culture filtrate was exchanged into buffer composed of 50 mM NaH₂PO₄, 300 mM NaCl, and 3 mM imidazole, pH 8, on a 10DG desalting column (Bio-Rad, Richmond, CA) and the enzymes then purified by metal affinity chromatography. The enzyme solution was incubated with 2 mL Ni-NTA agarose (Qiagen, Valencia, CA) for 1 hr at 4°C, the agarose washed twice with 8 mL wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, and 6 mM imidazole, pH 8), and the enzyme eluted twice with 1 mL elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, and 400 mM imidazole, pH 8). The purified enzymes were buffer-exchanged into 50 mM sodium acetate, pH 4.8, and stored in aliquots at -80°C. Protein stock concentrations varied from 0.1 to 0.3 mg/mL.

Protein analysis

For SDS-PAGE, 20 μ L of each enzyme sample was analyzed under denaturing conditions on a 4-12% acrylamide gel (Novex, Thermo Fisher, Waltham, MA). Molecular weight markers were Perfect Protein (EMD Millipore, Darmstadt, Germany). Protein concentration was measured with bicinchoninic acid (Sigma-Aldrich) using bovine serum albumin as the standard.

Protein deglycosylation

Proteins were chemically deglycosylated with trifluoromethanesulfonic acid (TFMS) (Sigma-Aldrich) (Sojar and Bahl 1987). Each lyophilized enzyme sample (0.1 mg) was treated

with 45 μ L anhydrous TFMS precooled to -20°C for 1 h under nitrogen on ice with shaking every 10 min. The mixtures were kept on ice and neutralized by slow addition of 120 μ L 60% pyridine (Sigma-Aldrich) pre-cooled to -20°C . Subsequently, the deglycosylated enzymes were diluted and buffer-exchanged to 50 mM sodium acetate, pH 4.8, with a 2-mL ZebaTM 7-kDa MWCO spin desalting column (Thermo Scientific).

Preparation of lignin-rich residues

Lignin-rich residues were prepared from alkaline hydrogen peroxide (AHP)-pretreated corn stover (Palonen et al. 2004; Rahikainen et al. 2011). Biomass was treated with a 3:1 mixture of two commercial cellulases (CTec2 and HTec2; 60 mg total protein/g glucan) in 50 mM sodium acetate, pH 4.8, plus 0.02% sodium azide. The biomass enzyme mixture was incubated at 50°C with shaking at 160 rpm. The enzyme was replaced with fresh enzyme every 24 h. After 72 h, the insoluble residues were collected by centrifugation at $6400 \times g$ for 15 min and washed three times with water adjusted to pH 2.5 with 1 N HCl. The material was then heated at 90°C for 30 min to inactivate the enzymes and digested with 1 mg/mL Pronase at 37°C in 0.1 M Tris and 10 mM CaCl_2 , pH 7.5, for 48 h to remove the bound enzymes. The residue was heated at 90°C to inactivate the protease and centrifuged ($6400 \times g$) to pellet the undigested material, which was then washed three times with water adjusted to pH 2.5, and lyophilized. This material is referred to as lignin. Lignin inhibition was tested at 10 mg/mL lignin, 10 mg/mL CMC, and 1 $\mu\text{g}/\text{mL}$ of each GH5_5 enzyme. Reactions were carried out in 50 mM sodium acetate, pH 4.8, and 50°C with rotation at 10 rpm, in triplicates, for 5 to 30 min. Released reducing sugars were determined with 3,5-dinitrosalicylic acid (DNS) using Glc as the standard (Miller 1959).

Enzyme assays

The standard assay conditions for the GH5_5 endo-glucanases were 0.5-1 $\mu\text{g}/\text{mL}$ enzyme, 10 mg/mL CMC, 50 mM sodium acetate, pH 4.8, 5 to 30 min assay time, 50°C, in triplicates. Reducing sugar production was measured using the DNS assay. The reactions (60 μL assay mixture and 60 μL DNS solution) were incubated in a 96-well PCR plate in a thermocycler at 95°C for 5 min. After cooling to room temperature on ice, 100 μL was transferred to a flat-bottom 96-well microplate and absorbance measured at 540 nm on a microplate reader (Uzbas et al. 2012).

Enzyme activity on insoluble cellulosic substrates was assayed using a modified filter paper assay (Xiao et al. 2004). Endo-glucanase concentration was 12 $\mu\text{g}/\text{mL}$, and, when added, the other enzyme concentrations were 12 $\mu\text{g}/\text{mL}$ CBH1, 4 $\mu\text{g}/\text{mL}$ CBH2, and 2 $\mu\text{g}/\text{mL}$ BG. Reactions contained 60 μL 50 mM sodium acetate, pH 4.8, plus one 0.7 cm disc of Whatman No. 1 filter paper. Reaction conditions were 50°C for 1 h, in triplicates. Reducing sugars in the supernatant were measured with DNS.

In the tests of substrate range, the enzyme concentration was 5 $\mu\text{g}/\text{mL}$. Oat β -glucan and β 1,4-mannan concentrations were 5 mg/mL and glucomannan and galactomannan were 2.5 mg/mL. Reactions were conducted in triplicates in 1.5 mL centrifuge tubes for the appropriate times to establish a linear reaction rate (5 min to 6 h, depending on substrate) at pH 4.8 and 50°C with rotation at 10 rpm.

For biomass digestion experiments, the AHP-pretreated corn stover concentration was 2 mg glucan/mL (equivalent to 5.8 mg total biomass/mL). Reactions (total volume 0.5 mL) contained 50 mM sodium acetate, pH 4.8, plus 0.02% sodium azide in 96-well deep-well plates. Total enzyme loading was 15 mg/gm glucan (30 $\mu\text{g}/\text{mL}$). The enzyme cocktail contained 21

µg/mL of the seven components of the 8-CSM plus 9 µg/mL of the GH5_5 endo-glucanase being tested. Each assay was run in quadruplicate at 50°C. Free Glc in the supernatant was measured by enzyme-linked colorimetry after 24, 48, and 72 hr (Banerjee et al. 2010b). Assays to measure the effect of galactomannan on enzyme digestion were conducted similarly. The galactomannan concentrations were 0, 0.05, 0.25, or 1 mg/mL and reaction time was 72 hr.

Results

*Expression of six GH5_5 enzymes in *T. reesei* Δ xyl1*

Earlier we showed that when compared side-by-side in fully-defined synthetic enzyme cocktails, the endo-β1,4-glucanase from *S. thermophile* (StCel5A) consistently produced more Glc from pretreated corn stover than its counterpart from *T. reesei*, TrCel5A (Ye et al. 2014; Figure 2.1). TrCel5A and StCel5A belong to clades 1 and 2, respectively, of subfamily 5 of glycoside hydrolase (GH) family 5 (GH5_5) (Aspeborg et al. 2012).

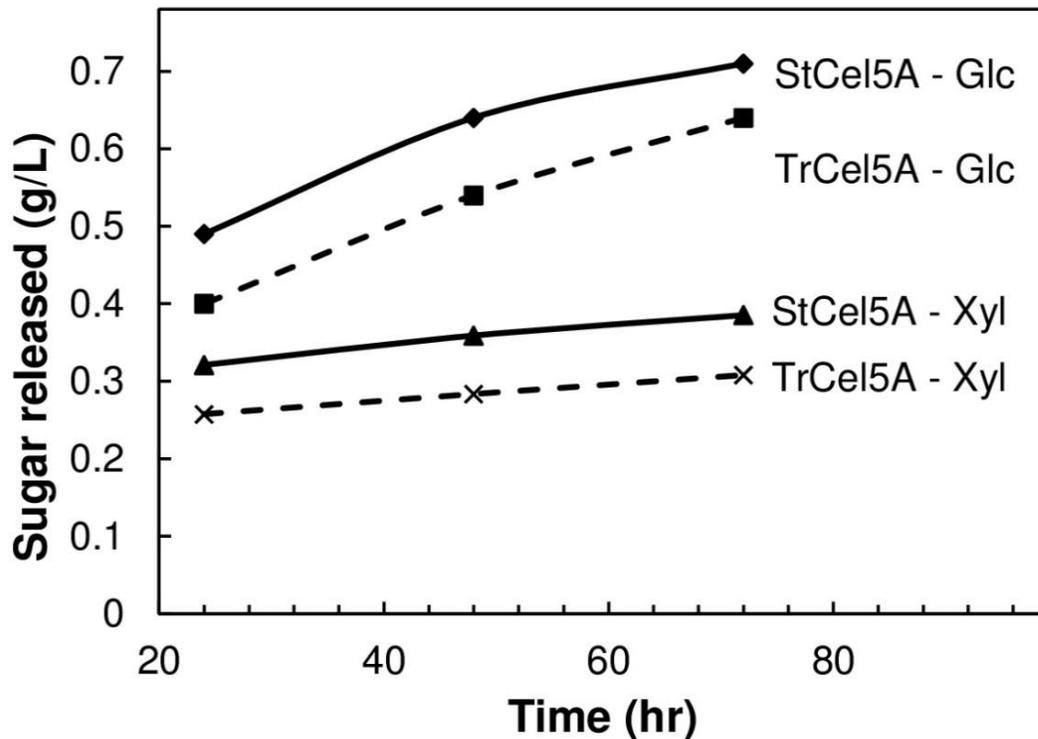


Figure 2.1 Yields of Glc and Xyl from corn stover as a function of time and source of Cel5A for 8-component synthetic enzyme mixture plus either StCel5A or TrCel5A. Reprinted from Ye et al. (2014) with permission.

In order to understand the basis of the superiority of StCel5A, we began by asking whether superiority was a general characteristic of subclade 2 of GH5_5. That is, in side-by-side comparisons, would endo-glucanases in clade 2 consistently outperform those in clade 1? To test this, additional members of the GH5_5_1 and GH5_5_2 clades were heterologously expressed in the identical host and subjected to comparative assays alone and in synthetic mixtures on both pure polysaccharide substrates and on real biomass. Two of them, HgCel5A and ApCel5A were, like TrCel5A and StCel5A, from fungi in the Ascomycota, and two others, PaCel5A and VvCel5A, were from the Basidiomycota. Phylogenetic analysis indicated that the two clades of GH5_5 diverged before these two major groups of fungi split evolutionarily (Ye et al. 2014). The

structural properties of the enzymes are described in Table 1. The discrepancies between predicted and observed molecular weights can be largely attributed to glycosylation (Figure 2.2).

In previous work, we used StCel5A obtained by heterologous expression in *Aspergillus niger* (Tambor et al. 2012; Ye et al. 2014). In the current study, all of the enzymes including TrCel5A and StCel5A were expressed in *T. reesei* $\Delta xyr1$ with a C-terminal His₆ tag. The gene *xyr1* encodes a transcription factor necessary for expression of most cellulases and xylanases, and thus *xyr1* knock-out strains have a low background of potentially interfering activities (Stricker et al. 2006). To suppress endogenous cellulase production even further, the heterologous genes were driven by the strong, constitutive promoter *cdna1* and the fungi were grown on Glc as sole carbon source (Uzbas et al. 2012). Culture filtrates of *T. reesei* $\Delta xyr1$ showed no activity against cellulase substrates including CMC, filter paper, p-nitrophenyl-lactoside, or p-nitrophenyl-glucoside, and only minor extraneous protein bands were visible by SDS-PAGE (Uzbas et al. 2012, and our data, not shown).

The enzymes were purified by immobilized metal affinity chromatography. Yields of the heterologous endo-glucanases after chromatography were 8-12 mg/L, except StCel5A, which yielded ~0.5 mg/L. Two proteins (ApCel5A and PaCel5A) had minor, lower molecular weight bands (Figure 2.2). These were determined by immunoblotting with an anti-His₆ antibody to be degradation products of the parent proteins and not co-purifying *T. reesei* proteins (data not shown). Significant glycosylation of all the proteins was demonstrated by shifts in apparent molecular weight after chemical deglycosylation (Figure 2.2).

Table 2.1 Structural characteristics of the six GH5_5 enzymes. In all proteins, the cellulose-binding module (CBM) and linker are at the N-terminus. The sequences of the three clade 1 enzymes (ApCel5A, PaCel5A, and TrCel5A), are 59–68% identical, the three clade 2 enzymes (HgCel5A, StCel5A, and VvCel5A) are 54–73% identical, and the two clades are 29–34% identical to each other.

Protein	ApCel5A	PaCel5A	TrCel5A	HgCel5A	StCel5A	VvCel5A
GH5_5 clade	1	1	1	2	2	2
Total length (aa)	430	417	418	388	389	389
Signal peptide (aa)	23	19	21	16	16	23
CBM (aa)	30	39	35	36	36	36
Linker peptide (aa)	51	31	34	31	32	23
Catalytic domain (aa)	326	328	328	305	305	307
Predicted MW^a (kD)	43	43	43	42	42	40
Apparent MW^b (kD)	65	57	58	58	58	46
N-glycosyl-ation sites^c	1	3	1	3	3	2
Disulfide bonds^d	5	6	6	3	3	3

aa = amino acid

^aAfter removal of the signal peptide

^bEstimated size from SDS-PAGE

^cPredicted by NetNGlyc 1.0 (Blom et al. 2004). The linkers are also expected to have *O*-glycosylation

^dPredicted by DiANNA (Ferrè and Clote 2006) and experimentally determined for TrCel5A (Lee et al. 2011)

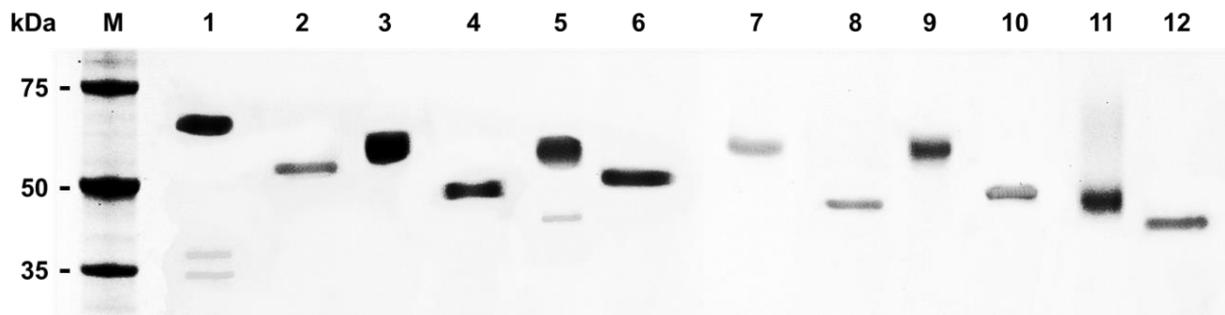


Figure 2.2 SDS-PAGE of six recombinant endo-β1,4-glucanases before (lanes 1, 3, 5, 7, 9, and 11) and after (lanes 2, 4, 6, 8, 10, and 12) chemical deglycosylation. Proteins were visualized with Coomassie blue staining. Lanes 1 and 2, ApCel5A; lanes 3 and 4, HgCel5A; lanes 5 and 6, PaCel5A; lanes 7 and 8, StCel5A; lanes 9 and 10, TrCel5A; lanes 11 and 12, VvCel5A. M, protein size markers (kilodaltons).

Enzyme activities on pure cellulose

The activities of the six endo-glucanases were first compared by themselves on CMC. Five of the six enzymes showed specific activities of ~32-63 $\mu\text{mol}/\text{min}/\text{mg}$, which is in the range of other GH5 endo- β 1,4-glucanases (Figure 2.3a). For example, HgCel5A expressed in *Aspergillus oryzae* has a specific activity of 32 $\mu\text{mol}/\text{min}/\text{mg}$ (Takashima et al. 1997), whereas the same enzyme produced in *T. reesei* had a specific activity of 46 $\mu\text{mol}/\text{min}/\text{mg}$ (Figure 2.3a). However, ApCel5A had a specific activity of 295 $\mu\text{mol}/\text{min}/\text{mg}$, 5-9 times higher than the others. This was confirmed by multiple assays with different batches of ApCel5A. It was not due to contamination with other synergistic enzymes such as CBH1, because the *T. reesei* ΔxyrI strain does not make other cellulose-active enzymes (Stricker et al. 2006; Uzbas et al. 2012) and because ApCel5A and all of the other GH5_5 preparations lacked any detectable background activity with the CBH1-specific substrate *p*-nitrophenyl- β -lactoside (data not shown).

When assayed by themselves on filter paper, the six enzymes had weak activity (9 to 23 μmol reducing sugar released per hour per mg enzyme), which is typical for endo-glucanases when assayed by themselves on insoluble substrates, and there was no pattern among clades (Figure 2.3b). Whereas ApCel5A was 5-9 times more active than the other enzymes on CMC, its activity was only 1.2-2.6 times higher on filter paper (Figure 2.3b). When the endo-glucanases were combined with a suite of other enzymes necessary for efficient conversion of cellulose to free Glc (i.e., CBH1, CBH2, and BG), sugar release was enhanced 3 to 4-fold, as expected because of the known synergistic interactions of exo- and endo-acting cellulases. Again, however, the six enzymes had similar specific activities, and there were no clear patterns differentiating clade 1 from clade 2. Notably, StCel5A was not significantly superior to TrCel5A

when tested alone on CMC, alone on filter paper, or in combination with the other cellulases on filter paper (Figure 2.3).

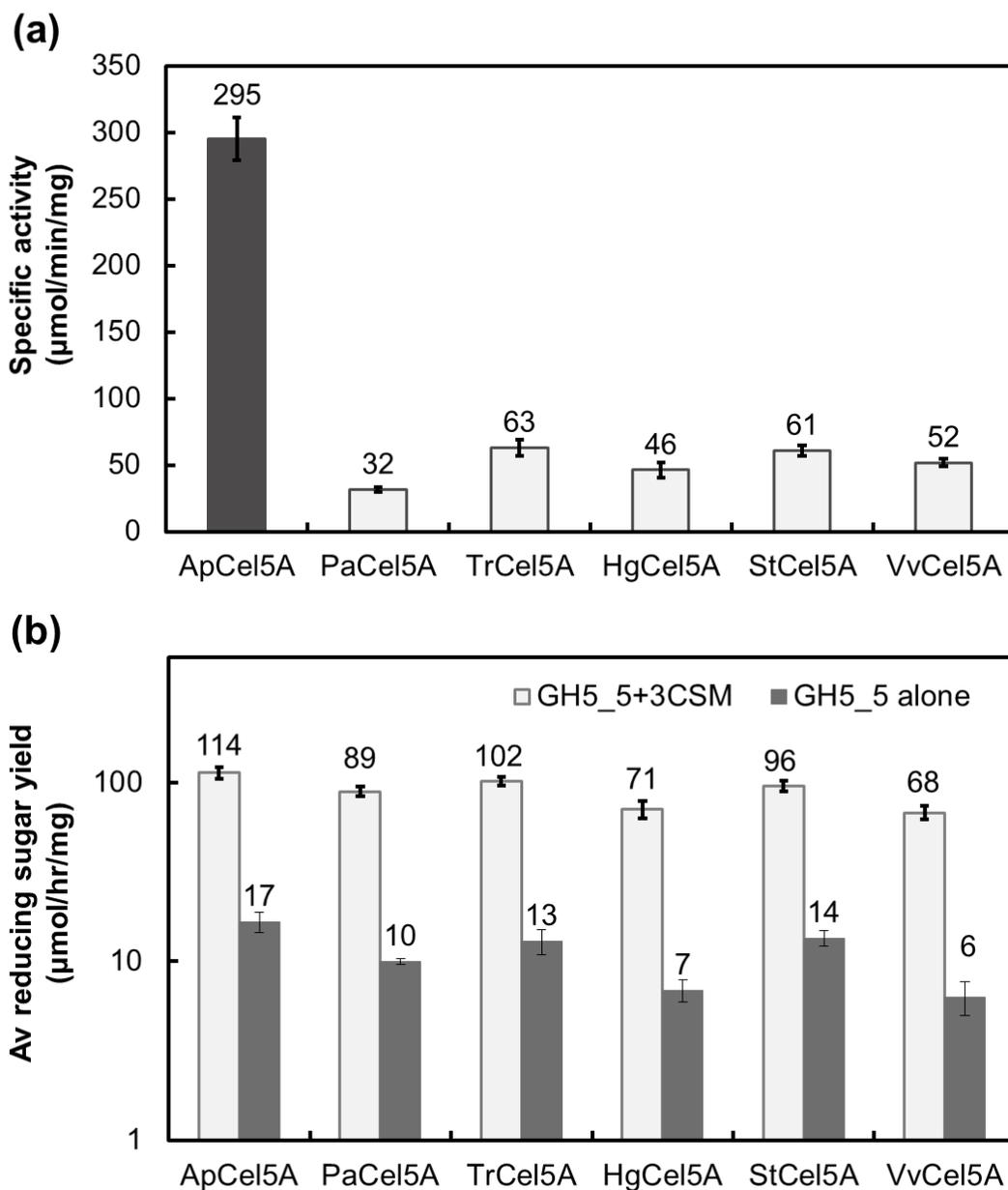


Figure 2.3 Activities of six endoglucanases on cellulosic substrates. (a) Specific activities of the six GH5_5 enzymes on CMC. (b) Activities of the six enzymes on filter paper, either alone (right bars) or when mixed with a synthetic mixture (3-CSM) containing CBH1, CBH2, and BG (left bars). The 3-CSM mixture without any added GH5_5 released $12.4 \pm 1.8 \mu\text{mol}/\text{hr}/\text{mg}$ (± 1 standard deviation, $n = 6$).

Hydrolysis of pretreated corn stover

The six GH5_5 endoglucanases were next compared against each other as part of an eight-component synthetic enzyme mixture on pretreated corn stover. The enzyme proportions in the mixture were based on our previous studies (Banerjee et al. 2010b). As found before, mixtures in which StCel5A was the only endo- β 1,4-glucanase outperformed mixtures in which TrCel5A was the only endoglucanase, at 48 and 72 hr (Figure 2.4). However, this superiority was not shared by the other enzymes in clade 2 (i.e., HgCel5A and VvCel5A). TrCel5A, which is in clade 1, produced more Glc than both of the other clade 1 enzymes as well as the alternate clade 2 enzymes at 48 and 72 hr (Figure 2.4).

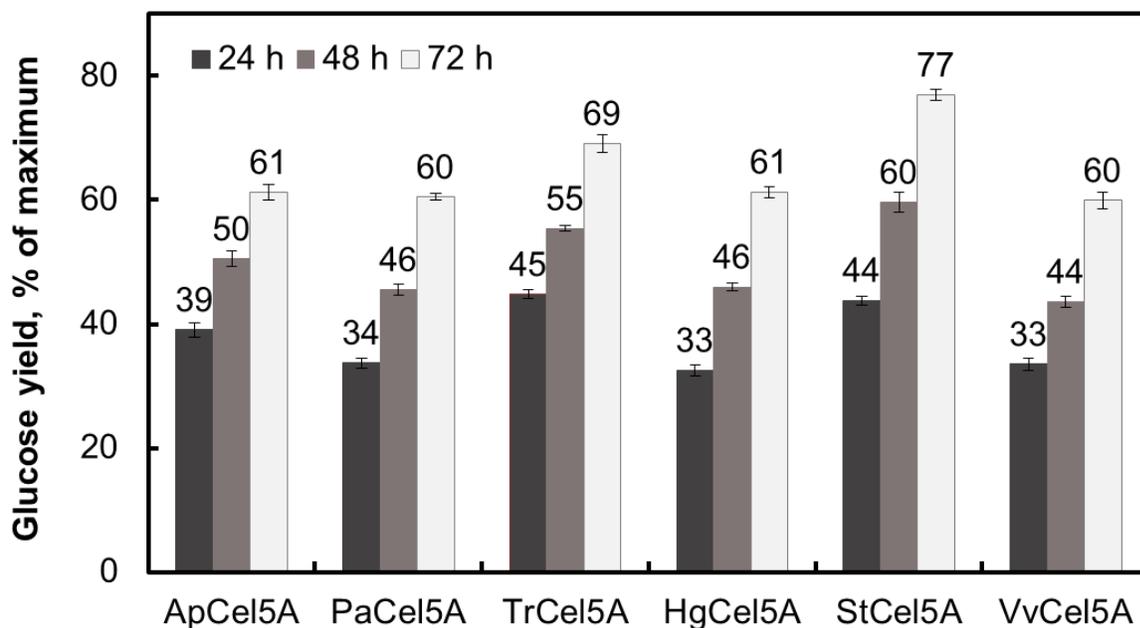


Figure 2.4 Yield of Glc from pretreated corn stover treated with a seven-component synthetic mixture plus different GH5_5 endo-glucanases at equal concentration. The base mixture contained CBH1, CBH2, BG, EX2, EX3, BX, and AA9. Total enzyme loading including GH5_5 was fixed at 15 mg/g glucan. Error bars indicate the standard deviation, n =4. The unsupplemented 7-component synthetic mixture released 54% \pm 1% of maximum theoretical Glc from corn stover after 72 hr incubation.

On corn stover, the activity of ApCel5A as part of a mixture was similar to PaCel5A, HgCel5A, or VvCel5A at 48 and 72 hr (Figure 3). Thus, the activities of the six enzymes on CMC or filter paper did not predict their activities in mixtures on corn stover.

Substrate specificities of six endo-glucanases

Previous research indicated that StCel5A has activity on β -glucan, mannan, galactomannan, xylan, and arabinoxylan, whereas of these substrates TrCel5A is only active on β -glucan (Ye et al. 2014; Figure 2.5). To test whether substrate specificity of the six endo-glucanases was correlated with either clade or superiority, the six enzymes were tested against β -glucan, glucomannan, β 1,4-mannan, and galactomannan. ApCel5A showed much stronger (4 to 22-fold higher) activity than the others on β -glucan (Figure 2.6), which is consistent with the results with CMC, since both substrates are soluble β 1,4-glucans. All six enzymes showed activity against glucomannan, which is an unbranched polymer of β 1,4-linked mannose and β 1,4-linked Glc (Katsuraya et al. 2003). HgCel5A showed the weakest activity on β -glucan but the strongest activity on glucomannan (Figure 2.6). Although ApCel5A was much more active than the others on oat β -glucan and CMC, it was less active than HgCel5A and StCel5A on glucomannan. All three enzymes in clade 2 but none of the clade 1 enzymes showed activity against mannan and galactomannan. HgCel5A was the most active on both substrates while VvCel5A showed the weakest activity on both (Figure 2.6). The activities of VvCel5A were 15-fold and 5.5-fold less than the activities of HgCel5A on mannan and galactomannan, respectively.

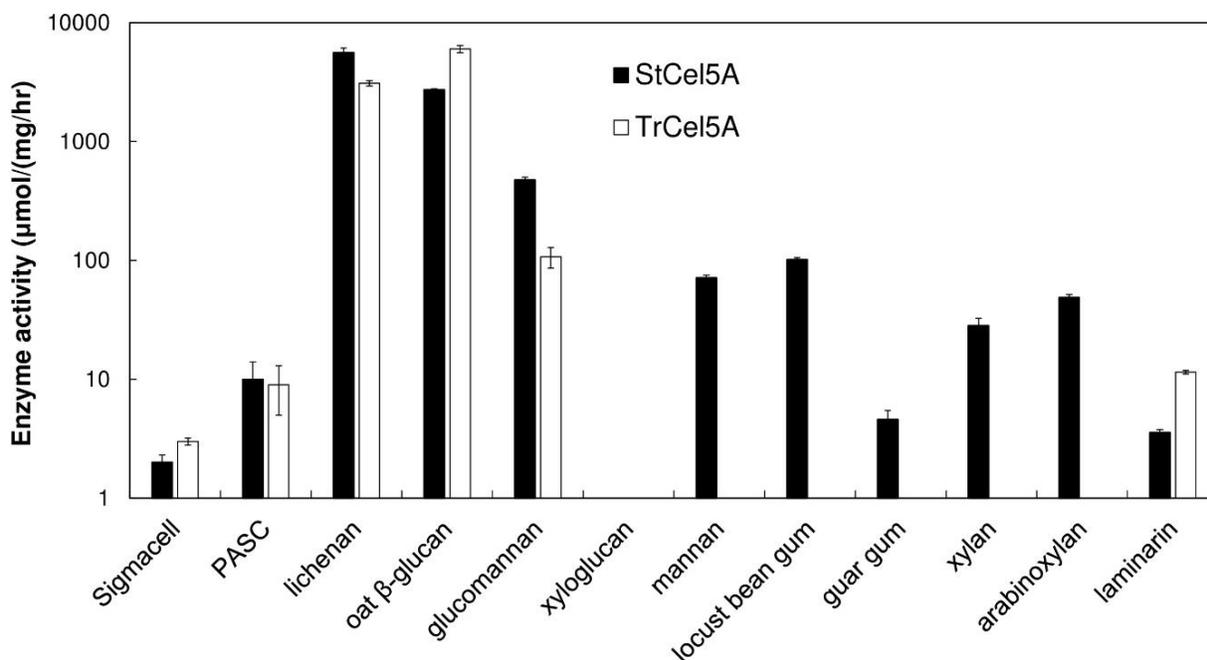


Figure 2.5 Activities of pure TrCel5A and StCel5A on defined substrates. Sigmacell is crystalline cellulose; PASC is phosphoric-acid swollen cellulose; lichenan and β -glucan are two types of mixed-linkage β 1,3- β 1,4 glucan, one from Iceland moss and the other from oats; xyloglucan is β 1,4-linked glucan with α 1,6-linked Xyl side chains from tamarind; mannan is insoluble β 1,4-mannan; locust bean gum and guar gum are two types of β 1,4-linked mannan with side chains of α 1,6-linked galactose; xylan is β 1,4-linked xylan; arabinoxylan is β 1,4-linked xylan with α 1,6-linked arabinose side chains; laminarin is predominantly β 1,3-linked glucan. Shaded bars, StCel5A; open bars, TrCel5A. TrCel5A had no detectable activity against mannan, either galactomannan, xylan, or arabinoxylan. Neither enzyme had detectable activity against xyloglucan. Experiment performed by Bingyao Li; reprinted from Ye et al. (2014) with permission.

The activities of StCel5A and TrCel5A on mannan and galactomannan are consistent with previous results (Ye et al. 2014; Figure 2.5). However, none of the enzymes used in the current work showed activity on xylan or arabinoxylan (data not shown), which is inconsistent with the earlier results (Ye et al. 2014; Figure 2.5). The probable explanation for this discrepancy is that the StCel5A used in the previous work was obtained by heterologous expression in wild type *A. niger* and might therefore have been contaminated with *A. niger* xylanase (Tambor et al. 2012). This is less likely to be the case for the StCel5A used in the current work because *xyl1* is

required for xylanase expression (Stricker et al. 2006) and the heterologous host was grown on Glc. Therefore, our earlier hypothesis that xylanase activity might explain the superiority of StCel5A over TrCel5A was not supported in the present study.

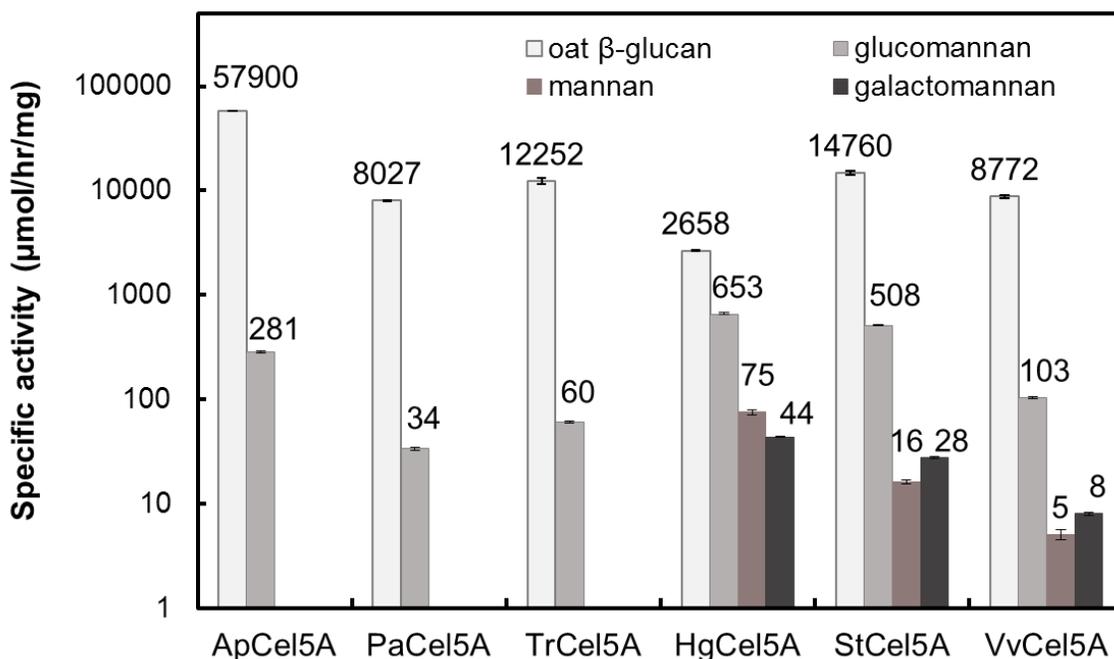


Figure 2.6 Activity of six GH5_5 enzymes (reducing sugar assay) by themselves on different polysaccharides (logarithmic scale). Enzyme concentrations were 5 $\mu\text{g}/\text{mL}$. The three enzymes on the left are in clade 1 and the three on the right are in clade 2. Reaction times were adjusted to ensure that all enzyme activities were in the linear range. Activities of ApCel5A, PaCel5A, and TrCel5A on mannan and galactomannan were zero.

Effect of lignin on GH5_5 enzyme activity

To understand why the performance of the six enzymes on biomass deconstruction did not correlate with their activities on pure cellulose, some potential factors were tested. Lignin is known to have a negative effect on enzymatic biomass deconstruction (Berlin et al. 2006; Rahikainen et al. 2011, 2013b). If the six GH5_5 enzymes were differentially sensitive to lignin inhibition, it could influence their performance in corn stover deconstruction. To test this

hypothesis, the activities of the six endo-glucanases were tested on CMC with and without the addition of lignin extracted from pretreated corn stover.

When tested in isolation on CMC, all six enzymes were somewhat inhibited by 10 mg/mL lignin (Figure 2.7). Lignin reduced VvCel5A activity by 32% and the other activities by 12-16% (Figure 2.7). Only the reductions of ApCel5A and VvCel5A were statistically significant ($p < 0.05$). There was no correlation between clade and degree of sensitivity to lignin, nor was StCel5A less sensitive than TrCel5A. Therefore, reduced inhibition by lignin is probably not the explanation for the superiority of StCel5A over TrCel5A and the other enzymes tested here.

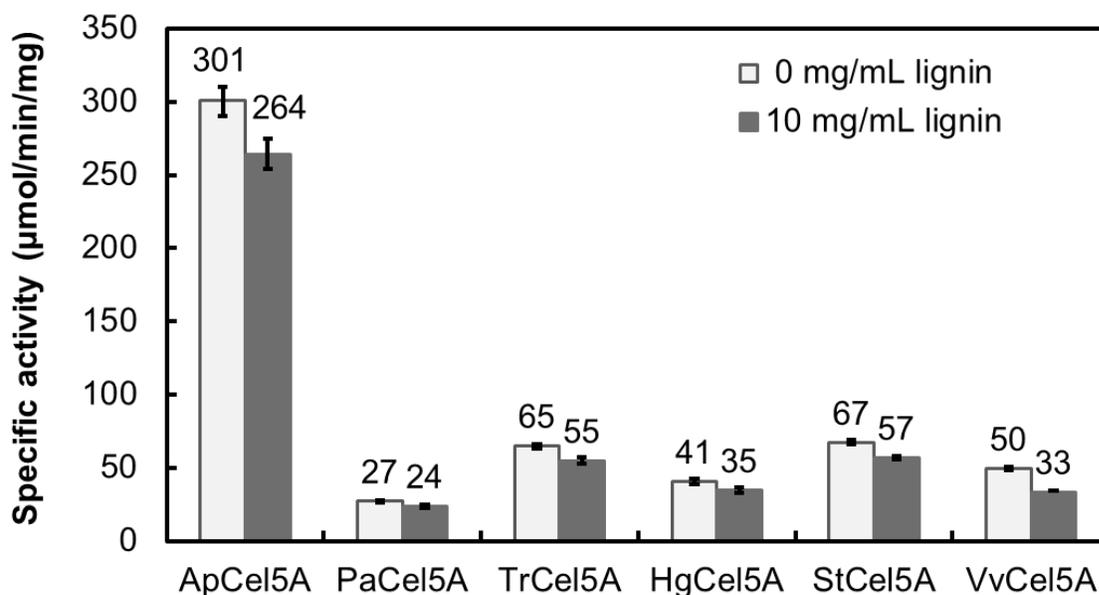


Figure 2.7 Specific activities (reducing sugar assay) of six GH5_5 enzymes on CMC with and without addition of lignin. CMC concentration was 10 mg/mL and enzyme concentration was 1 µg/mL.

Influence of mannan on enzymatic hydrolysis of corn stover

Clade 2 but not clade 1 endo-glucanases have low but clear β -mannanase activity (Figure 2.6). Mannan and manno-oligosaccharides were reported to be strong inhibitors of cellulases,

especially cellobiohydrolase 1 (CBH1) (Kumar and Wyman 2014; Xin et al. 2014). The corn stover used in the current study contained 0.6% mannan (Banerjee et al. 2010d). To test whether the mannanase activity of StCel5A contributes to its superiority by hydrolyzing inhibitory mannans and manno-oligosaccharides in corn stover, the activity of the seven-component mixture plus different GH5_5 endo-glucanases was measured in the presence of galactomannan. (Galactomannan was used instead of mannan because it is water-soluble). In the absence of galactomannan, the enzymes showed the same relative activities as before, i.e., StCel5A gave higher Glc yields than the other endo-glucanases, followed by TrCel5A (Figure 2.4, Figure 2.8a). The six endo-glucanases showed strong differences in sensitivity to galactomannan (Figure 2.8). Figure 2.8a shows the absolute Glc release and Figure 2.8b shows the same data plotted relative to the maximum activities of each enzyme mixture in the absence of galactomannan. The different mixtures fell into two groups, one of which was inhibited only 15-22% by 1 mg/mL galactomannan and the other of which was inhibited by 58-62% (Figure 2.8b). The sensitive endo-glucanases were significantly inhibited (~30%) even by 0.25 mg/mL galactomannan. The two most galactomannan-resistant endo-glucanases were in clade 2 (Figure 2.8b). The third clade 2 endo-glucanase, VvCel5A, was as sensitive as the three clade 1 enzymes. The behavior of VvCel5A might be related to the fact that it had the weakest mannanase activity of the clade 2 enzymes (Figure 2.6).

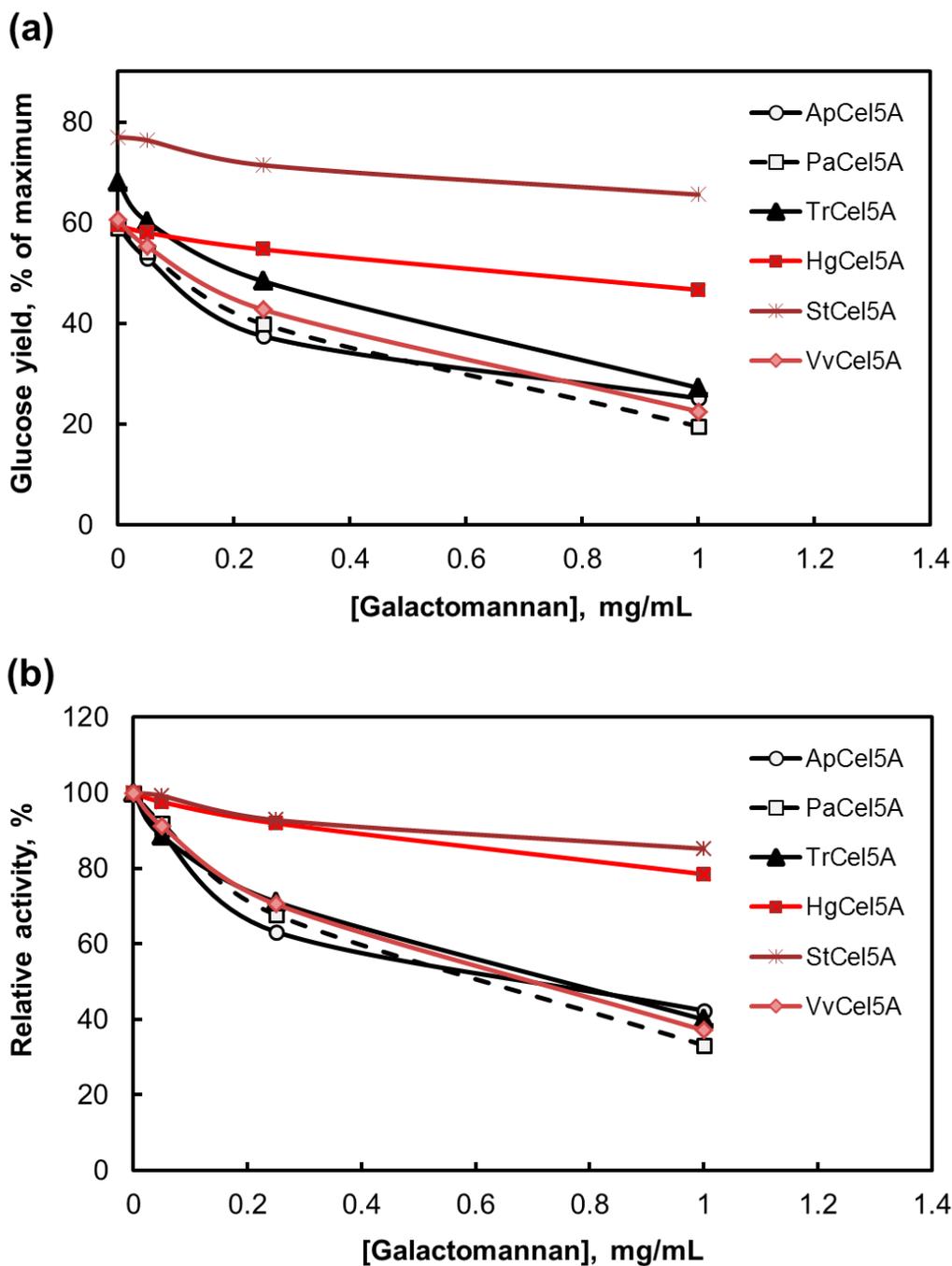


Figure 2.8 Inhibition by galactomannan of Glc release from pretreated corn stover by a seven-component synthetic enzyme mixture plus different GH5_5 endo-glucanases. (a) Glc yield after 72 h hydrolysis as a function of galactomannan concentration. **(b)** Same data normalized to 100% for the activity of each GH5_5 in the absence of galactomannan. Clade 1 enzymes are in black and clade 2 enzymes are in red. Standard deviation bars are not visible because they were smaller than the data markers.

Discussion

Six endo- β 1,4-glucanases of clades 1 and 2 of GH5_5 were expressed in a common heterologous host with low cellulase background. Of the three enzymes from clade 2 that were tested, only StCel5A was superior in terms of specific activity. Therefore, superiority is not a general trait of clade 2 enzymes but is a unique property of StCel5A. Xylanase activity, hypothesized earlier to account for the superiority of StCel5A over TrCel5A, was excluded in the current study as contributing to superiority because none of the six enzymes showed any xylanase activity when expressed in *T. reesei* Δ *xyl1*. Differential lignin sensitivity was also ruled out as contributing to superiority.

One of the enzymes, ApCel5A, was much more active than the others when assayed alone on CMC or β -glucan but not when assayed in combination with other synergistic enzymes on filter paper or corn stover. No apparent reasons for this emerged from our studies, but structurally a major difference between ApCel5A and the other tested enzymes is its exceptionally long linker region between the catalytic domain and the CBM. The ApCel5A linker consists of 51 amino acid residues while the other five linkers are 23-34 residues. Another possible explanation for why ApCel5A did not perform better in an eight-component synthetic enzyme mixture on corn stover is that other enzymes are limiting under these conditions.

Two of the three clade 2 enzymes (StCel5A and HgCel5A) were less inhibited by galactomannan compared to the clade 1 enzymes. However, even though StCel5A was less inhibited by mannan compared to TrCel5A, mannanase activity and mannan inhibition did not correlate with superiority. One clade 2 enzyme (VvCel5A) was sensitive to galactomannan, but this enzyme also had the weakest β -mannanase activity. Despite the fact that mannose is a minor

sugar in corn stover and other cereal biomass, mannans and oligo-mannosides are strong inhibitors of cellulase action (Kumar and Wyman 2014; Xin et al. 2014). As shown by GENPLAT, β -mannanase is the second most important enzyme (after amylase) for releasing free Glc from corn dried distillers' grains (DDG), despite DDG containing only 2.5% mannose (Banerjee et al. 2010d). Therefore, it appears that mannans have a stronger influence on cellulase digestion of biomass than their low levels would suggest. Possible explanations for the earlier and current results on the importance of mannan and β -mannanase is that the small amount of mannan occludes access of cellulases to cellulose (Wilkie 1979), or that mannanase activity relieves inhibition of cellulases by mannan and manno-oligosaccharides (Kumar and Wyman 2014; Xin et al. 2014). The latter hypothesis implies that all of the endo- β -glucanases tested in the current paper, and perhaps all family GH5 enzymes, are intrinsically sensitive to inhibition by mannans and/or manno-oligosaccharides.

Three conclusions can be drawn from the current work. First, even within one subfamily of glycoside hydrolases there can be significant natural variation in enzymatic behavior relevant to biomass deconstruction. This suggests that there is a high degree of functional diversity even among enzymes with high primary structure similarity, which it should be possible to exploit for the discovery and improvement of enzymes for industrial applications such as biomass conversion. Second, the six enzymes showed significant differences in behavior on pure substrates compared to biomass, and when tested alone compared to being combined into complete cellulase mixtures. This suggests that screening enzymes for key enzymatic traits, including specific activity, using pure enzymes on model substrates may not be a reliable strategy to find better enzymes. A third conclusion is that minor polysaccharide components in biomass might in some circumstances have a disproportionate influence on deconstruction

efficiency, and therefore accessory enzymes will be important components of effective enzyme mixtures.

CHAPTER 3: *CLADORRHINUM BULBILLOSUM* ALKALIPHILIC CELLULASES FOR BIOMASS DECONSTRUCTION

I would like to acknowledge Dr. Dina Jabbour for the screening, isolation and identification of the fungus *Cladorrhinum bulbiliosum*, and the contribution of Figure 3.2.

Abstract

Currently available cellulases for biomass deconstruction have a narrow pH optimum around pH 5. The fungus *Cladorrhinum bulbilosum* was previously isolated from soil samples and shown to degrade cellulose at pH 9. The genome and transcriptome under different growth conditions of *C. bulbilosum* were sequenced by the Department of Energy Joint Genome Institute. Compared to *T. reesei*, *C. bulbilosum* has more endo- and exo-cellulases in families GH6 and GH7 and endo-xylanases in GH10 and GH11 and many more lytic polysaccharide monooxygenases of family AA9 than *T. reesei* (35 compared to 4). An endo- β 1,4-glucanase (CbCel7E) and a β -glucosidase (CbCel3A) were purified from *C. bulbilosum* cultures and their pH profiles compared against their *T. reesei* homologs (TrCel7B and TrCel3A, respectively). CbCel7E and CbCel3A showed higher activities than their *T. reesei* homologs at pH 6-8, although they were less active than TrCel7B and TrCel3A at the optimal pH of each enzyme. Both TrCel7B and TrCel3A showed higher K_{cat} than their *C. bulbilosum* counterparts at their optimal pHs. Since non-productive binding of cellulases to lignin is reduced at higher pH, *C. bulbilosum* alkaliphilic cellulases are promising for increasing the enzyme efficiency of biomass conversion and for other enzyme applications performed at high pH.

Introduction

Lignin in lignocellulosic material adsorbs to and inhibits cellulases, hence reducing the extent and efficiency of biomass saccharification (Berlin et al. 2006; Palonen et al. 2004; Ooshima et al. 1990; Rahikainen et al. 2013a). The binding of cellulases to lignocellulose is

influenced by reaction pH – the higher the pH, the lower the non-productive binding (Lou et al. 2013; Rahikainen et al. 2013b). *Trichoderma reesei*-based commercial cellulase mixtures such as Cellic® CTec2 achieve their maximal activity at pH 5, where the nonproductive binding to lignin is approximately 1.8 mg protein per gram lignin (Lou et al. 2013). CTec2 adsorption to lignin decreases to about 0.7 mg protein per gram lignin at pH 6, but the enzyme mixture loses 40% of its activity (Cellic® CTec2 product brochure, Novozymes; Lou et al. 2013). To reduce non-productive binding of cellulases to lignin and thereby improve enzyme efficiency, it would be desirable to develop alkaliphilic cellulase systems that are active above pH 6, where nonproductive cellulase binding to lignin is reduced (Rahikainen et al. 2013b).

There are several other potential advantages to alkaliphilic cellulase systems: 1) a broader pH profile would eliminate the requirement for precise pH control during enzyme digestion; 2) enzyme digestion at higher pH would be more compatible with alkaline pretreatments; 3) enzyme digestion at higher pH would reduce microbial contamination (Beckner et al. 2011). Besides biomass degradation, alkaline cellulases could have applications in laundry detergents and the paper and pulp industry (Horikoshi 1999; Ito 1997; Kenealy and Jeffries 2003; Wang et al. 2017).

A few studies have reported alkaliphilic cellulases from bacteria and fungi such as *Bacillus* species, *Marinobacter* sp., *Aspergillus niger* and *Volvariella volvacea* (Fukumori et al. 1985; Hong et al. 2001; Li et al. 2011; Horikoshi et al. 1984; Shanmughapriya et al. 2010). However, no alkaline cellulase mixtures have been studied in particular for their potential use in lignocellulosic biomass deconstruction. To obtain an alkaliphilic cellulase mixture, one strategy is to isolate the individual major cellulases from fungi that are able to produce cellulases at alkaline pH, and then to construct a synthetic cocktail (Banerjee et al. 2010b,c,d) in which the

components are in their optimal proportions. To release glucose from lignocellulosic substrates, such a cocktail must contain at least one cellobiohydrolase 1 (CBH1), one endo- β 1,4-glucanase (EG), and one β -glucosidase (BG). Other enzymes that are probably important, depending on the source of the biomass, include additional CBH's (e.g., CBH2), auxiliary enzymes such as AA9, xylanases, and xylosidases (Banerjee et al. 2010b,c,d).

In a search for alkaliphilic enzymes, previous lab member Dina Jabbour isolated and described a fungus *Cladorrhinum bulbillosum*. She showed that it could grow on lignocellulose better than *T. reesei* at pH 9-10 and that the bulk cellulase activity of its crude culture filtrates was better than *T. reesei* at high pH. The goal of my study was to annotate the cell wall degrading enzymes in the genome and transcriptome of *C. bulbillosum*, and to characterize in greater depth the individual cellulases produced by *C. bulbillosum*, with the ultimate goal of making a synthetic mixture of cellulases and accessory enzymes that would show a broader pH profile compared to *T. reesei*.

Materials and methods

Prediction of carbohydrate-active enzymes in the C. bulbillosum, T. reesei and P. anserina genomes

The Carbohydrate-Active enZymes (CAZymes) in the *C. bulbillosum* were predicted using dbCAN (Yin et al. 2012). For comparison, the same program was run with the genomes of *Trichoderma reesei* and *Podospora anserina*, which is in the same family as *C. bulbillosum*. The genome sequences were searched against a database containing catalytic domain models of all CAZyme family members. A protein was considered to be a CAZyme if the aligned sequence covered more than 30% of a domain model and the E-value was $< 1 \times 10^{-5}$. If the aligned sequence

covered more than 30% of a model but consisted of less than 80 amino acid residues (e.g., carbohydrate-binding modules), the criterion of an E-value $< 1 \times 10^{-3}$ was used to consider a protein as a CAZyme (Yin et al. 2012). The annotations have been inserted into the *C. bulbiliosum* record at JGI using the community annotation system (<https://genome.jgi.doe.gov/Clabul1/Clabul1.home.html>).

Fungus growth

Cladorrhinum bulbiliosum was maintained on V8 juice agar (163 ml V8 juice plus 5 gm CaCO₃ in one liter water). *C. bulbiliosum* cellulase induction medium contained 0.5 g/L KH₂PO₄, 0.1 g/L MgSO₄ · 7H₂O, 0.1 g/L NaCl, 5 g/L yeast extract, 10 mg/L FeSO₄ · 7H₂O, 10 mg/L MnSO₄, 10 mg/L ZnSO₄, and 10 g/L Avicel pH-101 cellulose (Sigma-Aldrich, St. Louis, MO; Kumar et al. 2000).

Enzymes, substrates and buffer system

Endo- β 1,4-glucanase (TrCel7B; AAA34212) and β -glucosidase (TrCel3A; AAA18473) from *T. reesei* were expressed in *Pichia pastoris* as described earlier (Banerjee et al. 2010b,c). 4-Nitrophenyl- β -D-cellobioside, 4-nitrophenyl- β -D-glucoside, and 4-nitrophenyl- β -D-lactoside were purchased from Sigma-Aldrich (St. Louis, MO). Britton-Robinson universal buffer was prepared according to Mongay and Cerda (1974) with modifications. The stock buffer contained 7 g/L boric acid, 15.32 g/L citric acid · H₂O, 7.38 g/L Na₂HPO₄ and 20 g/L NaOH. Universal buffer from pH 3 to 10 was produced by adjusting the stock buffer with 5 M HCl.

Expression and purification of C. bulbillosum cellulases

Plugs of V8 agar medium with fresh *C. bulbillosum* mycelia were added to 1-L flasks, each of which contained 250 mL cellulase induction medium and 50 mM Tris buffer, pH 9. Flasks were incubated at 200 rpm for 3 days at room temperature. The *C. bulbillosum* cultures were filtered through two layers of Miracloth (Calbiochem, San Diego, CA), concentrated approximately 15 times on a rotary evaporator under vacuum at 35°C, and exchanged into 20 mM pH 8 Tris buffer by membrane filtration with 10 kDa molecular weight cut off (Vivaflow® 50 Polyethersulfone membrane; Sartorius AG, Göttingen, Germany).

Individual cellulases were purified by chromatography using an Agilent HPLC 1200 system equipped with UV detector and fraction collector. Chromatography columns (Tosoh Bioscience, LLC, King of Prussia, PA) were all 8.0 mm x 7.5 cm in dimension and packed with TSKgel® DEAE-5PW (anion exchange), TSKgel® SP-5PW (cation exchange), or TSKgel® Phenyl-5PW (hydrophobic interaction). For anion exchange chromatography, solution A was 20 mM Tris, pH 8, and solution B was 20 mM Tris, pH 8, plus 0.6 M NaCl. For cation exchange, solution A was 20 mM sodium acetate, pH 4, while solution B was the same buffer with 0.6 M NaCl. For hydrophobic interaction chromatography, solution A was 0.1 M KH₂PO₄, pH 7, plus 1.7 M (NH₄)₂SO₄ while solution B was deionized H₂O. The elution conditions for ion exchange chromatography were 0% solution B for 5 min, followed by a linear gradient from 0% to 100% B over 30 min. For hydrophobic interaction chromatography, solution B gradient was increased from 0% to 100% in 30 min and held at 100% for 10 min. In all cases, enzymes were eluted at a flow rate of 1 mL/min, and 1 mL fractions were collected for 35-45 min.

Proteomic analysis (as provided by MSU RTSF)

Proteins in the purified cellulase samples were analyzed by mass spectrometry (MS) at the MSU Research Technology Support Facility. Chromatographic fractions with purified cellulases were first analyzed by SDS-PAGE. After staining with Coomassie Blue R-250, protein bands were excised and were digested in-gel according to Shevchenko et al. (1997) with modifications. First, gel bands were dehydrated using 100% acetonitrile, incubated at 56°C for 45 min with 10 mM dithiothreitol in 100 mM ammonium bicarbonate, dehydrated again, and incubated in the dark for 20 min with 50 mM iodoacetamide in 100 mM ammonium bicarbonate. Gel bands were then washed with ammonium bicarbonate and dehydrated again. Sequencing grade modified trypsin was prepared to 10 µg/mL in 50 mM ammonium bicarbonate, and ~50 µL of the trypsin solution was added to each gel band so that the gel was completely submerged. Bands were then incubated at 37°C overnight. Peptides were extracted from the gel by water bath sonication in a solution of 60% acetonitrile and 1% trichloroacetic acid, and vacuum dried to ~2 µL. Peptides were then re-suspended in 20 µL solution of 2% acetonitrile and 0.1% trifluoroacetic acid.

The extracted peptides were separated by a Waters nanoAcquity UPLC. First, 10 µL peptide solution were automatically injected by a Waters nanoAcquity Sample Manager and loaded at 4 µL/min for 5 minutes onto a Waters Symmetry C18 peptide trap (5 µm, 180 µm x 20 mm) in 5% acetonitrile and 0.1% formic acid. The bound peptides were then eluted onto a Waters BH130 C18 column (1.7 µm, 100 µm x 150 mm). For reverse-phase chromatography, buffer A was 0.1% formic acid in water, and buffer B was 0.1% formic acid in acetonitrile. The peptides were eluted at a constant flow rate of 1 µL/min over 16 min. The elution conditions

were solution B gradient from 5% to 30% in 8 min, ramped up to 90% at 9 min and held for 1 min, then dropped back to 5% B at 10.1 min.

Eluted peptides were sprayed into a Thermo Fisher LTQ Linear Ion trap mass spectrometer outfitted with a MICHROM Bioresources ADVANCE nano-spray source. The top five ions in each survey scan were then subjected to data-dependent zoom scans followed by low energy collision induced dissociation (CID). The resulting MS/MS spectra were converted to peak lists in Mascot Distiller, v2.4.3.3, using the default LTQ instrument parameters. Peak lists were searched against all sequences available in the *C. bulbiliosum* protein sequence database (downloaded from the Joint Genome Institute, <http://www.jgi.doe.gov/>) appended with common laboratory contaminants (downloaded from www.thegpm.org, cRAP project) using the Mascot searching algorithm, v2.4. Mascot parameters were set as follows: allowing up to 2 missed tryptic sites; fixed modifications of carbamidomethyl cysteine; variable modification of oxidation of methionine; peptide tolerance of ± 200 ppm; MS/MS tolerance of 0.6 Da; peptide charge state limited to +2/+3. The Mascot output was then analyzed using Scaffold Q+S, v4.3.0, to probabilistically validate protein identifications. Assignments validated in Scaffold with <1% FDR were considered true.

Protein concentrations of purified cellulases were measured according to Bradford (1976) using bovine serum IgG as standard.

Enzyme assays

Chromatographic fractions were tested for CBH1, EG, and BG activities by addition of 2 μ L of each fraction into reactions containing 50 mM Tris, pH 8, and 2.5 mM of the appropriate substrate (4-nitrophenyl-lactoside, 4-nitrophenyl-cellobioside, or 4-nitrophenyl-glucoside,

respectively; Banerjee et al. 2010b). Each 80 μL reaction was incubated at 50°C for 10 min and terminated by addition of 120 μL 1 M Na_2CO_3 . Optical density was measured at 405 nm using a microplate reader. A standard curve was constructed using 4-nitrophenol.

To obtain pH profiles of the GH family 7 EG from *C. bulbillosum* and its homolog from *T. reesei*, 80 μL reactions containing 4 $\mu\text{g}/\text{mL}$ EG, 2.5 mM 4-nitrophenyl cellobioside, and 50 μL Britton-Robinson universal buffer from pH 3 to 10 were incubated for 20 min at 50°C. All reactions were within the linear range. Each reaction was conducted in triplicates. The reaction conditions to determine the pH profiles of purified GH family 3 BG from *C. bulbillosum* and *T. reesei* were the same except that 0.5 $\mu\text{g}/\text{mL}$ BG was added in each reaction and the reaction time was 5 min. To determine kinetic parameters, 40 ng BG or 320 ng EG was added to reactions containing 0-15 mM substrate and universal buffer at the optimal pH of each enzyme. Triplicate reactions were performed for 5 to 20 min at 50°C and were always confirmed to be within the linear range. Standard Michaelis-Menten enzyme kinetics were assumed.

Results

Cladorrhinum bulbillosum (Figure 3.1) was one of the three fungi that were isolated that could grow on cellulose at pH 9 (D. Jabbour, unpublished data). Relative activities of the *C. bulbillosum* crude culture and the commercial enzyme mixture Cellic[®] CTec2/HTec2 derived from *T. reesei* (Tr) (Novozymes, Franklinton, NC) were compared on filter paper over a range of pH from 5 to 10 (Figure 3.2). Both enzyme mixtures were most active at pH 5, but *C. bulbillosum* maintained 88% of its activity at pH 7 and 67% at pH 10. In contrast, the

CTec2/HTec2 mixture maintained only 35% of its activity at pH 7 and 6% at pH 10. *C. bulbilosum* cellulolytic enzymes showed activity over a broader pH range.



Figure 3.1 *C. bulbilosum* growing on V8 juice agar, 10 days old.

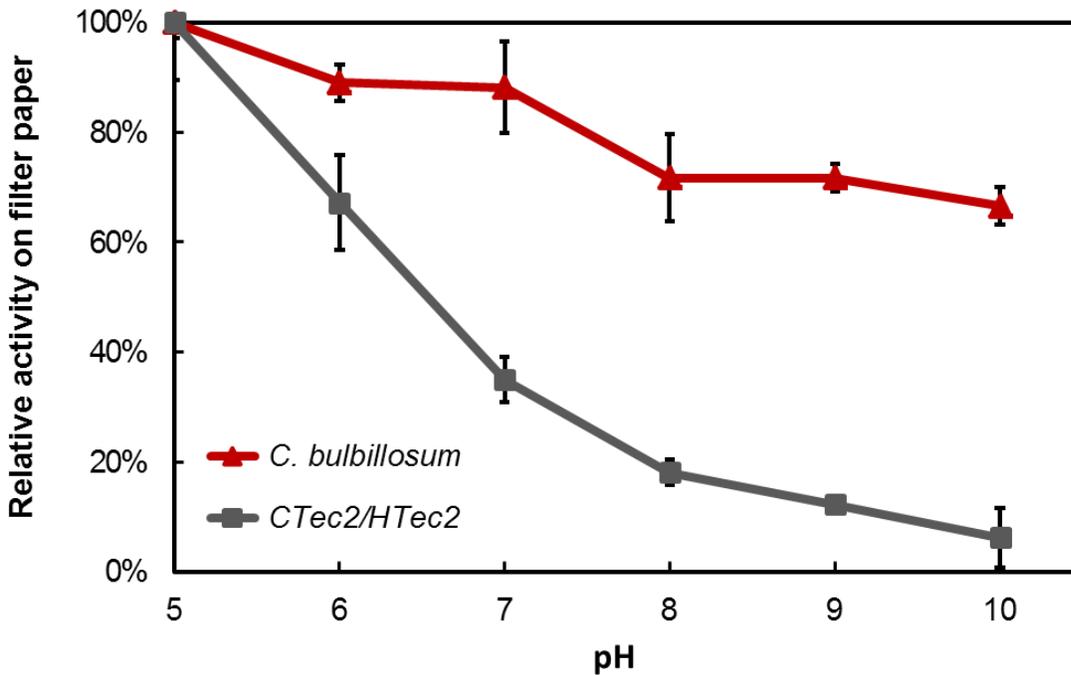


Figure 3.2 Comparison of relative activities of *C. bulbillosum* crude culture filtrate and commercial enzyme mixture CTec2/HTec2. Whatman No. 1 filter paper was the substrate (experiment performed by Dina Jabbour).

CAZymes in the *C. bulbillosum*, *T. reesei* and *P. anserina* genomes

The genome and transcriptomes of *C. bulbillosum* were sequenced by the DOE Joint Genome Institute (JGI), and the CAZymes annotated (Lombard et al. 2013). Annotated genome and transcriptome data are available on the JGI website

<https://genome.jgi.doe.gov/Clabul1/Clabul1.home.html>.

CAZymes in the *C. bulbillosum* genome were compared against those in *T. reesei*, which is the major source of current available commercial enzymes. CAZymes in *C. bulbillosum* were also compared to those in *Podospora anserina*, which is in the same family as *C. bulbillosum* (family Lasiosphaeriaceae) and whose genome has been sequenced and annotated. Additionally, its CAZymes have been studied (Bennati-Granier 2015; Couturier et al. 2011; Espagne et al.

2008; Poidevin et al. 2013, 2014). The number of different CAZyme families in all three genomes was estimated by dbCAN (Yin et al. 2012). The CAZyme profiles for *C. bulbillosum* and *P. anserina* are similar but different from the profile for *T. reesei* (Figure 3.3). All three fungi have similar numbers of glycosyltransferases and polysaccharide lyases. However, *C. bulbillosum* and *P. anserina* have slightly more glycoside hydrolases (GHs), which contain wall-polysaccharide-hydrolyzing enzymes (Nagendran et al. 2009). In addition, *C. bulbillosum* and *P. anserina* have approximately 3.4 times more carbohydrate esterases, and 2.1 times more auxiliary activity enzymes than *T. reesei*. For noncatalytic carbohydrate binding module (CBM), *P. anserina* has the highest number of CBMs while *T. reesei* has the lowest. Overall, *C. bulbillosum* has 1.3 times more CAZymes than *T. reesei*.

The three fungi were also compared for their number of enzymes and CBMs in eight families known to contribute to biomass degradation (Banerjee et al. 2010c, Nagendran et al. 2009; Figure 3.4). Similar to the CAZyme profile, *C. bulbillosum* and *P. anserina* have a similar number of enzymes and CBMs in the eight families. All three fungi have similar number of GH3 and GH5 enzymes. However, *C. bulbillosum* and *P. anserina* contain many more enzymes in the other six families compared to *T. reesei*. Overall, *C. bulbillosum* and *P. anserina* contain about 2.6 times more major lignocellulose-degrading enzymes, and 1.9 times more CBMs than *T. reesei*. Particularly striking is the greater number of AA9 genes in the *C. bulbillosum* and *P. anserina* genomes compared to *T. reesei* (more than 30 vs. 4). This family contains the lytic polysaccharide monooxygenases (LPMOs) which cleave cellulose chains oxidatively (Horn et al. 2012; Vaaje-Kolstad et al. 2010). In general, *C. bulbillosum* contains more lignocellulose-degrading enzymes in its genome compared to *T. reesei*.

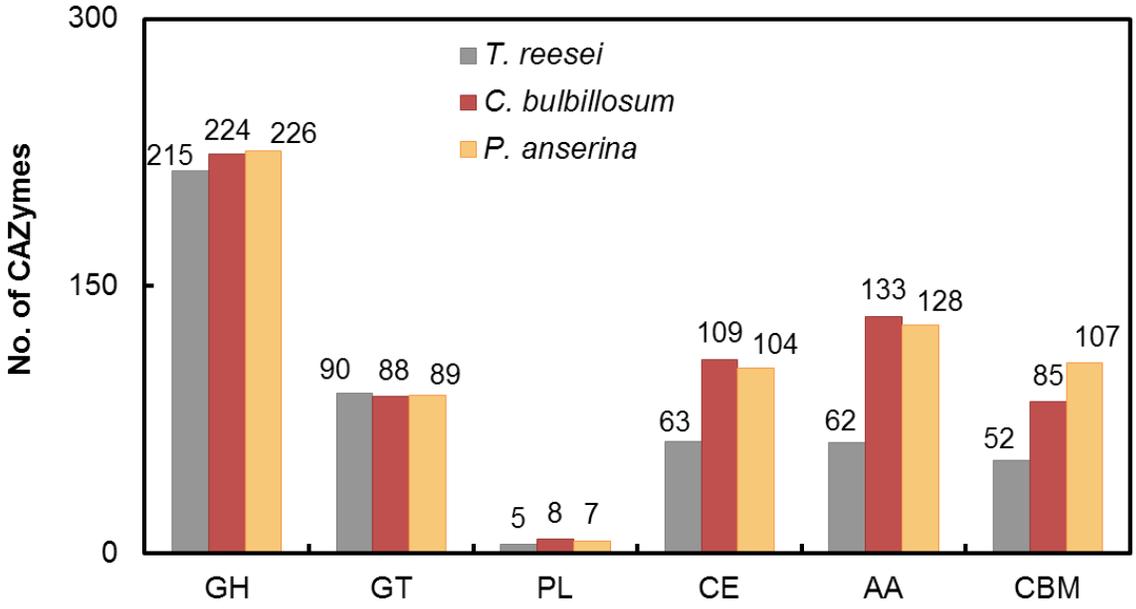


Figure 3.3 Comparison of number of Carbohydrate-Active enZymes (CAZymes) encoded in the *C. bulbilosum*, *T. reesei* and *P. anserina* genomes. GH stands for Glycoside Hydrolase, GT for Glycosyl Transferase, PL for Polysaccharide Lyase, CE for Carbohydrate Esterase, AA for Auxiliary Activity, and CBM for Carbohydrate-Binding Module.

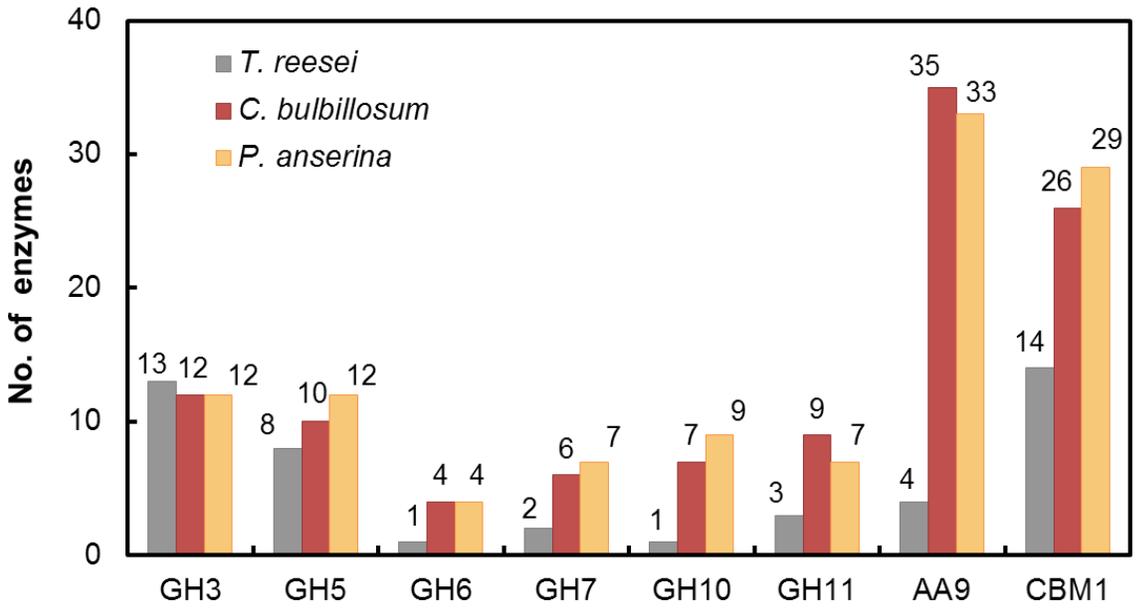


Figure 3.4 Comparison of number of enzymes in eight CAZyme families that contain many of the most important enzymes and CBMs known to be necessary for efficient biomass conversion.

Purification of C. bulbillosum cellulases

C. bulbillosum culture filtrate was first fractionated by anion exchange chromatography (AEX; Figure 3.5). Each fraction was tested for EG, BG and CBH1 activity. High BG activity was detected in the early fractions (fractions 5-8), which indicated that BG(s) did not bind to the AEX column. In contrast, high CBH1 activity was detected in fractions 22-23. However, fractions showing high BG or CBH1 activities also showed moderate EG activity. BG(s) and CBH1(s) were separated from each other after the first chromatography step, but they were not separated from EG(s).

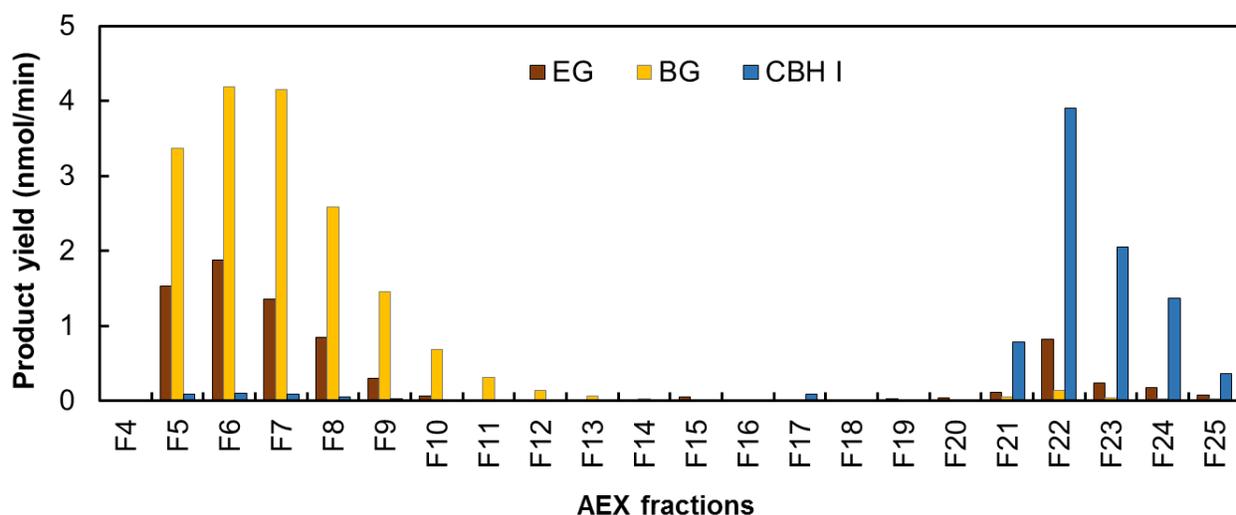


Figure 3.5 First chromatographic fraction (anion exchange) of *C. bulbillosum* culture filtrate. Two microliters of each fraction were assayed for EG, BG, and CBH1 activity using as substrates 4-nitrophenyl cellobioside, 4-nitrophenyl glucoside, and 4-nitrophenyl lactoside, respectively.

To separate BG(s) and EG(s), AEX fractions 5-7 showing high BG activity were pooled and further fractionated by cation exchange chromatography (CEX; Figure 3.6). The highest BG activity was detected in CEX fraction 16 while EG activity was in fractions 23-24. After cation

exchange chromatography, BG(s) was separated from EG(s). To identify purified BG(s), CEX fractions 16 from multiple rounds of chromatography were combined for proteomic analysis.

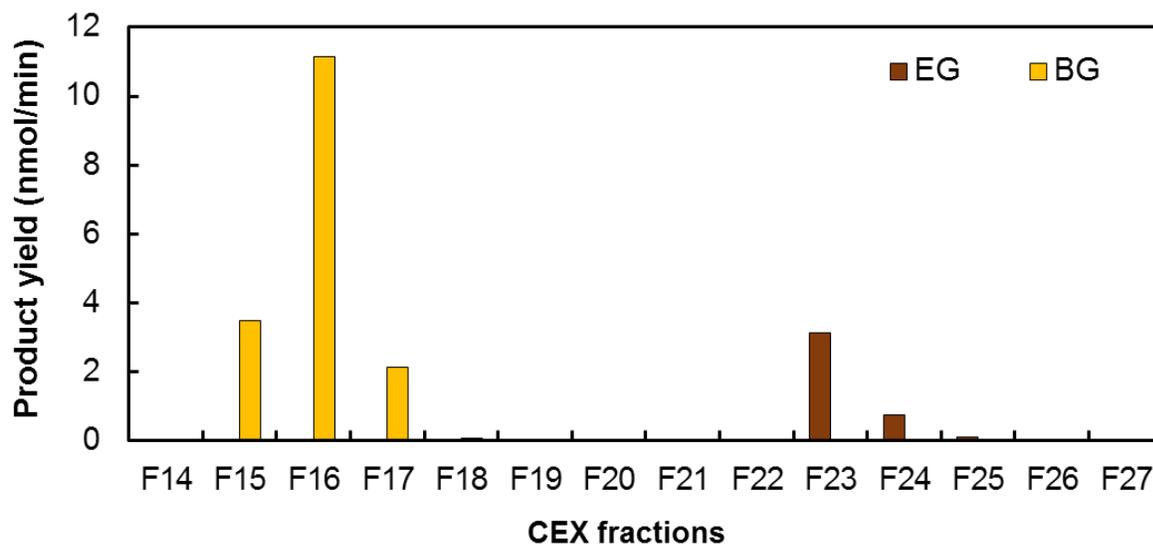


Figure 3.6 Cation exchange chromatography (CEX) of AEX fractions 5-7 (see Figure 3.5) to separate *C. bulbiliosum* BG(s) and EG(s). Each fraction was assayed for EG and BG activity using 4-nitrophenyl cellobiose and 4-nitrophenyl glucoside, respectively, as substrates.

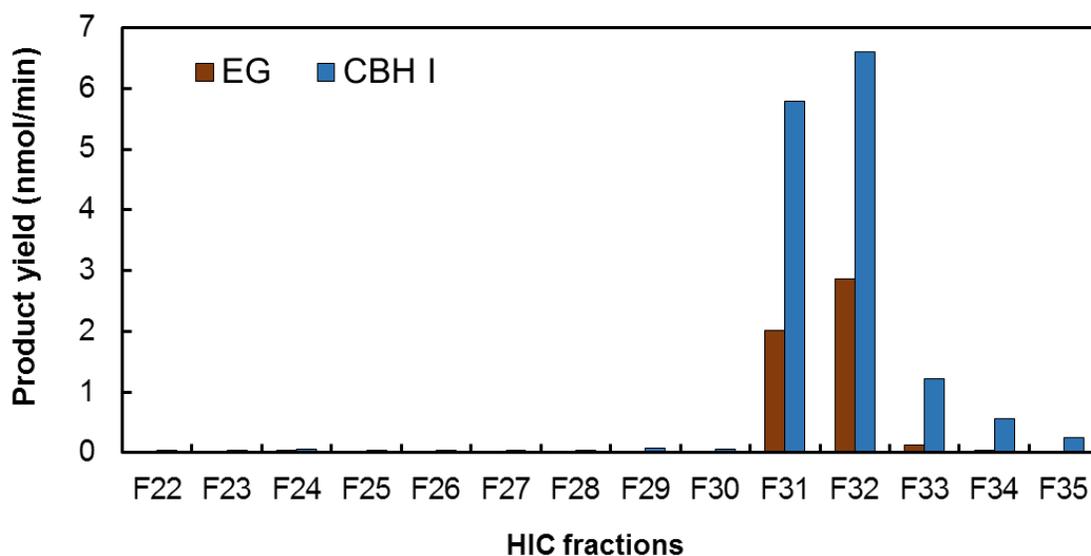


Figure 3.7 Hydrophobic interaction chromatography of AEX fraction 23 (see Figure 3.5) to separate *C. bulbiliosum* CBH1(s) and EG(s). Substrates for EG and CBH1 were 4-nitrophenyl cellobiose and pNP-lactoside.

To separate CBH1(s) and EG(s), AEX fraction 22 (Figure 3.5) from different rounds of chromatography were pooled and subjected to hydrophobic interaction chromatography (HIC; Figure 3.7). EG and CBH1 activities were detected in the same fractions (HIC fractions 31-32). As a result, EG(s) and CBH1(s) could not be separated from each other by HIC. HIC fractions 31-32 showing highest EG and CBH1 activity were combined for proteomic analysis.

Proteomic analysis of the purified C. bulbillosum cellulases

Proteins in purified BG and partially purified CBH1 samples were analyzed by LC-MS and searched against the JGI *C. bulbillosum* DJ3 v1.0 genome database. GH3 BG (CbCel3A) was the dominant protein in the purified BG sample (Table 3.1), accounting for 98% of the total proteins. Trace amount of an AA7 oligosaccharide oxidase and a GH7 enzyme were also identified in the sample. A GH7 enzyme, CbCel7E, was the major protein in the purified CBH1 sample, accounting for 80% of the total proteins (Table 2). Besides CbCel7E, 2% of a CBH1 (CbCel7A) and 11% of other CAZymes were also present in the sample. BLAST of CbCel7E indicated that CbCel7E showed 71% similarity over 88% of the protein with a well characterized *Humicola grisea* EG (Genbank accession no. P56680). Therefore, CbCel7E was considered an EG instead of a CBH1. Testing of a GH7 endoglucanase from *T. reesei*, TrCel7B, against 4-nitrophenyl lactoside indicated that TrCel7B was also active on this substrate (data not shown). This explains why a GH7 EG but not a CBH1 was dominant in the putative CBH1 sample and why the highest EG and CBH1 activity were both detected in the same HIC fractions (Figure 3.7).

Table 3.1 All proteins identified in purified BG sample at >95% confidence (Scaffold). Protein IDs were from JGI *C. bulbiliosum* DJ3 v1.0 database. Identified proteins were sorted by normalized total spectra counts. MW stands for molecular weight.

Protein ID	Protein name	Function	MW (kDa)	CAZyme family	Spectral counts	Coverage (%)
850450	CbCel3A	β -glucosidase	78	GH3	665	78%
428267	CbAA7A	oligosaccharide oxidase	71	AA7	7	13%
844573	CbCel7E	endo- β 1,4-glucanase	48	GH7	4	9%

Table 3.2 All proteins identified in partially purified CBH1 sample at >95% confidence (Scaffold). Protein IDs were from JGI *Cladorrhinum bulbiliosum* DJ3 v1.0 database. Identified proteins were sorted by normalized total spectra counts. GMC oxidoreductase stands for glucose-methanol-choline oxidoreductase and MW for molecular weight.

Protein ID	Protein name	Function	MW (kDa)	CAZyme family	Spectral counts	Coverage (%)
844573	CbCel7E	endo- β 1,4-glucanase	48	GH7	397	67%
310615	CbGH125A	exo- α -1,6-mannosidase	60	GH125	14	21%
878606	CbAA5A	glyoxal oxidase-like protein	117	AA5	11	13%
848016	Unknown	β -lactamase-like protein	67	N/A	11	9%
894551	CbCel7A	cellobiohydrolase 1	56	GH7	11	17%
862051	CbCE2A	acetyl xylan esterase	48	CE2	9	11%
845425	CbCel5D	endo- β 1,4-mannanase	46	GH5	8	14%
949261	Unknown	aminopeptidase	55	N/A	8	13%
884977	Unknown	aminopeptidase	56	N/A	7	11%
864683	CbAA3E	cellobiose dehydrogenase	86	AA3	7	9%
894356	Unknown	amine oxidase	53	N/A	7	13%
847174	CbGH55A	exo- β -1,3-glucanase	92	GH55	6	6%

Table 3.2 (Cont'd)

Protein ID	Protein name	Function	MW (kDa)	CAZyme family	Spectral counts	Coverage (%)
853112	Unknown	Unknown	41	N/A	3	5%
852081	Unknown	Unknown	45	N/A	3	6%
850450	CbCel3A	β -glucosidase	78	GH3	2	3%
952736	Unknown	actin	42	N/A	2	7%
953816	CbAA3B	GMC oxidoreductase	67	AA3	2	5%

Comparing pH profiles of C. bulbillosum and T. reesei cellulases

Activities of the purified *C. bulbillosum* BG (CbCel3A) and EG (CbCel7E) were compared against the activities of their homologs from *T. reesei*, TrCel3A and TrCel7B (Figure 3.8), respectively, at different pHs. The pH optimum of CbCel7E was pH 6 while that of TrCel7B was pH4; CbCel7E was active between pH 4 and 10 while TrCel7B was active between 3 and 7 (Figure 3.9). Although TrCel7B activity was 1.4 times the activity of CbCel7E at their optimal pHs, CbCel7E showed higher specific activity than TrCel7B above pH 6. At pH 6, CbCel7E showed approximately 2.6 times as much activity as TrCel7B (Figure 3.9a). TrCel7B was not active above pH 7, but CbCel7E maintained 72% of its activity at pH 7 and 33% at pH 8 (Figure 3.9b). In general, CbCel7E was active over a broader pH range and showed higher activity above pH 6.

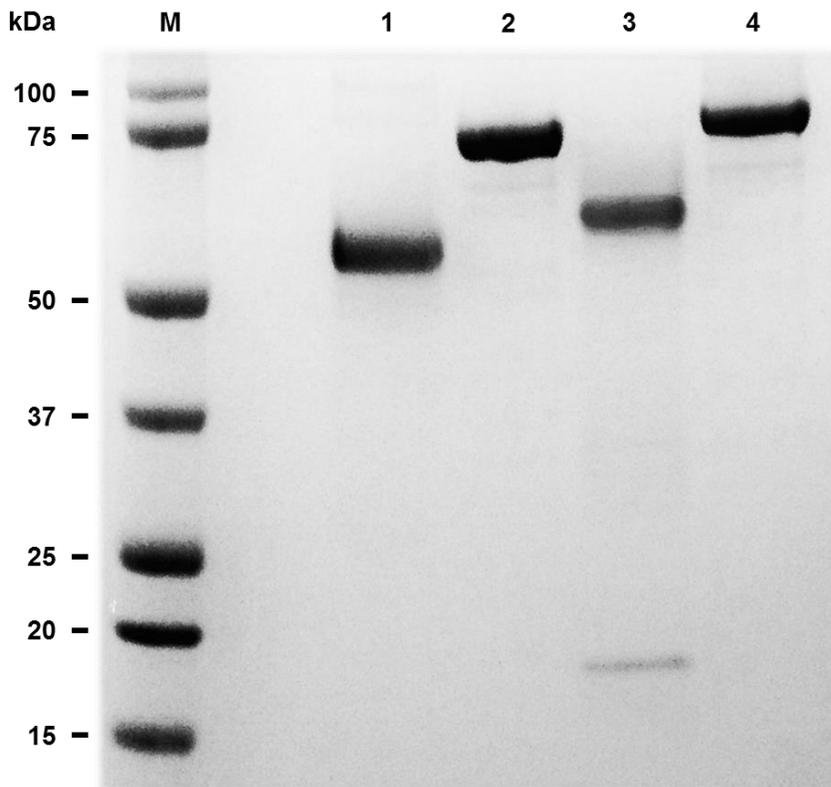


Figure 3.8 SDS-PAGE of purified *C. bulbillosum* and *T. reesei* cellulases. Gel was stained with Coomassie blue. Proteins from lane 1 to 4 are CbCel7E, CbCel3A, TrCel7B, and TrCel3A, respectively. M is the protein size markers (in kDa).

Both CbCel3A and TrCel3A were active between pH 3 and 9. The optimal pH for CbCel3A was pH 5 while that for TrCel3A was pH 6 (Figure 3.10). TrCel3A showed 1.5 times more activity than CbCel3A at their optimal pHs, but CbCel3A was more active above pH 6 (Figure 3.10a). It maintained 74% activity at pH 7 and 23% activity at pH 8, while TrCel3A maintained 16% activity at pH 7 and only 3% at pH 8 (Figure 3.10b). Comparing to the two cellulases from *T. reesei*, the two *C. bulbillosum* cellulases showed higher activities between pH 6-8, although they were less active than their *T. reesei* homologs at their respective optimal pHs.

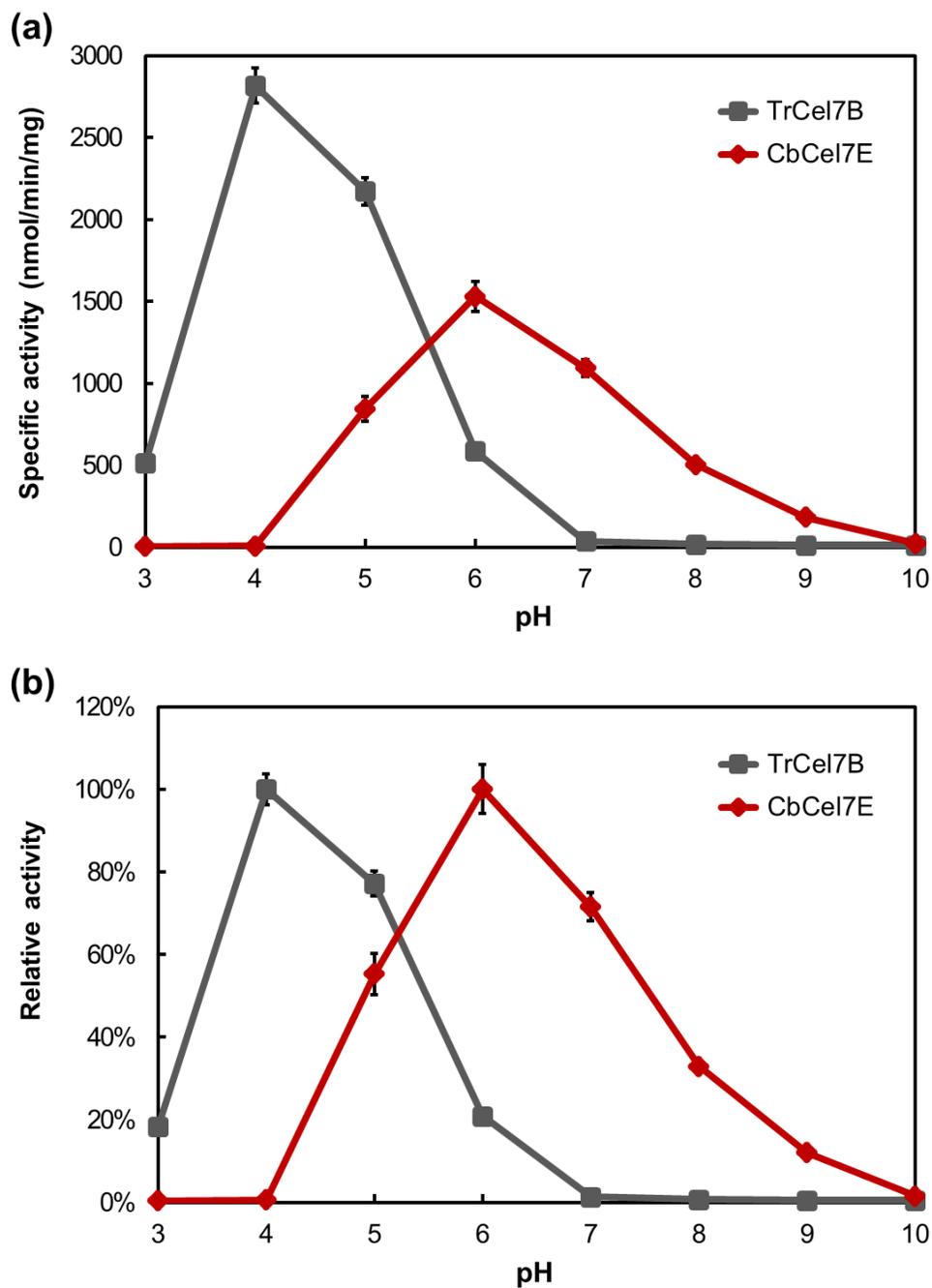


Figure 3.9 pH profiles of CbCel7E and TrCel7B. All reactions were performed in universal buffer ranging from pH 3 to 10. (a) Specific activities of CbCel7E and TrCel7B. (b) Relative activities of CbCel7E and TrCel7B. The activities of the individual enzymes at their pH optima were set to 100%.

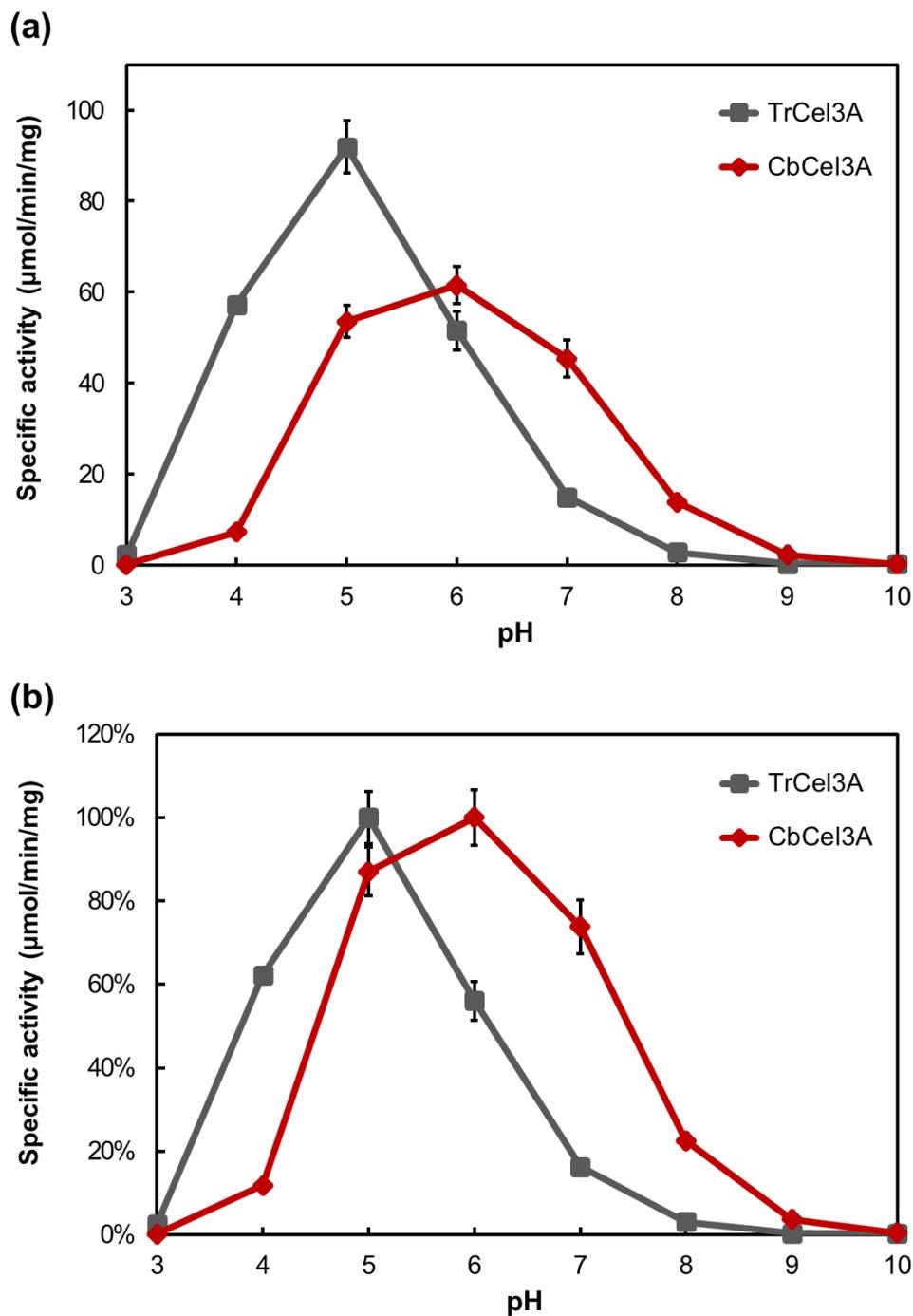


Figure 3.10 pH profiles of CbCel3A and TrCel3A. All reactions were performed in universal buffer ranging from pH 3 to 10 within linear range. (a) Specific activities of CbCel3A and TrCel3A. (b) Relative activities of CbCel3A and TrCel3A. The activities of the individual enzymes at their pH optima were set to 100%.

Comparison of kinetic parameters of C. bulbillosum and T. reesei cellulases

CbCel7E and TrCel7B showed very different kinetic parameters when the substrate was 4-nitrophenyl cellobiose. The K_m of TrCel7B was approximately 3.4 times higher than CbCel7E (Table 3.3). The k_{cat} of TrCel7B was about four times that of CbCel7E (Table 3.3, Figure 3.11). In contrast, CbCel3A and TrCel3A showed similar K_m and relatively similar K_{cat} (Table 3.3, Figure 3.12). The K_m of both BGs was about 0.4 mM, while the k_{cat} of TrCel3A was approximately 1.4 times higher than CbCel3A. Compared to the two EGs, the two BGs showed lower K_m and higher k_{cat} .

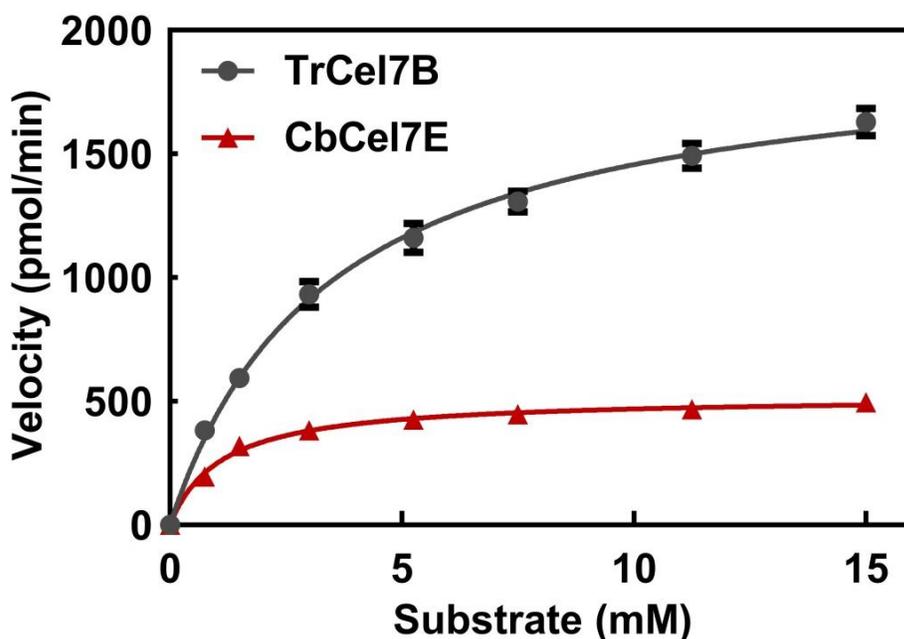


Figure 3.11 Initial reaction velocities of CbCel7E and TrCel7B at different concentrations of 4-nitrophenyl cellobioside. All reactions were performed in triplicates at the optimal pH of each enzyme. Enzyme kinetic curves were generated by GraphPad Prism v7.03.

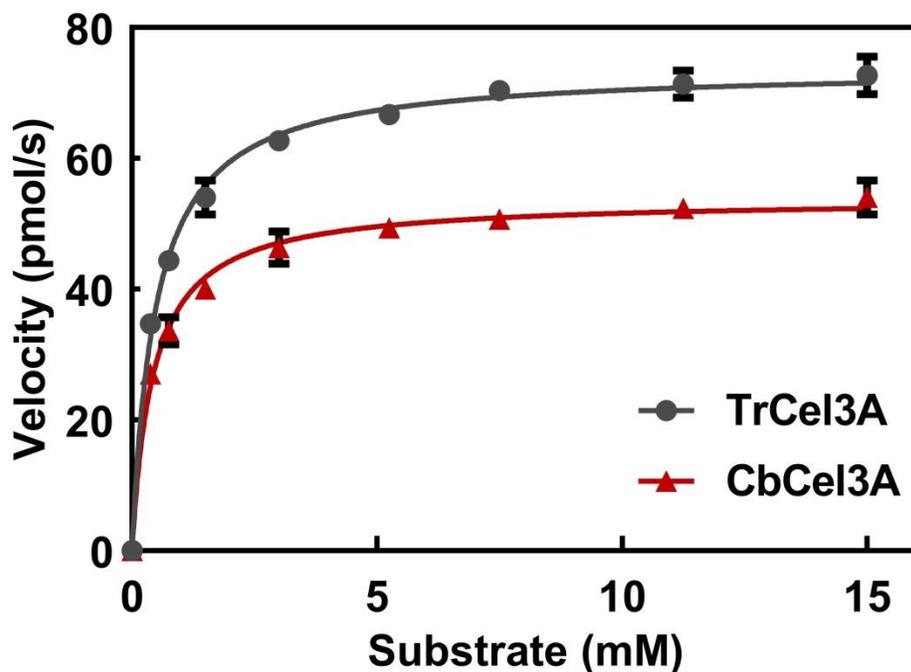


Figure 3.12 Initial reaction velocities of CbCel3A and TrCel3A at different substrate concentrations. All reactions were performed in triplicates at the optimal pH of each enzyme. Enzyme kinetic curves were generated by GraphPad Prism v7.03.

Table 3.3 Kinetic parameters of four *C. bulbilosum* and *T. reesei* cellulases. Kinetic parameters were calculated by GraphPad Prism v7.03. pNP-C stands for 4-nitrophenyl cellobioside and pNP-G for 4-nitrophenyl glucoside.

Enzyme	Substrate	K_m (mM)	K_m 95% confidence interval (mM)	K_{cat} (min^{-1})	K_{cat} 95% confidence interval (min^{-1})
TrCel7B	pNP-C	3.43	3.02-3.89	279	265-292
CbCel7E	pNP-C	1.10	0.92-1.32	74	72-77
TrCel3A	pNP-G	0.48	0.43-0.54	8687	8519-8858
CbCel3A	pNP-G	0.43	0.38-0.49	6096	5956-6239

Discussion

Current *T. reesei*-based lignocellulose-degrading enzyme mixtures show their maximal activities around pH 5. However, a large portion of cellulases adsorb to and are inhibited by lignin present in the lignocellulosic biomass at pH 5 (Berlin et al. 2006; Palonen et al. 2004; Ooshima et al. 1990; Rahikainen et al. 2013a). Recent studies reveal that cellulase adsorption to lignin is largely reduced above pH 6 where *T. reesei*-based enzyme mixtures lose most of their activity (Lou et al. 2013; Rahikainen et al. 2013b). To overcome the nonproductive binding of cellulases to lignin and hence increase biomass degradation efficiency, one strategy is to replace or supplement current *T. reesei*-based enzyme mixtures with alkaliphilic enzymes that can efficiently degrade biomass above pH 6.

While working in the laboratory, Dina Jabbour isolated a fungus, *Cladorrhinum bulbillosum*, that could live on and degrade cellulose at pH 9. Compared to the enzyme mixture Cellic® CTec2/HTec2, a second-generation commercial cellulase mixture from Novozymes, *C. bulbillosum* crude culture filtrate maintained more activity above pH 6. To obtain and characterize the major cellulases in the *C. bulbillosum* secretome, *C. bulbillosum* culture filtrate was fractionated by anion exchange, cation exchange, and hydrophobic interaction chromatography. As a result, a GH7 endo- β 1,4-glucanase (CbCel7E) and a GH3 β -glucosidase (CbCel3A) were purified by chromatography.

CbCel7E and CbCel3A were compared against their *T. reesei* homologs, TrCel7B and TrCel3A, respectively, for their pH profiles and catalytic abilities. The two *T. reesei* cellulases were active around pH 4-6 while the two *C. bulbillosum* cellulases were active around pH 5-8. Both *T. reesei* enzymes showed higher K_{cat} than their *C. bulbillosum* counterparts at their optimal

pHs. However, two *C. bulbillosum* cellulases were more active between pH 6-8. Therefore, these two *C. bulbillosum* cellulases may degrade lignocellulosic biomass more efficiently than their *T. reesei* homologs at pH 6-8.

Two purified *C. bulbillosum* cellulases were not active above pH 9. However, *C. bulbillosum* crude culture filtrate maintained more than 60% activity at pH 10. It is possible that other unpurified cellulases such as GH7 CBHs, GH5 EGs and AA9 LPMOs are active between pH 9-10. To obtain other *C. bulbillosum* cellulolytic enzymes and construct a synthetic cocktail for biomass degradation, one strategy would be to heterologously express *C. bulbillosum* cellulase genes in *T. reesei*. Another strategy would be to induce the expression of different *C. bulbillosum* enzymes using different carbon sources.

In summary, *C. bulbillosum* crude culture filtrates and purified cellulases both showed higher activities than *T. reesei* enzymes above pH 6. Alkaliphilic cellulolytic enzymes from *C. bulbillosum* are a promising way to increase efficiency of lignocellulosic biomass deconstruction at high pH.

CHAPTER 4: REAL-TIME IMAGING REVEALS THAT LYTIC POLYSACCHARIDE MONOOXYGENASE PROMOTES CELLULASE ACTIVITY BY INCREASING CELLULOSE ACCESSIBILITY

The material in this chapter has been published: Song B[§], Li B[§], Wang X, Shen W, Park S, Collings C, Feng An, Smith SJ, Walton JD, Ding SY (2018) Real-time imaging reveals that lytic polysaccharide monooxygenase promotes cellulase activity by increasing cellulose accessibility. *Biotechnology for biofuels*, 11(1): 41

I would like to acknowledge Dr. Bo Song for his contribution to the results shown in Figures 4.4 - 4.7 and Figure 4.9, and to the writing of the manuscript entitled “Real-Time Imaging Reveals Lytic Polysaccharide Monooxygenase Promotes Cellulase Activity by Increasing Cellulose Accessibility”.

[§]Authors contributed equally to the work.

Abstract

The high cost of enzymes is one of the key technical barriers that must be overcome to realize the economical production of biofuels and biomaterials from biomass. Supplementation with lytic polysaccharide monooxygenase (LPMO) can increase the efficiency of cellulase mixtures for biomass conversion. In this study, we combined enzymatic assays and real time imaging using atomic force microscopy (AFM) to study the molecular interactions of an LPMO (TrAA9A, formerly known as TrCel61A) and a cellobiohydrolase 1 (TlCel7A) with bacterial microcrystalline cellulose (BMCC) as a substrate. Cellulose conversion by TlCel7A alone was enhanced from 46% to 54% by addition of TrAA9A. Conversion by a mixture of TlCel7A, endoglucanase, and β -glucosidase was increased from 79% to 87% by using TrAA9A-pretreated BMCC. AFM imaging demonstrated that individual TrAA9A molecules exhibited intermittent random movement along, across, and penetrating into the ribbon-like microfibril structure of BMCC, concomitant with the release of a small amount of oxidized sugars and the splitting of large cellulose ribbons into fibrils with smaller diameters. The dividing effect of the cellulose microfibril occurred more rapidly when TrAA9A and TlCel7A were added together compared to TrAA9A alone; TlCel7A alone caused no separation. Therefore, TrAA9A increases the accessible surface area of BMCC by separating large cellulose ribbons, and thereby enhances cellulose hydrolysis yield. By providing the first direct observation of LPMO action on a cellulosic substrate, this study sheds new light on the mechanisms by which LPMO enhances biomass conversion.

Introduction

Non-food plant biomass is a sustainable source of fermentable sugars for the production of biofuels and chemicals (Bozell and Peterson 2010; Himmel et al. 2007; Kumar and Bhattacharya 2008). Cellulose, the main component of lignocellulosic biomass, forms rigid microfibrils composed of well-organized linear β -1,4-glucan chains. Due to its homogeneity and abundance, cellulose is of major interest in producing monomeric sugars for biorefineries (Himmel et al. 2007; Horn et al. 2012; Ragauskas et al. 2006). However, the cellulose chains in a microfibril are strongly held together by hydrogen bond networks and van der Waals forces, so that the majority of these chains are not readily accessible to cellulases. In addition, in plant biomass, microfibrils are organized into closely associated bundles and are embedded in a non-cellulosic matrix of lignin and hemicelluloses, which further reduce the efficiency of biomass conversion (Himmel et al. 2007; Mansfield et al. 1999).

The most widespread method of cellulose conversion requires three types of cellulases to act synergistically to convert cellulose to glucose: endo- β -1,4-glucanases (EGs) cleave the internal bonds in the cellulose chain, cellobiohydrolases (CBHs) processively hydrolyze cellulose from the chain ends to produce cellobiose, and β -glucosidases (BGs) hydrolyze cellobiose to glucose (Lynd et al. 2002). Lytic polysaccharide monooxygenases (LPMOs) are a recently discovered class of enzymes that stimulate biomass hydrolysis and hence improve the efficiency of biomass conversion (Banerjee et al. 2010d; Harris et al. 2010; Hu et al. 2014; Phillips et al. 2011). Unlike cellulases that cleave glycosidic bonds by hydrolysis (Davies and Henrissat 1995), LPMOs are copper-dependent enzymes that lyse polysaccharide chains by oxidation at either the C1 or the C4 carbon of the glucose unit in the presence of an external

electron donor (Kracher et al. 2016; Quinlan et al. 2011; Vaaje-Kolstad et al. 2010; Westereng et al. 2011, 2015). However, the mechanism by which the addition of LPMOs to cellulase mixtures enhances the overall yield of glucose has not been entirely elucidated.

Atomic force microscopy (AFM), particularly high-speed AFM, can image with sub-nanometer resolution under aqueous conditions at video rates. It has been used to visualize the interactions between enzymes and substrates in real time (Dufrêne et al. 2017). AFM has also been used to image the structural changes of cellulose after treatment by LPMO, thereby showing disintegrating and fibrillation of cellulose microfibrils (Villares et al. 2017). However, how the interaction of LPMO with the cellulose surface causes these changes has not been elucidated. In this study, we used high-speed AFM to monitor *in situ* interactions of an LPMO (TrAA9A, formerly known as TrCel61A) from *Trichoderma reesei* and a CBH1 (TlCel7A) from *T. longibrachiatum*, alone and together, with bacterial microcrystalline cellulose (BMCC). Our goal was to visualize in real time the molecular motion of LPMO and its effects on the structure of the cellulose surface, in order to deepen our understanding of LPMO function in cellulose hydrolysis.

Materials and methods

General chemicals

All chemicals and reagents, unless specifically noted, were purchased from Sigma–Aldrich (St. Louis, MO).

Bacterial strain and medium

Gluconacetobacter xylinus (also known as *Komagataeibacter xylinus*) strain ATCC 53524 from the American Type Culture Collection (Manassas, VA) was maintained on Hestrin and Schramm (HS) medium agar. HS medium contains 20 g/L glucose, 5 g/L peptone, 5 g/L yeast extract, 2.7 g/L Na₂HPO₄, 1.15 g/L citric acid monohydrate and 15 g/L agar. The pH was adjusted to 6.0 with NaOH or HCl (Hestrin and Schramm 1954).

Enzymes

TlCel7A from *Trichoderma longibrachiatum* (catalog no. E-CBH1) was purchased from Megazyme, Ltd., Bray, Ireland. TrAA9A (GenBank CAA71999), endo- β 1,4-glucanase (EG; AAA34212) and β -glucosidase (BG; AAA18473) from *T. reesei* were expressed in *Pichia pastoris* as previously described (Banerjee et al. 2010b). Pronase E (catalog no. P2714) was purchased from Sigma-Aldrich (St. Louis, MO).

Production of BMCC

The bacterial microcrystalline cellulose (BMCC) was produced based on previous studies with modifications (Krystynowicz et al. 2002; McKenna et al. 2009). HS medium (50 mL) in a 250-mL flask was inoculated with *G. xylinus* colonies and incubated at 30°C for 1-3 days without shaking until a cellulose pellicle was visible. The culture was then vigorously shaken for 30 min to release active cells embedded in the pellicle. After that, 450 mL fresh HS medium in a plastic tray (18 x 6 x 2.5 inch) was inoculated with the culture broth and incubated in stationary culture for another 3-5 days until the cellulose pellicle had grown over the entire medium surface.

The cellulose pellicle was collected by filtration through two layers of Miracloth (Calbiochem, San Diego, CA). The pellicle was then washed repeatedly with distilled deionized water (ddH₂O) to rinse off as much of the medium as possible. The washed cellulose was boiled in 1% NaOH with stirring for 30 min to remove the remaining bacteria. Finally, the alkali-treated cellulose was washed several times with ddH₂O until the water pH reached 7.0, freeze dried, and cut into 1-cm strips. Native BMCC without alkali-treatment was stored in 0.02% sodium azide at 4°C for AFM imaging.

Preparation of BMCC samples for AFM imaging

BMCC samples were prepared by hand-cutting fresh and never dried BMCC film. A bright field microscope was used to select the samples with relative uniform surface. Only thin BMCC layers of ~10-50 μm in thickness were used. The cellulose samples were washed several times with deionized water and placed on poly-lysine coated glass slides (Thermo Fisher Scientific Inc, Waltham, MA, USA) with enough deionized water to cover the samples. The samples were kept in the water during the whole imaging process.

AFM operation

All the AFM experiments were conducted at room temperature on a Dimension AFM equipped with a Nanoscope controller V (Fastscan, Bruker Nano, Santa Barbara, CA) and an acoustic and vibration isolation system. Probes used were SCANASYST-FLUID+ (Bruker, Camarillo, CA USA) for imaging in fluid. The AFM operation software (Nanoscope v9.1) was used to control the scan size, setpoint, and gain. Before AFM imaging, the scanner was carefully calibrated using a calibration kit (Bruker, Camarillo, CA, USA) to make sure all the

measurements were very close to their actual values. In order to dynamically capture the movement of enzyme molecules, the scan rate was normally set at 10-20 Hz for continuous observation at low resolution (256 x 128 pixels or 128 x 64 pixels). However, when monitoring the penetration by single TrAA9A molecules, the scan rate was lowered to 2 Hz to obtain images with better resolution. For other static observations, the scan rate was 1 Hz at a resolution of 1024 x 1024 pixels.

AFM image processing

All AFM image data were analyzed off-line with Nanoscope Analysis v1.8 software (Bruker Nano, Santa Barbara, CA). The height and peak force error images were plane-fitted at 1 order for the images in all figures. The color bar was manually modified according to the best presence of each image. The images used for height and width measurements were raw data without any processing and all width and height measurements were conducted with the “section” function in the Nanoscope Analysis software.

Preparation of TrAA9A-treated BMCC

BMCC (2 mg/mL) was incubated with 100 µg/mL TrAA9A, 1 mM L-ascorbic acid and 0.02% sodium azide for 72 h in 50 mM, pH 4.8, sodium acetate buffer. All reactions were conducted in triplicate at 150 rpm and 50°C in a shaking incubator. After 72 h incubation, the BMCC residue was recovered by filtering through two layers of Miracloth (Calbiochem, San Diego, CA). Reaction filtrates were saved for TrAA9A product analysis. The TrAA9A remaining on the cellulose residue was removed as described in a previous study with modifications (Yang et al. 2006). The residue was first washed several times with ddH₂O and

then incubated with 10 mg/mL Pronase E overnight at 100 rpm, 37 °C in 50 mM Tris, pH 7.5, to remove the remaining TrAA9A. The BMCC residue was collected by filtering through Miracloth and washed with ddH₂O, 1 M NaCl, and ddH₂O in order to remove Pronase E. Finally, the BMCC residue was freeze-dried and stored at 4°C.

TrAA9A product analysis by mass spectrometry (method details provided by MSU RTSF)

TrAA9 reaction filtrates obtained after 72 h reaction with BMCC were centrifuged at 4°C for 20 min in 70% ethanol to precipitate TrAA9A, and the supernatant was collected for electrospray ionization (ESI) mass spectrometry (MS) analysis. The supernatant of BMCC incubated under the same condition without TrAA9A was also prepared in the same way as a control. Samples were analyzed on a Waters Xevo G2-XS Q-TOF system coupled to a Waters I-Class UPLC system. Carbohydrates were separated by an ACQUITY UPLC BEH Amide column (2.1 x 100 mm, 1.7 µm) maintained at 40°C, with an injection volume of 10 µL. Solvent A for chromatography was 10 mM ammonium formate, and solvent B was 100% acetonitrile. The solutes were eluted at 0.2 mL/min starting at 95% B, followed by a linear gradient to 35% B over 14 min. The proportion of solvent B was maintained at 35% for 2 min and then increased back to 95% and kept for 4 min for re-equilibration.

The operation conditions for mass spectrometer was a capillary voltage of 3.0 kV, sample cone voltage of 80 V, source temperature of 100 °C, desolvation temperature of 350 °C, and desolvation gas flow of 600 L/hr. Mass spectra were acquired in positive ion mode across the 50-2000 m/z range. Data processing was performed using MassLynx v4.1 software.

Enzymatic degradation of BMCC

To test the synergism between TrAA9A and TlCel7A on cellulose degradation, 2 mg/mL BMCC was incubated with 34 µg/mL TlCel7A and 6 µg/mL TrAA9A, alone or together, in 10 mL reactions containing 1 mM L-ascorbic acid and 0.02% sodium azide. Reactions were conducted in triplicate for 72 h at 150 rpm, 50°C in 50 mM sodium acetate, pH 4.8. To measure the cellulose conversion rate, 100 µL reaction supernatant was incubated with 40 µg/mL β-glucosidase for 30 min at 50°C in 50 mM sodium acetate, pH 4.8. Free glucose was measured by enzyme-linked colorimetry as described (Banerjee et al. 2010b).

To test the digestibility of TrAA9A-treated cellulose, TrAA9A-treated BMCC was compared against untreated BMCC for enzymatic hydrolysis at pH 4.8. For all reactions, cellulose loading was 2 mg/mL and cellulase mixture loading 100 µg/mL. The cellulase mixture contained TlCel7A, EG, and BG in a 6:3:1 ratio. Each reaction was performed in triplicate at 150 rpm, 50°C in 0.02% sodium azide. Glucose yields were measured during 72 h incubation.

Results

Enhancement of TlCel7A hydrolysis of BMCC by TrAA9A

CBH1 is the major cellulase that degrades crystalline cellulose (Lynd et al. 2002; Teeri 1997). LPMO has been reported to enhance CBH1 hydrolysis of pretreated biomass (Banerjee et al. 2010c; Harris et al. 2010). Under our experimental conditions, a mixture of TrAA9A and TlCel7A converted 8% more cellulose (46% to 54% of the maximum potential cellulose conversion) than TlCel7A alone (Figure 4.1). This result was consistent with previous work showing that hydrolysis of cellulose in pretreated corn stover was increased by approximately

6% after adding a *Thielavia terrestris* LPMO to *T. reesei* CBH1 (Harris et al. 2010). We therefore considered that TrAA9A used in this study exhibited similar activity in enhancing cellulose hydrolysis by TlCel7A.

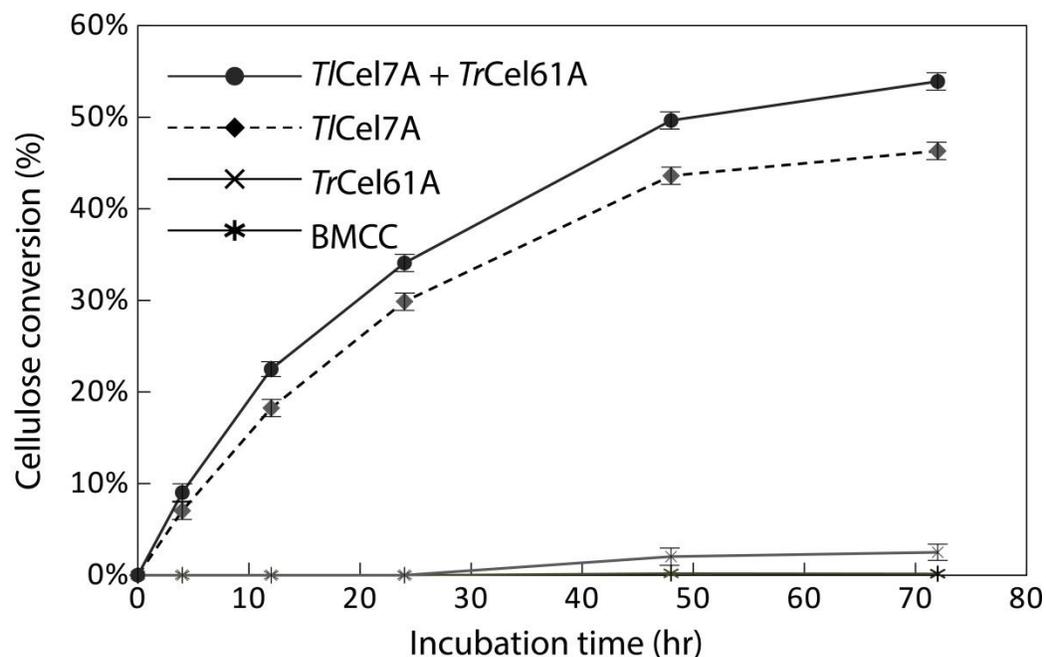


Figure 4.1 Cellulose conversion by TrAA9A and TlCel7A. BMCC (2 mg/mL) was digested with TlCel7A (34 $\mu\text{g/mL}$) alone or together with TrAA9A (6 $\mu\text{g/mL}$). The reaction was carried out at 50°C with agitation at 150 rpm. The error bars represent the standard deviation of triplicates.

MS analysis of TrAA9A products

LPMO degrades cellulose to cello-oligosaccharides and oxidized cello-oligomers. Three types of LPMOs have been described based on their oxidation products (Figure 4.2). Type 1 enzymes oxidize the C1 carbon of the glucose unit and produce aldonolactone and its hydrated product aldonic acid (Agger et al. 2014; Forsberg et al. 2011; Vaaje-Kolstad et al. 2010; Westereng et al. 2011). Type 2 enzymes oxidize the C4 carbon and generate a gem-diol intermediate and 4-ketoaldose (Isaksen et al. 2014; Li et al. 2012; Phillips et al. 2011). Type 3

oxidizes both C1 and C4 carbons and can generate the C1 oxidation product, the C4 oxidation product, and the C1-C4 double oxidized products (Agger et al. 2014; Beeson et al. 2012; Bennati-Granier et al. 2015; Frommhagen et al. 2016; Li et al. 2012; Vu et al. 2013). C1 or C4 oxidation alone can lead to the breakage of the cellulose chain.

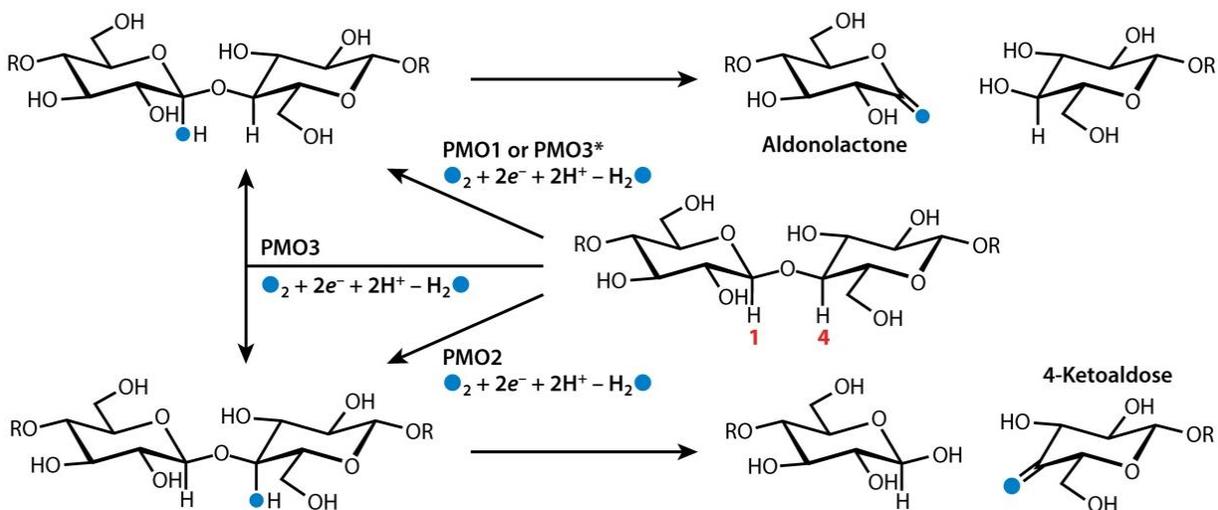


Figure 4.2 Regioselective hydroxylation of cellulose by lytic polysaccharide monooxygenases. PMO stands for polysaccharide monooxygenase. Reprinted from Beeson et al. (2015) with permission.

To further confirm the activity of TrAA9A, reaction products after 72 h incubation of TrAA9A with BMCC were analyzed by mass spectrometry (Figure 4.3). After 72 h incubation, BMCC alone released a small amount of cellobiose (DP2) but no other cellodextrin or oxidized sugar (Figure 4.3a). In contrast, TrAA9A generated cellobiose (DP2), cellotriose (DP3) and two types of oxidation products derived from DP2 and DP3 after 72 h reaction (Figure 4.3b, b, d). Cellodextrin and oxidized sugar products were all associated with Na^+ in the mass spectra probably because the reactions were conducted in sodium acetate buffer (Bey et al. 2013). Although there was a small amount of cellobiose in the negative control, its relative intensity was

much lower than that from TrAA9A-containing reactions. Therefore, cellobiose is probably one of the products of TrAA9A. In contrast to many other studies, cello-oligomers and oxidized cello-oligomers of higher degree of polymerization (DP) were not detected after 72 h treatment with TrAA9A. One possible explanation is that TrAA9A might also be active on cello-oligosaccharides (Isaksen et al. 2014) and therefore that the cellodextrins of higher DP were oxidized after 72 h. Due to the identical masses between gem-diol and aldonic acid ($DPx + 16$ amu, DPx represents cellodextrin), and between ketoaldose and aldonolactone ($DPx - 2$ amu), MS analysis alone cannot determine the type of TrAA9A activity (Isaksen et al. 2014).

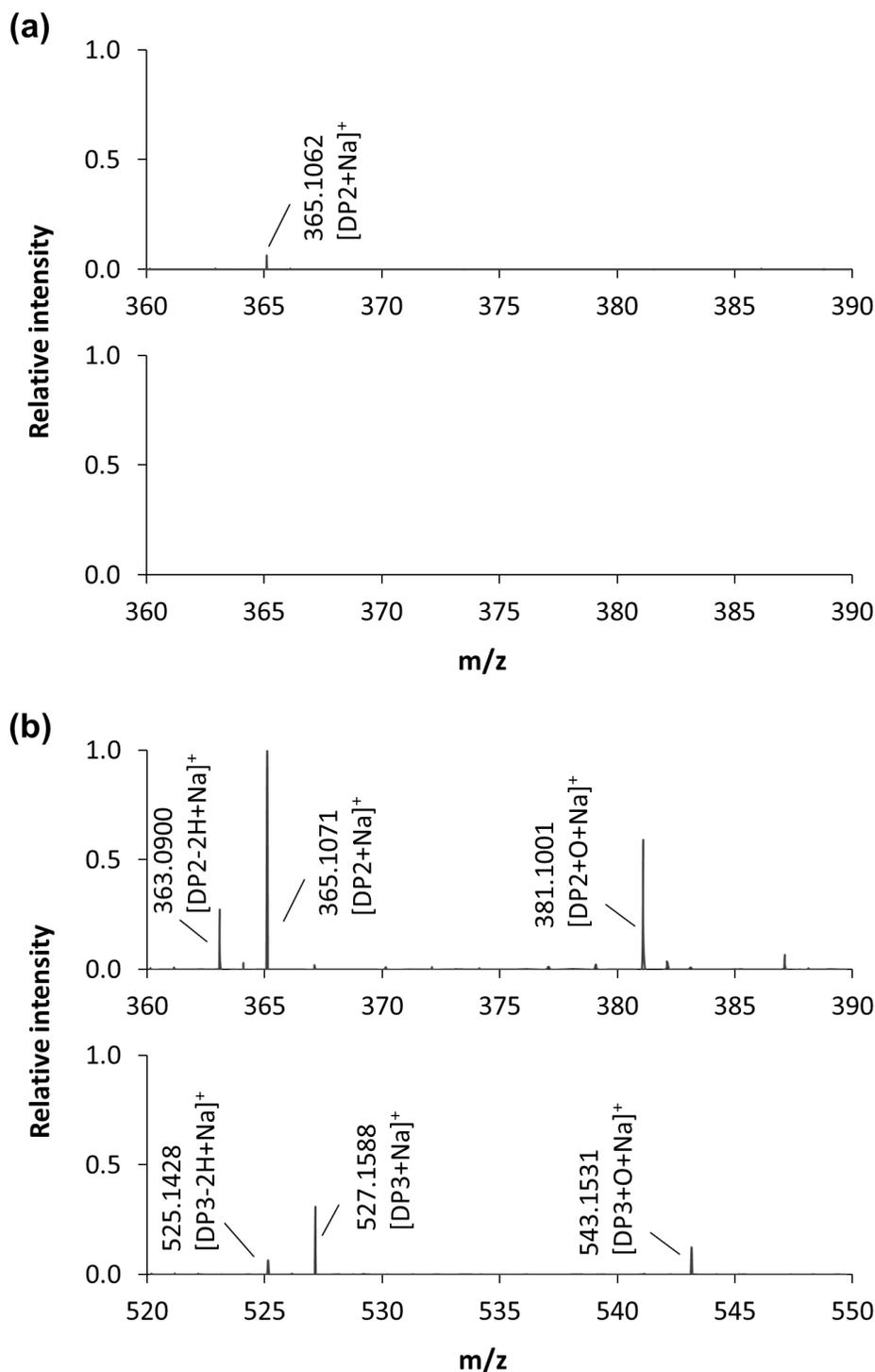
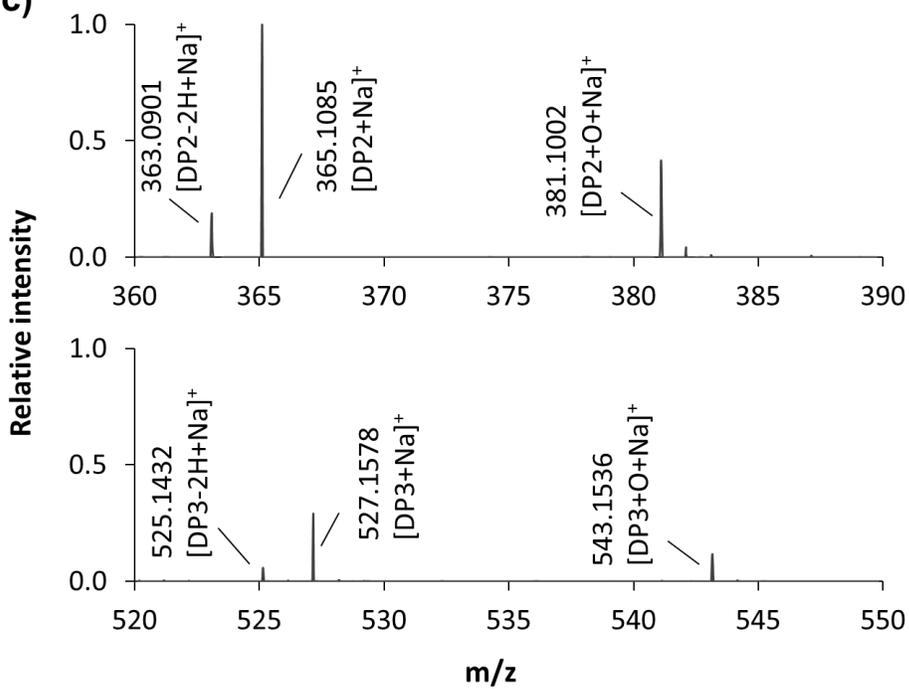


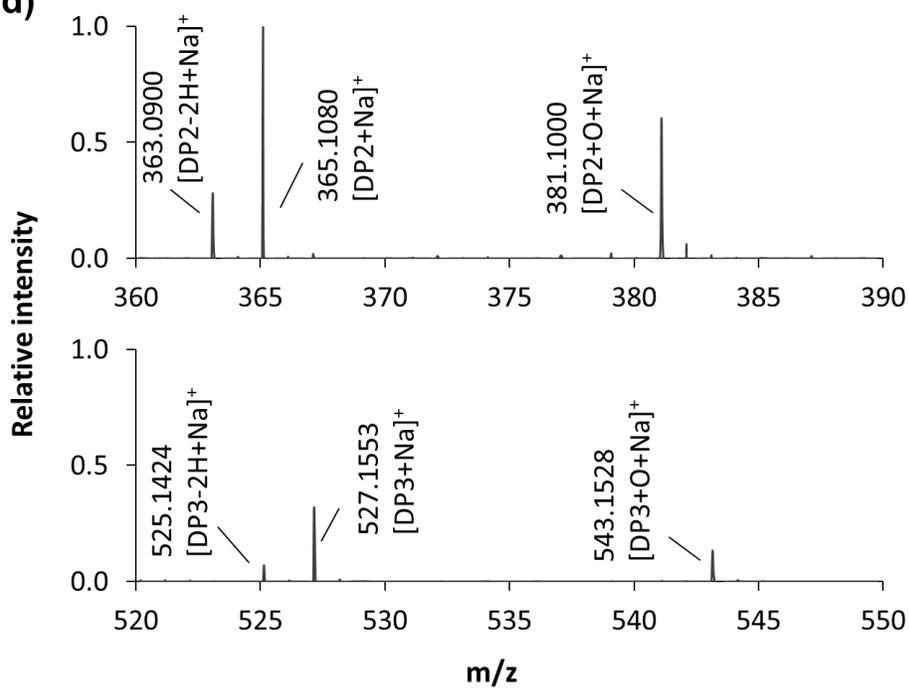
Figure 4.3 ESI-MS analysis of oxidized cellobiose and cellotriose released from BMCC by TrAA9A. (a) Small amounts of cellobiose were detected in the negative control. (b), (c) and (d) Cellobiose (DP2), cellotriose (DP3) and two oxidation products derived from DP2 and DP3 were detected in all replicates of TrAA9A-containing samples.

Figure 4.3 (Cont'd)

(c)



(d)



Molecular action of TrAA9A on BMCC

To visualize the interaction of TrAA9A with BMCC under AFM, a small piece of BMCC was transferred onto a glass slide pre-coated with poly-lysine. As a negative control, BMCC was at first imaged in the same buffer used for the enzyme reactions. In the BMCC control, we observed large ribbon-like cellulose structures with a relatively smooth surface. Cellulose ribbons were 30-150 nm in width. About 10 min after adding TrAA9A, particles of 6-10 nm in diameter accumulated on the surface of the BMCC ribbon (Figure 4.4). TrAA9A was predicted to be ~6 nm in diameter based on the structure of *TaGH61* (Chen and Makhatadze 2015; Quinlan et al. 2011), and we therefore considered these particles to be bound TrAA9A molecules.

Continuous imaging revealed that TrAA9A molecules initially bound randomly to the BMCC surface, and individual TrAA9A molecules exhibited a “stop-and-go” behavior. This means that the enzyme moved along the cellulose but stopped intermittently before starting again (Igarashi et al. 2011). This is thought to be due to temporary blocks that prevent the enzyme from progressing. Interestingly, the enzyme molecules stayed in the approximately same position for a much longer time than they spent moving along the cellulose. TrAA9A moved at 0.25 ± 0.13 nm/s ($n = 40$) in three patterns: in both directions along one cellulose ribbon, across one ribbon, or moving from one ribbon to another (Figure 4.4a, b and c, respectively; details in video S1-S3 in Song et al. 2018). The observed speed of TrAA9A was one order of magnitude slower than previously observed for CBH1 moving on *Valonia* microcrystalline cellulose (TrCel7A; Igarashi et al. 2011). This might be related to the slow rate of oxidization catalyzed by TrAA9A (Westereng et al. 2011). When the imaging period extended to ~7 h, some TrAA9A molecules seemed to form into groups that moved together towards a particular cellulose ribbon where eventually many molecules accumulated (Figure 4.4c and Video S3 in Song et al. 2018).

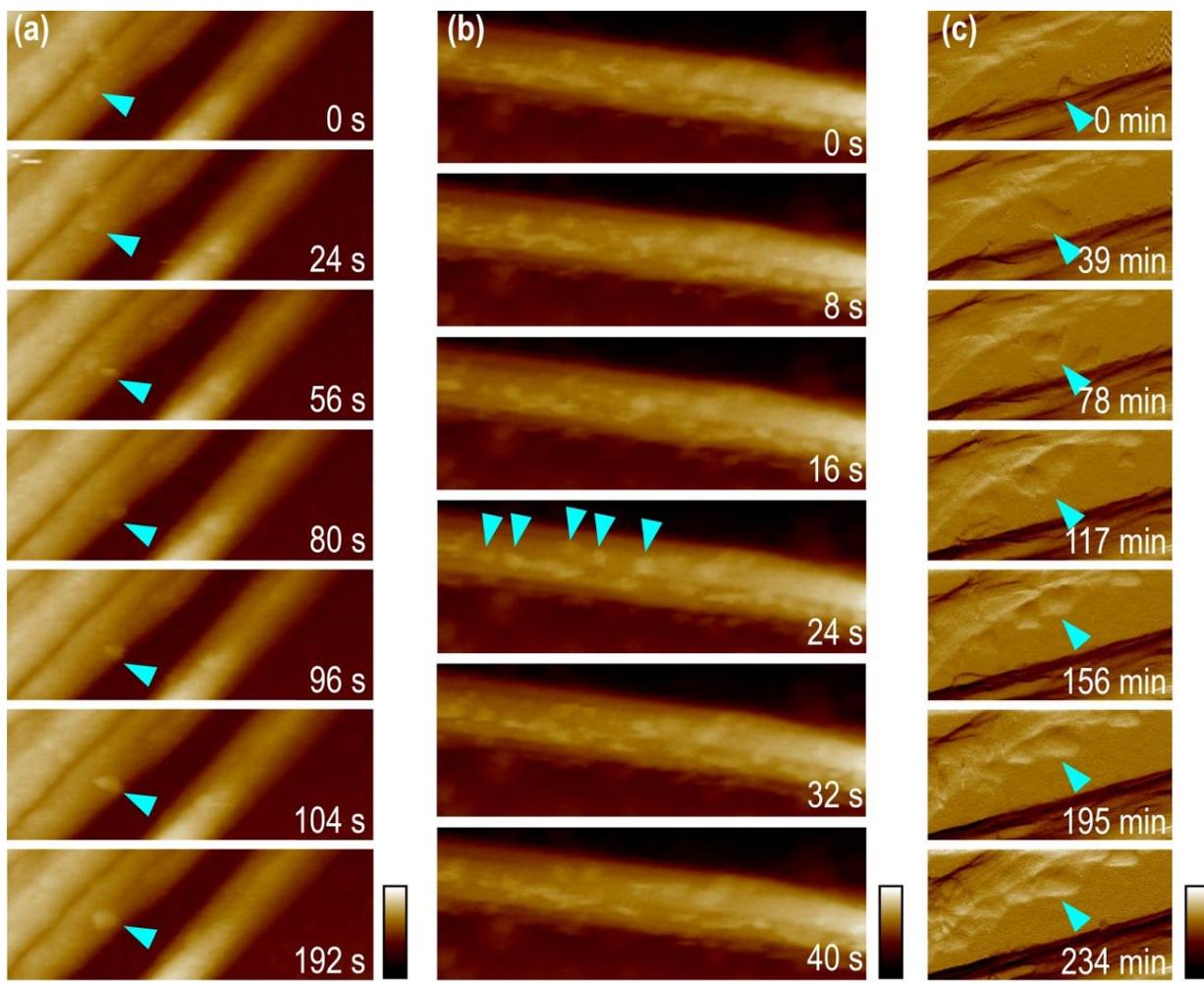


Figure 4.4 TrAA9A molecules move and diffuse randomly on BMCC. Time-lapse images from video S1-3 show TrAA9A molecules moving across (a), along (b), and diffusing (c) between cellulose ribbons. Cyan arrows indicate individual enzyme molecules. Scale bar = 150 nm. Color bars are 55 nm in (a) and (b), and 100 pN in (c).

Previously, AFM imaging revealed that TrCel7A binds and reacts on the hydrophobic face of crystalline cellulose (Liu et al. 2011) and moves processively from the reducing end of the glucan chain to the non-reducing end (Igarashi et al. 2011). In contrast, TrAA9A moves in both directions along and also across the ribbons. The different binding and movement patterns of TrAA9 and TrCel7A observed in this study could be explained by the differences in their molecular structures. The binding face of LPMO is flat (Beeson et al. 2015; Hemsworth et al.

2015; Karkehabadi et al. 2008), whereas CBHs like TlCel7A adopt a tunnel-shape active site (Divne et al. 1998) that can hold several glucose residues of the cellulose chain and thereby constrain movement to a linear fashion.

By monitoring the height change in the region that individual TrAA9A enzymes moved across, we observed the unexpected phenomenon of enzyme molecules penetrating inside the ribbon (Figure 4.5 and video S4 in Song et al. 2018). After addition of TrAA9A, there was an initial increase in height due to the bound enzyme (Figure 4.5c). Subsequently, TrAA9A molecules penetrated inside the cellulose ribbon and were hardly visible after 27 min. The enzyme gradually came back to the surface after 96 min, moved along the cellulose surface, and diffused away after 210 min. As TrAA9A moved, groove-shape features appeared on the cellulose microfibril surface and the existing grooves deepened and widened. As a result, more obvious edges appeared on the cellulose surface after 2 h incubation with TrAA9A (Figure 4.6 and video S5 in Song et al. 2018).

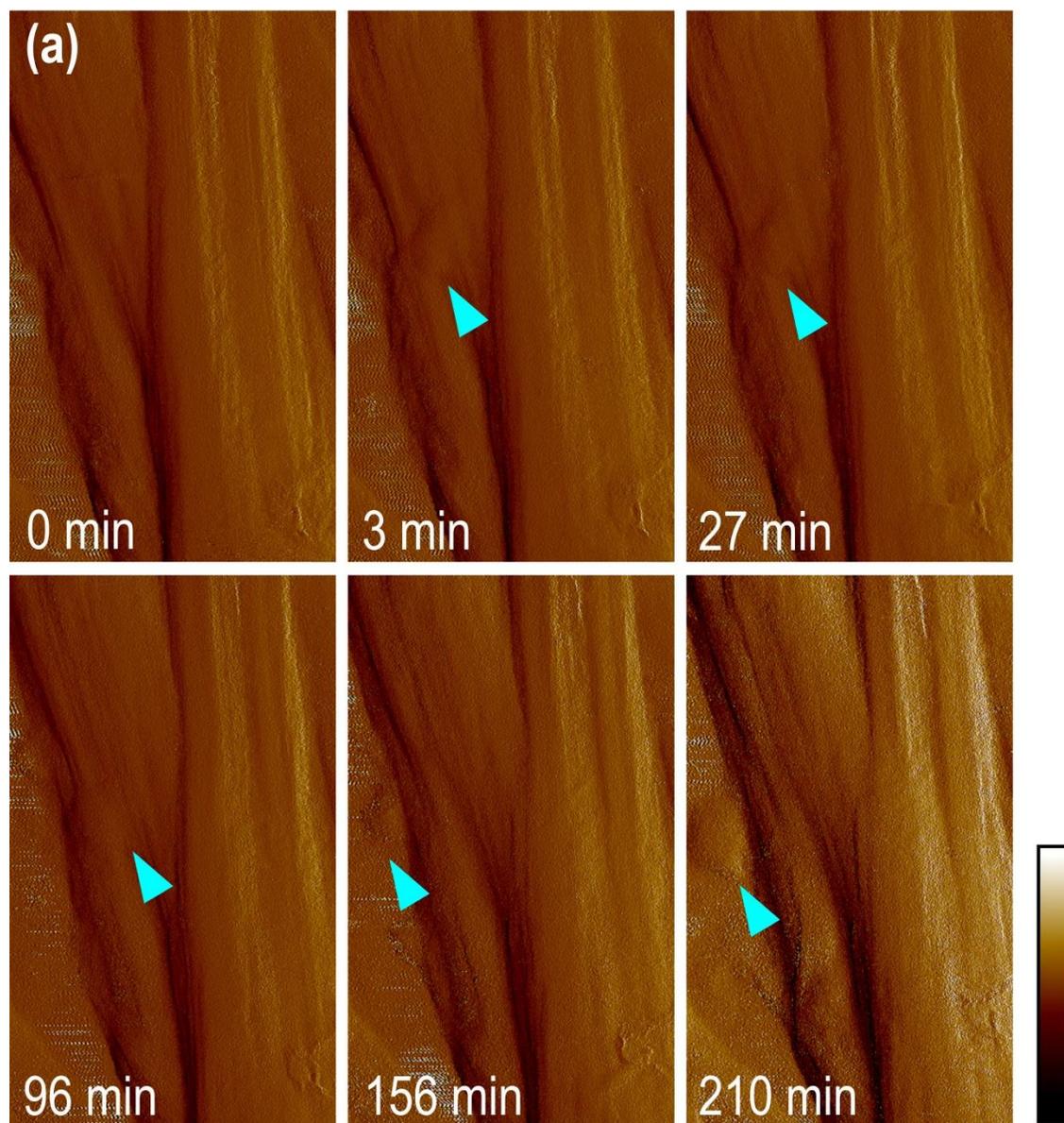


Figure 4.5 TrAA9A penetrating and moving inside a BMCC ribbon. Time-lapse AFM peak force error (a) and height (b) images showing TrAA9A (indicated by the cyan arrows) moving in and out of the surface of a cellulose ribbon (See video S4 in Song et al. 2018 for more information). (c) Relative height measured across a TrAA9A molecule during 210 min incubation. Scale bar is 100 nm; color bar is 1.1 nN in (a) and 75 nm in (b).

Figure 4.5 (Cont'd)

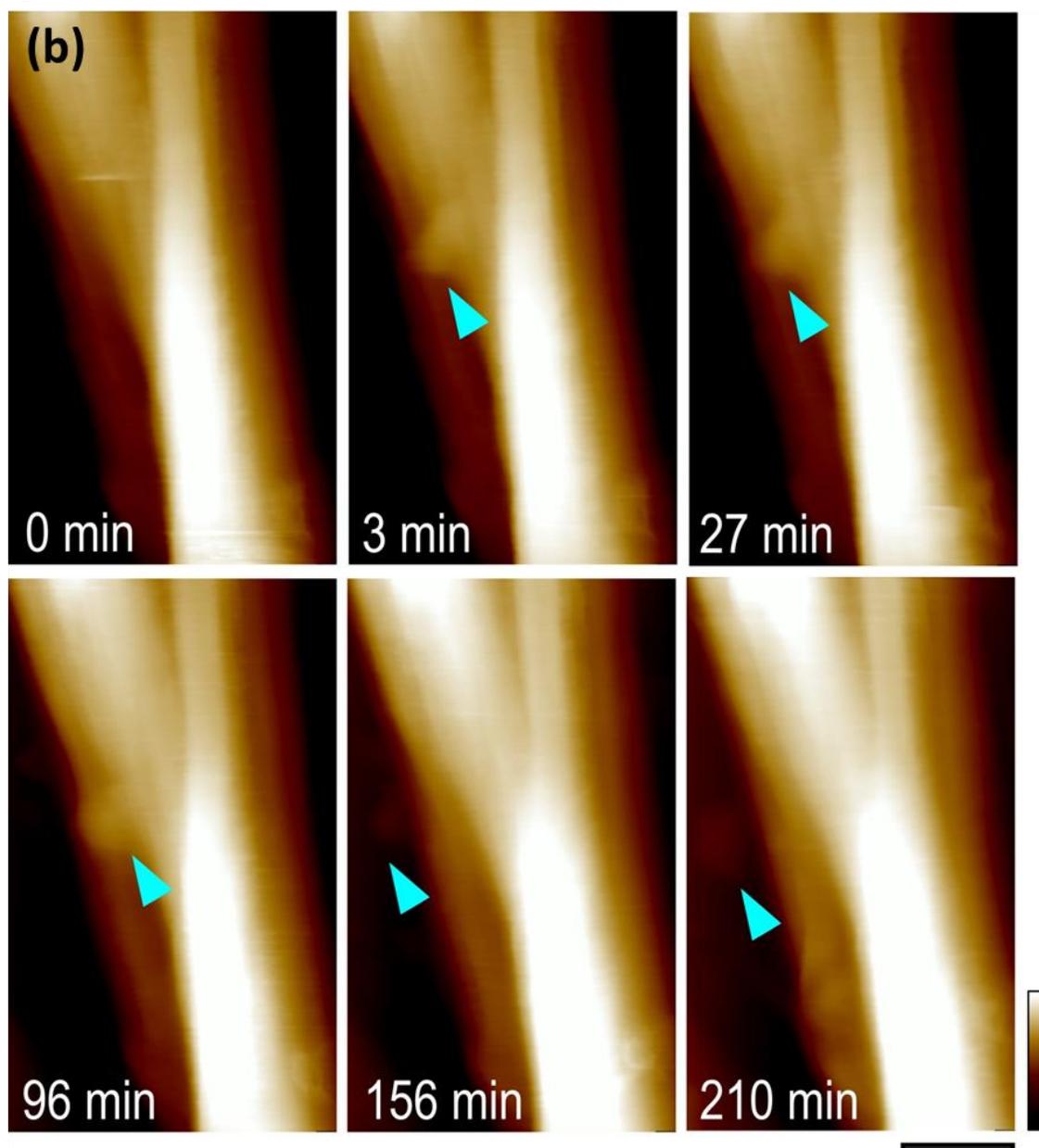


Figure 4.5 (Cont'd)

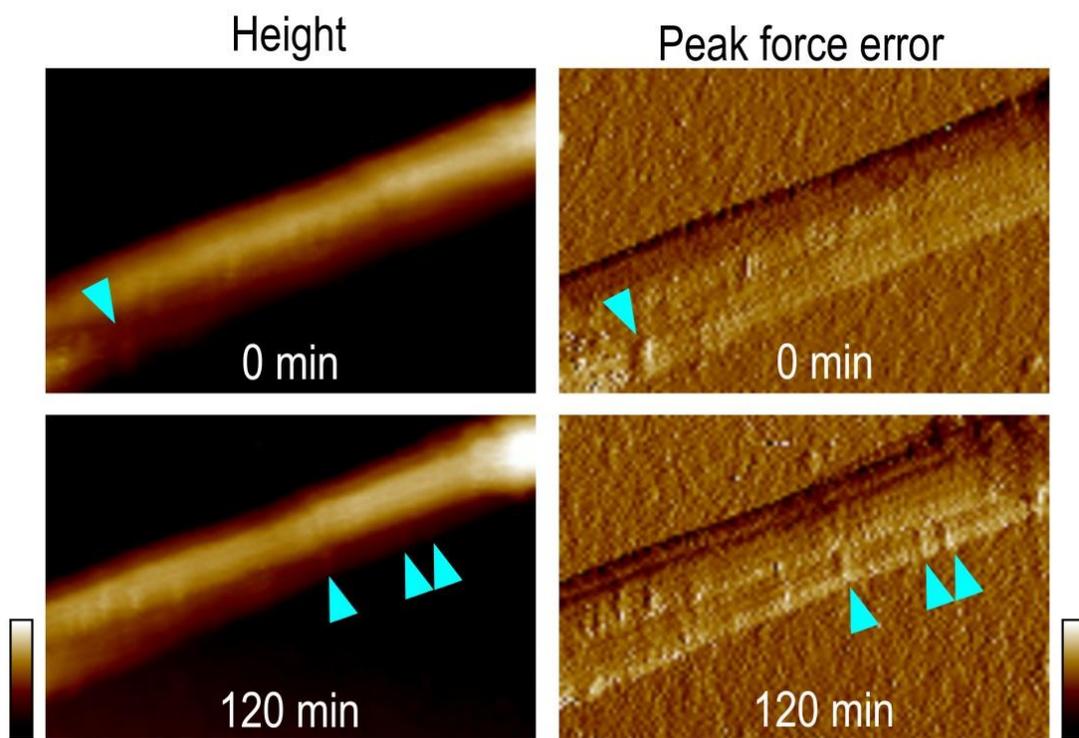
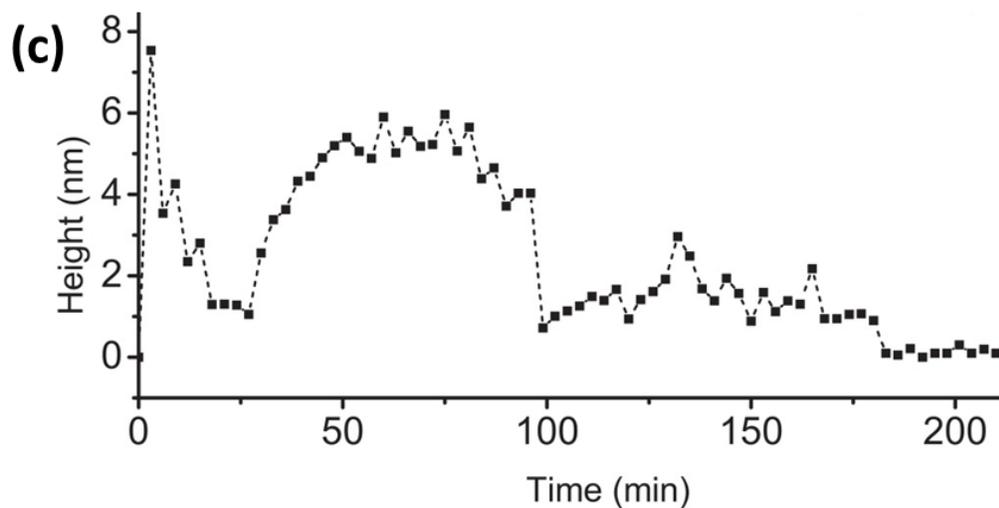


Figure 4.6 Changes in morphology of BMCC ribbon when incubated with TrAA9A (indicated by cyan arrows) during 2 h continuous AFM observation. Pictures were taken from Video S5. Scale bar is 50 nm. Color bar is 50 nm and 830 pN in height (left) and peak force error (right) channels, respectively.

Disassembly of BMCC ribbon into smaller fibrils by TrAA9A

The width of cellulose microfibrils was measured during incubation with TrAA9A (Figure 4.7). After 4 h treatment, the average width of the cellulose ribbons slightly decreased from 101.8 ± 14.9 nm ($n = 20$) to 89.8 ± 16.7 nm ($n = 16$), but the surface roughness increased greatly, with obvious grooves and edges observed on the surface (Figure 4.7b) compared to a relatively smooth surface before treatment (Figure 4.7a). After 24 h incubation, we observed splitting of the ribbon-like cellulose structures along its long axis into smaller microfibrils of 52.9 ± 11.7 nm ($n = 24$) in width, which was approximately 52% of their original ribbon width. Similar structural changes were also observed recently after treating bleached softwood Kraft pulp with LPMO (Villares et al. 2017). We postulate that the ability of TrAA9A to penetrate inside the cellulose ribbon leads to the dividing of a large cellulose ribbon into multiple smaller microfibrils. Considering that a BMCC ribbon is composed of multiple cellulose elementary fibrils (CEF) that can be as small as 2-4 nm (Martínez-Sanz et al. 2015, 2016), it is possible that TrAA9A oxidation disrupts the glucan chains between these CEFs.

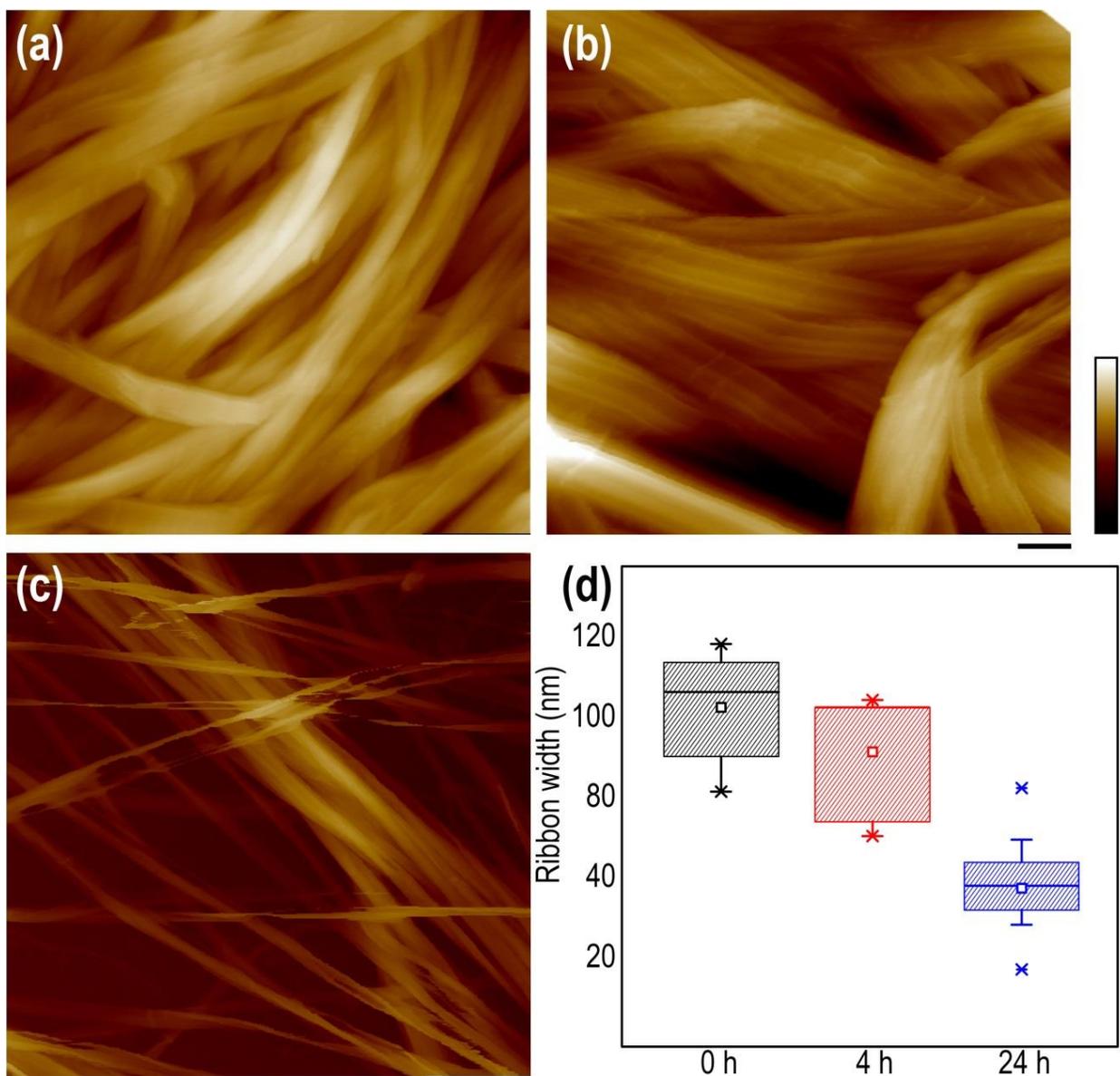


Figure 4.7 Disassembly of BMCC ribbon into smaller fibrils by TrAA9A. Atomic force micrographs of BMCC before (a), after 4 h (b) and 24 h (c) incubation with TrAA9A. (d) The width (and standard deviation) of BMCC microfibrils before and after TrAA9A treatment. Scale bar is 150 nm and color bar is 200 nm.

Synergism between TrAA9A and TlCel7A during hydrolysis of BMCC

We further studied the overall BMCC structural changes by incubating TlCel7A with and without TrAA9A. After 72 h incubation with both enzymes, BMCC strips were broken into small pieces (Figure 4.8). In contrast, treatment with TlCel7A alone resulted in the shortening of BMCC strips floating on the top of the reaction buffer. There was no significant structural change to BMCC visible by eye after TrAA9A treatment alone (Figure 4.8).

AFM was used to visualize the structural changes of individual BMCC ribbons at a nanometer scale after treatment with these two enzymes (Figure 4.9). For the first 20 min, the average height and width of the cellulose ribbon slightly increased, which was likely due to the binding of TrAA9A and TlCel7A onto the cellulose surface. After 7 h incubation, the height of the BMCC ribbon decreased while the width increased. A significant decrease in height and increase in width was observed from 106 min to 149 min, which was concomitant with the disintegration of the large cellulose ribbon into smaller microfibrils. Between 213 and 255 min, the cellulose ribbon completely separated into small microfibrils of ~4 nm in diameter, as the ribbon height continued to decrease and the width to increase (Figure 4.9a; details in video S6 in Song et al. 2018). The diameter of the small microfibril (~4 nm) indicated they could contain an individual or a couple of fundamental CEFs (Martínez-Sanz et al. 2015, 2016).

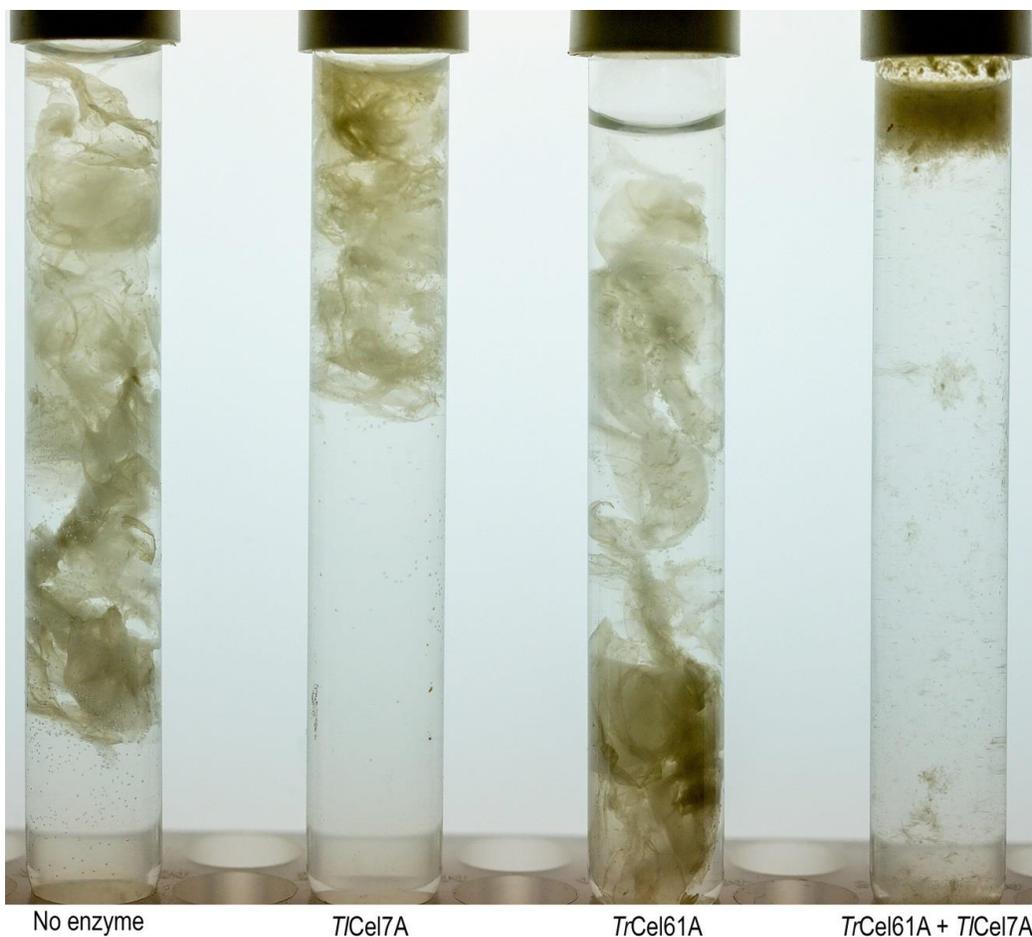


Figure 4.8 Morphology change of BMCC after 72 h incubation with or without enzymes. BMCC was cut into 1-cm strips. Reaction conditions were the same as Figure 4.1. There were no obvious morphological changes of cellulose strips after TrAA9A (TrCel61A) treatment alone. In contrast, after 72 h incubation with TlCel7A, the length of the strip shortened to approximately half of its original length with almost 50% cellulose degradation (Figure 4.1). After 72 h incubation with both TrAA9A and TlCel7A, the cellulose strip was broken into small insoluble particles mostly floating on the surface.

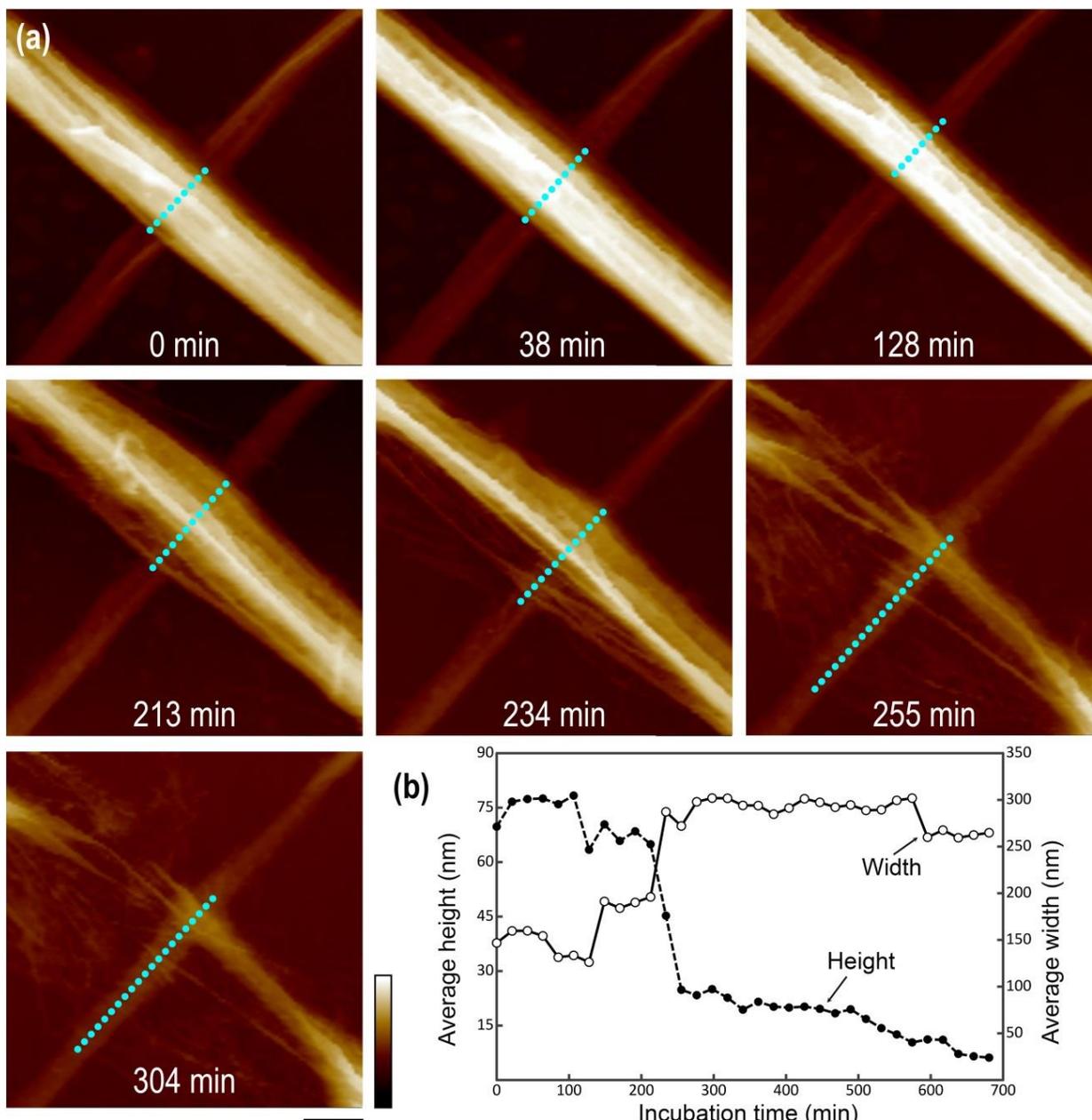


Figure 4.9 Synergism between TrAA9A and TlCel7A during hydrolysis of BMCC. (a) Time-lapse images from video S6 (Song et al. 2018) showing changes in the appearance of the cellulose ribbon during incubation with TrAA9A and TlCel7A. (b) Height and width measurements. Cyan dashed lines in (a) indicate the overall width change of a BMCC ribbon during enzymatic treatment. Scale bar is 50 nm and color bar is 100 nm.

After 4 h of incubation, the resulting cellulose microfibrils were significantly thinner when they were treated with TrAA9A together with TlCel7A (Figure 4.9a) compared to when

treated with TrAA9A alone (Figure 4.7). Interestingly, cellulose hydrolysis by a cellulase mixture composed of TrCel7A, EG, and BG increased from 79% to 87% when using TrAA9A-pretreated BMCC, compared with untreated BMCC (Figure 4.10). This enhancement is similar to the result when TrAA9A and TrCel7A were added together to untreated BMCC, suggesting that the synergy of TrAA9A and cellulases during cellulose hydrolysis may not be attributed to a direct molecular interaction between these enzymes.

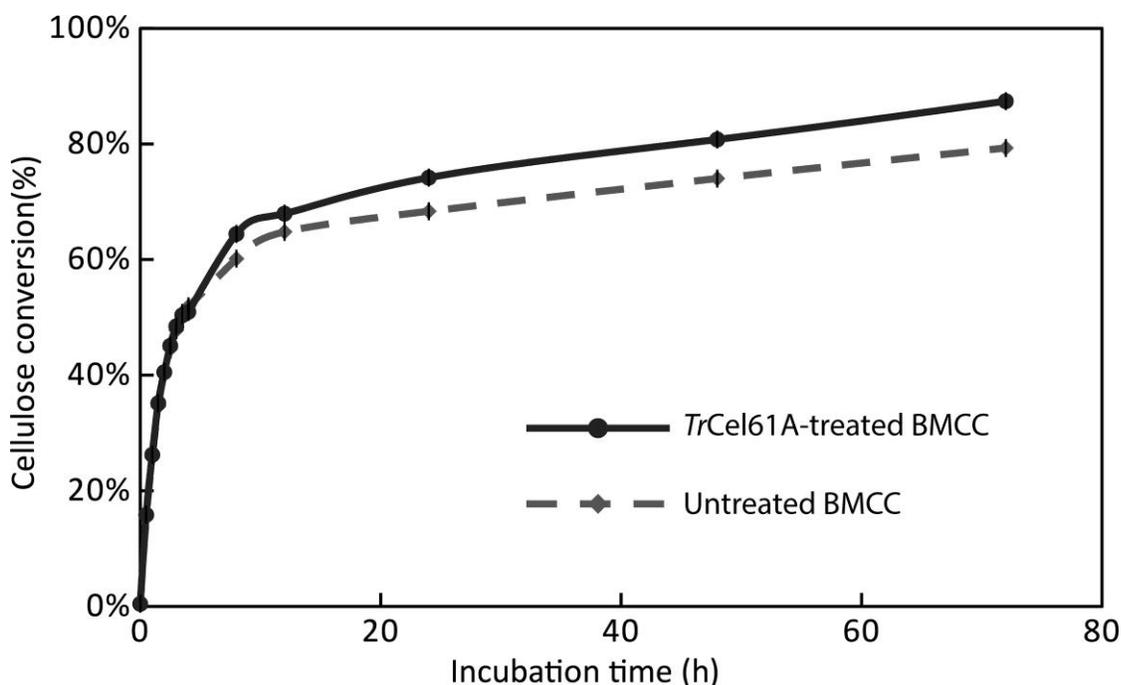


Figure 4.10 Comparison of cellulase hydrolysis of untreated and TrAA9A-treated BMCC. Cellulose loading was 2 mg/mL and total cellulase loading was 100 μ g/mL with TrCel7A: EG: BG ratio of 6:3:1. The reactions in triplicate were conducted in 50 mM sodium acetate, pH 4.8, at 150 rpm and 50°C for 72 h.

Discussion

We used high-speed AFM imaging to visualize in real time enzymatic hydrolysis of cellulose by cellulases and an LPMO, TrAA9A. Our results are consistent with previous studies

in which cellulases, i.e., CBHs and Egs, are responsible for the majority of cellulose hydrolysis, and LPMO promotes the efficiency of cellulases. We observed that LPMO reacts specifically to disintegrate large cellulose ribbons into small microfibrils with little oxidized sugar products, whereas the action CBH1 has been considered to be a “peel-off” process (Liu et al. 2011, White and Brown 1981). We hypothesize that LPMO and CBH1 may attack different structures of cellulose. BMCC ribbons are considered to contain amorphous cellulose and CEFs that are composed of well-organized cellulose glucan chains (Martínez-Sanz et al. 2015, 2016). It is likely that LPMO attacks the crystalline proportion of cellulose and increases the surface accessibility of CEFs for cellulase enzymes. If this is true, LPMO would not enhance cellulase activity when cellulose accessibility is already at a maximum. Indeed, it has been previously demonstrated that the degree of LPMO enhancement is negatively correlated with cellulose accessibility (Hu et al. 2014). Similarly, LPMO has been found to increase the degradation of insoluble xylan (Kim et al. 2016), which is one of the major matrix polysaccharides in the plant cell wall, suggesting that LMPO can also oxidize sugars in other polysaccharides that cross-link to each other and to cellulose, thus exposing the well-organized cellulose microfibrils to enhanced hydrolysis by cellulases.

CHAPTER 5: FUTURE DIRECTIONS

The high enzyme loading required for biomass conversion to fermentable sugars is one of the biggest hurdles to producing bioethanol economically from lignocellulosic substrates. In my Ph.D. research, I explored novel fungal cellulolytic enzymes of high efficiency in order to reduce the amount of enzymes needed for efficient biomass degradation. My work showed that two bi-function GH5_5_2 endoglucanases/mannanases were able to improve the overall performance of enzyme mixtures used for biomass conversion, possibly by degrading the strong enzyme inhibitor mannan. Since several hemicelluloses are proved to inhibit enzymatic degradation of lignocellulosic substrates, degradation of hemicelluloses during cellulose hydrolysis may help improve the overall efficiency of enzyme mixtures.

Several hemicelluloses have been shown to inhibit enzymatic degradation of lignocellulose. A feasible strategy to degrade these enzyme inhibitors is to prospect for bi- or multi-functional enzymes that degrade not only cellulose but also hemicelluloses. Another advantage of using multi-functional enzymes is to reduce the work required to optimize the composition of enzyme mixtures for the degradation of different biomass feedstocks. The GH5 family contains more than 10,000 enzymes active on cellulose and various other wall polysaccharides (<http://www.cazy.org/GH5.html>). This large family may contain multispecific enzymes for efficient biomass degradation. As shown in my work, GH5_5_2 enzymes possess endo-glucanase and mannanase activity. TmCel5A from subfamily 25 of GH5 family also exhibits activity on cellulose and mannan (Wu et al. 2011). Another example is CtCel5E from subfamily 4 that is active on both cellulose and xylan (Yuan et al. 2015). Xylan, xylooligosaccharides, mannan and mannooligosaccharides are reported to be strong cellulase

inhibitors. To find cellulases that are active on xylan and/or mannan, it is worthwhile to characterize enzymes from these three subfamilies.

Our work on TrAA9A, which is an LPMO, showed that TrAA9A improved TICel7B efficiency in cellulose hydrolysis by splitting bacterial cellulose ribbons into smaller cellulose fibrils, thereby increasing the accessible cellulose surface for enzymatic degradation. Interestingly, TrAA9A also shows synergism with xylanase on insoluble xylan degradation (Kim et al. 2016). To find more LPMOs that not only synergize with cellulases but also with enzymes required for hemicellulose or lignocellulose deconstruction, it is worthwhile to characterize other LPMOs from AA9 family and test their synergism with hemicellulases on the degradation of different hemicelluloses, especially on known enzyme inhibitors such as xylan and mannan. It is also necessary to study the mechanism of the synergism between LPMO and other hemicellulases for future enzyme engineering to obtain highly efficient biomass-degrading enzymes.

For *C. bulbillosum* alkaliphilic cellulolytic enzymes, my work showed that two *C. bulbillosum* cellulases were more active than their *T. reesei* homologs at pH 6-8 on a soluble substrate. To explore the potential of *C. bulbillosum* enzymes on lignocellulosic substrate degradation, it is important to purify other cellulolytic enzymes including CBH1, CBH2, and LPMO, construct enzyme mixtures, and compare them against the enzyme mixtures of their *T. reesei* homologs on lignocellulose deconstruction at high pH.

Although cellulolytic enzymes have been studied for more than 70 years for biomass degradation, thorough digestion of lignocellulosic biomass at economic enzyme loading is still challenging due to biomass recalcitrance, which is partially caused by the structure and composition of the plant cell wall. Instead of lowering enzyme loading for biomass conversion,

another way to reduce bioethanol production cost is to reduce the cost to produce large amount of the enzymes. Currently, commercial enzymes are obtained by fermentation of filamentous fungi such as *T. reesei* where the production facility-related expenses are estimated to account for approximately half of the bioethanol production cost (Klein-Marcuschamer et al. 2012). A novel idea would be to lower the enzyme production cost by expressing the cellulose-degrading enzymes in crop plants, due to the lower cost of cultivating plants compared to fungal fermentation. Cellulose-degrading enzymes could be expressed and sequestered in plant organelles such as the vacuole or chloroplast. After harvesting, these enzymes could be extracted from the plants without loss of their activities. Another idea would be to design biofuel plants that can self-degrade after harvesting. One important step that makes this idea possible is the successful engineering of poplar to produce lignin that is more amenable to degradation (Wilkerson et al. 2014). In summary, more research is required to produce bioethanol economically from lignocellulosic biomass.

REFERENCES

REFERENCES

- Agarwal AK (2007) Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Progress in energy and combustion science*, 33(3): 233-271.
- Agbor VB, Cicek N, Sparling R, Berlin A and Levin DB (2011) Biomass pretreatment: Fundamentals toward application. *Biotechnology advances* 29:675–685.
- Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R et al. (2014) Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proceedings of the national academy of sciences*, 111(17): 6287-6292.
- Allen F, Andreotti R, Eveleigh DE, Nystrom J (2009) Mary Elizabeth Hickox Mandels, 90, bioenergy leader. *Biotechnology for biofuels*, 2(1): 22.
- Alvira, P, Tomás-Pejó E, Ballesteros MJ, Negro MJ (2010) Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresource technology*, 101(13): 4851-4861.
- Anderson JE, DiCicco DM, Ginder JM, Kramer U, Leone TG, Raney-Pablo HE et al. (2012) High octane number ethanol–gasoline blends: Quantifying the potential benefits in the United States. *Fuel*, 97: 585-594.
- Arantes V, Saddler JN (2011) Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnology for biofuels*, 4(1): 3.
- Arora R, Manisseri C, Li C, Ong MD, Scheller HV, Vogel K et al. (2010) Monitoring and analyzing process streams towards understanding ionic liquid pretreatment of switchgrass (*Panicum virgatum L.*). *BioEnergy Research*, 3(2): 134-145.
- Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC evolutionary biology*, 12(1): 186.
- Banerjee G, Scott-Craig JS, Walton JD (2010a) Improving enzymes for biomass conversion: a basic research perspective. *Bioenergy research*, 3(1): 82-92.
- Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Aslam N, Walton JD (2010b) Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. *Biotechnology and bioengineering*, 106(5): 707-720.

- Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Bongers M, Walton JD (2010c) Synthetic multi-component enzyme mixtures for deconstruction of lignocellulosic biomass. *Bioresource technology*, 101(23): 9097-9105.
- Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Walton JD (2010d) Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnology for biofuels*, 3(1): 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(1): 16.
- Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual review of microbiology*, 58: 521-554.
- Beckham GT, Bomble YJ, Matthews JF, Taylor CB, Resch MG, Yarbrough JM et al. (2010) The O-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophysical journal*, 99(11): 3773-3781.
- Beckner M, Ivey ML, Phister TG (2011) Microbial contamination of fuel ethanol fermentations. *Letters in applied microbiology*, 53(4), 387-394.
- Beeson WT, Phillips CM, Cate JHD, Marletta MA (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *Journal of the american chemical society*, 134(2): 890-892.
- Beeson WT, Vu VV, Span, E A, Phillips CM, Marletta, MA (2015) Cellulose degradation by polysaccharide monooxygenases. *Annual review of biochemistry*, 84: 923-946.
- Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S et al. (2015) Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnology for biofuels*, 8: 90-90.
- Berlin A, Balakshin M, Gilkes N, Kadla J, Maximenko V, Kubo S, Saddler J (2006) Inhibition of cellulase, xylanase and β -glucosidase activities by softwood lignin preparations. *Journal of biotechnology*, 125(2): 198-209.
- Bey M, Zhou SM, Poidevin L, Henrissat B, Coutinho PM, Berrin JG et al. (2013) Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospora anserina*. *Applied and environmental microbiology*, 79(2): 488-496.
- Bischof RH, Ramoni J, Seiboth B (2016) Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microbial cell factories*, 15(1): 106.

- Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*, 4(6): 1633-1649.
- Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annual review of plant biology*, 54(1): 519-546
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical journal*, 382(3), 769-781.
- Bozell JJ, Petersen GR (2010) Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy’s “Top 10” revisited. *Green chemistry*, 12(4): 539-554.
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, 72(1-2): 248-254.
- Buchanan BB, Gruissem W, Jones RL (2000) *Biochemistry & molecular biology of plants*. Rockville, MD: American Society of Plant Physiologists.
- Campbell MM, Sederoff RR (1996) Variation in lignin content and composition - mechanisms of control and implications for the genetic improvement of plants. *Plant physiology*, 110(1): 3.
- Chang VS, Holtzapple MT (2000) Fundamental factors affecting biomass enzymatic reactivity. *Applied biochemistry and biotechnology*, 84(1): 5-37.
- Chapman L (2007) Transport and climate change: a review. *Journal of transport geography*, 15(5): 354-367.
- Chen CR, Makhatadze GI (2015) Protein Volume: calculating molecular van der Waals and void volumes in proteins. *BMC bioinformatics*, 16(1): 101.
- Chen H, Hayn M, Esterbauer H (1992) Purification and characterization of two extracellular β -glucosidases from *Trichoderma reesei*. *Biochimica et biophysica acta - protein structure and molecular enzymology*, 1121(1-2): 54-60.
- Chen HZ (2005) *Biotechnology of lignocellulose*. New York, NY: Springer.
- Cherubini F (2010) The biorefinery concept: using biomass instead of oil for producing energy and chemicals. *Energy conversion and management*, 51(7): 1412-1421.
- Chniti S, Djelal H, Hassouna M, Amrane A (2014) Residue of dates from the food industry as a new cheap feedstock for ethanol production. *Biomass and bioenergy*, 6: 66-70.

Chundawat SP, Vismeh R, Sharma LN, Humpala JF, da Costa Sousa L, Chambliss CK et al. (2010) Multifaceted characterization of cell wall decomposition products formed during ammonia fiber expansion (AFEX) and dilute acid based pretreatments. *Bioresource technology*, 101(21): 8429-8438.

Chundawat SP, Beckham GT, Himmel ME, Dale BE (2011a) Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annual review of chemical and biomolecular engineering*, 2: 121-145.

Chundawat SP, Bellesia G, Uppugundla N, da Costa Sousa L, Gao D, Cheh AM et al. (2011b) Restructuring the crystalline cellulose hydrogen bond network enhances its depolymerization rate. *Journal of the american chemical society*, 133(29): 11163-11174.

Converse AO, Ooshima H, Burns DS (1990) Kinetics of enzymatic hydrolysis of lignocellulosic materials based on surface area of cellulose accessible to enzyme and enzyme adsorption on lignin and cellulose. *Applied biochemistry and biotechnology*, 24(1): 67-73.

Couturier M, Haon M, Coutinho PM, Henrissat B, Lesage-Meessen L et al. (2011) *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. *Applied and environmental microbiology*, 77(1): 237-246.

Davies G, Henrissat B (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, 3(9): 853-859.

de Vries RP, Visser J (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiology and molecular biology reviews*, 65(4): 497-522.

Demirbaş A (2001) Biomass resource facilities and biomass conversion processing for fuels and chemicals. *Energy conversion and management*, 42(11): 1357-1378.

Ding SJ, Ge W, Buswell JA (2001) Endoglucanase I from the edible straw mushroom, *Volvariella volvacea*. *FEBS journal*, 268(22): 5687-5695.

Ding SJ, Ge W, Buswell JA (2002) Secretion, purification and characterisation of a recombinant *Volvariella volvacea* endoglucanase expressed in the yeast *Pichia pastoris*. *Enzyme and microbial technology*, 31(5): 621-626.

Ding SY, Himmel ME (2006) The maize primary cell wall microfibril: a new model derived from direct visualization. *Journal of agricultural and food chemistry*, 54(3): 597-606.

Divne C, Sinning I, Ståhlberg J, Pettersson G, Bailey M, Siika-aho M et al. (1993) Crystallization and preliminary X-ray studies on the core proteins of cellobiohydrolase I and endoglucanase I from *Trichoderma reesei*. *Journal of molecular biology*, 234(3): 905-907.

Divne C, Ståhlberg J, Teeri TT, Jones TA (1998). High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *Journal of molecular biology*, 275(2), 309-325.

Draude KM, Kurniawan CB, Duff SJB (2001) Effect of oxygen delignification on the rate and extent of enzymatic hydrolysis of lignocellulosic material. *Bioresource technology*, 79(2): 113-120.

Duan CJ, Feng JX (2010) Mining metagenomes for novel cellulase genes. *Biotechnology letters*, 32(12): 1765-1775.

Duchesne LC, Larson DW (1989). Cellulose and the evolution of plant life. *Bioscience*, 39(4): 238-241.

Dufrêne YF, Ando T, Garcia R, Alsteens D, Martinez-Martin D, Engel A et al. (2017) Imaging modes of atomic force microscopy for application in molecular and cell biology. *Nature nanotechnology*, 12(4): 295-307.

Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, Porcel BM et al. (2008) The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome biology*, 9(5): R77.

Fan LT, Lee YH, Beardmore DR (1981) The influence of major structural features of cellulose on rate of enzymatic hydrolysis. *Biotechnology and bioengineering*, 23(2): 419-424.

Feely RA, Doney SC, Cooley SR (2009) Ocean acidification: Present conditions and future changes in a high-CO₂ world. *Oceanography*, 22(4): 36-47.

Ferrè F, Clote P (2006) DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic acids research*, 34(suppl_2): W182-W185.

Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein science*, 20(9): 1479-1483.

Ford predicts fuel from vegetation (1925, September 20). *The New York Times*. Retrieved from <http://query.nytimes.com/mem/archive-free/pdf?res=9C00E4DB153FE631A25753C2A96F9C946495D6CF>

Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SWA, Vincken JP et al. (2016) Lytic polysaccharide monoxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnology for biofuels*, 9(1): 186.

Fukumori F, Kudo T, Horikoshi K (1985) Purification and properties of a cellulase from alkalophilic *Bacillus sp.* No. 1139. *Microbiology*, 131(12): 3339-3345.

Gibson LJ (2012) The hierarchical structure and mechanics of plant materials. *Journal of the royal society interface*, 9: 2749-2766. DOI: 10.1098/rsif.2012.0341.

- Gomez LD, Steele-King CG, McQueen-Mason SJ (2008) Sustainable liquid biofuels from biomass: the writing's on the walls. *New phytologist*, 178(3): 473-485.
- Grabber JH (2005) How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop science*, 45(3): 820-831.
- Gray KA, Zhao L, Emptage M (2006) Bioethanol. *Current opinion in chemical biology*, 10(2): 141-146.
- Grethlein HE (1985) The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates. *Nature biotechnology*, 3(2): 155-160.
- Gruber F, Visser J, Kubicek CP, de Graaff LH (1990) The development of a heterologous transformation system for the cellulolytic fungus *Trichoderma reesei* based on a pyrG-negative mutant strain. *Current genetics*, 18(1): 71-76.
- Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC et al. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*, 49(15): 3305-3316.
- Hatfield RD, Ralph J, Grabber JH (1999) Cell wall cross-linking by ferulates and diferulates in grasses. *Science of food and agriculture*, 79: 403-407.
- Hemsworth GR, Johnston EM, Davies GJ, Walton PH (2015) Lytic polysaccharide monoxygenases in biomass conversion. *Trends in biotechnology*, 33(12): 747-761.
- Henriksson H, Ståhlberg J, Isaksson R, Pettersson G (1996) The active sites of cellulases are involved in chiral recognition: a comparison of cellobiohydrolase 1 and endoglucanase 1. *FEBS letters*, 390(3): 339-344.
- Henrissat B, Driguez H, Viet C, Schülein M (1985) Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. *Nature biotechnology*, 3(8): 722-726.
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical journal*, 280(2): 309-316.
- Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical journal*, 293(3): 781-788.
- Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Current opinion in structural biology*, 7(5): 637-644.
- Hestrin S, Schramm M (1954) Synthesis of cellulose by *Acetobacter xylinum*. 2. Preparation of freeze-dried cells capable of polymerizing glucose to cellulose. *Biochemical journal*, 58(2): 345.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315(5813): 804-807.

Hong J, Tamaki H, Akiba S, Yamamoto K, Kumagai H (2001). Cloning of a gene encoding a highly stable endo- β -1, 4-glucanase from *Aspergillus niger* and its expression in yeast. *Journal of bioscience and bioengineering*, 92(5): 434-441.

Horikoshi K, Nakao M, Kurono Y, Sashihara N (1984) Cellulases of an alkalophilic *Bacillus* strain isolated from soil. *Canadian journal of microbiology*, 30(6): 774-779.

Horikoshi K (1999) Alkaliphiles: some applications of their products for biotechnology. *Microbiology and molecular biology reviews*, 63(4): 735-750.

Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink V (2012). Novel enzymes for the degradation of cellulose. *Biotechnology for biofuels*, 5(1): 45.

Hu J, Arantes V, Pribowo A, Gourlay K, Saddler JN (2014) Substrate factors that influence the synergistic interaction of AA9 and cellulases during the enzymatic hydrolysis of biomass. *Energy and environmental science*, 7(7): 2308-2315.

Huang R, Su R, Qi W, He Z (2010) Understanding the key factors for enzymatic conversion of pretreated lignocellulose by partial least square analysis. *Biotechnology progress*, 26(2): 384-392.

Igarashi K, Koivula A, Wada M, Kimura S, Penttilä M, Samejima M (2009) High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *Journal of biological chemistry*, 284(52): 36186-36190.

Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T et al. (2011) Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science*, 333(6047): 1279-1282.

Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *Journal of biological chemistry*, 289(5): 2632-2642.

Ito S (1997) Alkaline cellulases from alkaliphilic *Bacillus*: enzymatic properties, genetics, and application to detergents. *Extremophiles*, 1(2): 61-66.

Jørgensen H, Kristensen JB, Felby C (2007) Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels, bioproducts and biorefining*, 1(2): 119-134.

Kar Y, Deveci H (2006) Importance of P-series fuels for flexible-fuel vehicles (FFVs) and alternative fuels. *Energy sources, Part A*, 28(10): 909-921.

- Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M (2008) The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *Journal of molecular biology*, 383(1): 144-154.
- Karkehabadi S, Helmich KE, Kaper T, Hansson H, Mikkelsen NE, Gudmundsson M et al. (2014) Biochemical characterization and crystal structures of a fungal family 3 β-glucosidase, Cel3A from *Hypocrea jecorina*. *Journal of biological chemistry*, 289(45): 31624-31637.
- Katsuraya K, Okuyama K, Hatanaka K, Oshima R, Sato T, Matsuzaki K (2003) Constitution of konjac glucomannan: chemical analysis and ¹³C NMR spectroscopy. *Carbohydrate polymers*, 53(2): 183-189.
- Kenealy WR, Jeffries TW (2003) Enzyme processes for pulp and paper: a review of recent developments. In: Goodell B, Nicholas DD, Schultz TP (eds.) *Wood deterioration and preservation: Advances in our changing world* (pp. 17-43). American Chemical Society.
- Kim IJ, Youn HJ, Kim, KH (2016) Synergism of an auxiliary activity 9 (AA9) from *Chaetomium globosum* with xylanase on the hydrolysis of xylan and lignocellulose. *Process biochemistry*, 51(10): 1445-1451.
- Kim S, Dale BE (2004) Global potential bioethanol production from wasted crops and crop residues. *Biomass and bioenergy*, 26(4): 361-375.
- Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW (2012). The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnology and bioengineering*, 109(4): 1083-1087.
- Klemm D, Heublein B, Fink HP, Bohn A (2005) Cellulose: fascinating biopolymer and sustainable raw material. *Angewandte chemie international edition*, 44(22): 3358-3393.
- Kluepfel D, Shareck F, Mondou F, Morosoli R (1986) Characterization of cellulase and xylanase activities of *Streptomyces lividans*. *Applied microbiology and biotechnology*, 24(3): 230-234.
- Koullas D, Christakopoulos P, Kekos D, Macris BJ, Koukios E (1990). Effect of cellulose crystallinity on the enzymic hydrolysis of lignocellulosics by *Fusarium oxysporum* cellulases. *Cellulose chemistry and technology*, 24: 469-474.
- Kracher, D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K et al. (2016) Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science*, 352(6289): 1098-1101.
- Kraulis PJ, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J et al. (1989) Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry*, 28(18): 7241-7257.

Krystynowicz A, Czaja W, Wiktorowska-Jeziarska A, Gonçalves-Miśkiewicz M, Turkiewicz M, Bielecki S (2002) Factors affecting the yield and properties of bacterial cellulose. *Journal of industrial microbiology and biotechnology*, 29(4): 189-195.

Kumar CS, Bhattacharya S (2008) Tamarind seed: properties, processing and utilization. *Critical reviews in food science and nutrition*, 48(1): 1-20.

Kumar P, Barrett DM, Delwiche MJ, Stroeve P (2009) Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial and engineering chemistry research*, 48(8): 3713-3729.

Kumar R, Dahiya JS, Singh D, Nigam P (2000) Production of endo-1, 4- β -glucanase by a biocontrol fungus *Cladorrhinum foecundissimum*. *Bioresource technology*, 75(1): 95-97.

Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of industrial microbiology & biotechnology*, 35(5): 377-391.

Kumar R, Mago G, Balan V, Wyman CE (2009) Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. *Bioresource technology*, 100(17): 3948-3962.

Kumar R, Wyman CE (2014) Strong cellulase inhibition by mannan polysaccharides in cellulose conversion to sugars. *Biotechnology and bioengineering*, 111(7): 1341-1353.

Lamed R, Bayer EA (1988) The cellulosome of *Clostridium thermocellum*. *Advances in applied microbiology*, 33: 1-46.

Lee TM, Farrow MF, Arnold FH, Mayo SL (2011) A structural study of *Hypocrea jecorina* Cel5A. *Protein science*, 20(11): 1935-1940.

Li J, Tang C, Shi H, Wu M (2011) Cloning and optimized expression of a neutral endoglucanase gene (*ncel5A*) from *Volvariella volvacea* WX32 in *Pichia pastoris*. *Journal of bioscience and bioengineering*, 111(5): 537-540.

Li X, Beeson WT, Phillips CM, Marletta MA, Cate JHD (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure*, 20(6): 1051-1061.

Liao W, Wen Z, Hurley S, Liu Y, Liu C, Chen S (2005) Effects of hemicellulose and lignin on enzymatic hydrolysis of cellulose from dairy manure. In: Davison BH, Evans BR, Finkelstein M, McMillan JD (eds.) *Twenty-Sixth Symposium on Biotechnology for Fuels and Chemicals* (pp. 1017-1030). Humana Press.

Lieth H (1975) Primary production of the major vegetation units of the world. In: Lieth H, Whittaker RH (eds.) *Primary productivity of the biosphere* (pp. 203-215). Springer, Berlin-Heidelberg.

- Lin H, Li W, Guo C, Qu S, Ren N (2011) Advances in the study of directed evolution for cellulases. *Frontiers of environmental science and engineering in China*, 5(4): 519-525.
- Linder M, Teeri TT (1996) The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. *Proceedings of the national academy of sciences*, 93(22): 12251-12255.
- Liu YS, Baker JO, Zeng Y, Himmel ME, Haas T, Ding SY (2011) Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. *Journal of biological chemistry*, 286(13): 11195-11201.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2013) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research*, 42(D1): D490-D495.
- Lou H, Zhu JY, Lan TQ, Lai H, Qiu X (2013) pH-Induced lignin surface modification to reduce nonspecific cellulase binding and enhance enzymatic saccharification of lignocelluloses. *ChemSusChem*, 6(5): 919-927.
- Lynd LR, Weimer PJ, Van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and molecular biology reviews*, 66(3): 506-577.
- Mandels M, Reese ET (1957) Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. *Journal of bacteriology*, 73(2): 269.
- Mandels M, Weber J (1969) The production of cellulases. In: *Cellulases and Their Applications*. In: Hajny GJ & ET Reese (eds.), American Chemical Society, Washington DC, 391-414.
- Mansfield SD, Mooney C, Saddler JN (1999) Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnology progress*, 15(5): 804-816.
- Martínez-Sanz M, Lopez-Sanchez P, Gidley MJ, Gilbert EP (2015) Evidence for differential interaction mechanism of plant cell wall matrix polysaccharides in hierarchically-structured bacterial cellulose. *Cellulose*, 22(3): 1541-1563.
- Martínez-Sanz M, Gidley MJ, Gilbert EP (2016) Hierarchical architecture of bacterial cellulose and composite plant cell wall polysaccharide hydrogels using small angle neutron scattering. *Soft matter*, 12(5): 1534-1549.
- McKenna BA, Mikkelsen D, Wehr JB, Gidley MJ, Menzies NW (2009) Mechanical and structural properties of native and alkali-treated bacterial cellulose produced by *Gluconacetobacter xylinus* strain ATCC 53524. *Cellulose*, 16(6): 1047.
- Merino ST, Cherry J (2007). Progress and challenges in enzyme development for biomass utilization. In: Olsson L (ed.) *Biofuels. Advances in biochemical engineering/biotechnology*, vol. 108, pp. 95-120. Springer Berlin Heidelberg.

- Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical chemistry*, 31(3): 426-428.
- Mongay C, Cerda V (1974) VI/A Britton-Robinson buffer of known ionic strength. *Annali di Chimica*, 64: 409-412.
- Moon RJ, Martini A, Nairn J, Simonsen J, Youngblood J (2011) Cellulose nanomaterials review: structure, properties and nanocomposites. *Chemical society reviews*, 40(7): 3941-3994.
- Mooney CA, Mansfield SD, Touhy MG, Saddler JN (1998) The effect of initial pore volume and lignin content on the enzymatic hydrolysis of softwoods. *Bioresource technology*, 64(2): 113-119.
- Nagendran S, Hallen-Adams HE, Paper JM, Aslam N, Walton JD (2009) Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus *Amanita bisporigera*, based on the secretome of *Trichoderma reesei*. *Fungal genetics and biology*, 46(5): 427-435.
- Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic acids research*, 28(1): 292-292.
- Nazhad MM, Ramos LP, Paszner L, Saddler JN (1995) Structural constraints affecting the initial enzymatic hydrolysis of recycled paper. *Enzyme and microbial technology*, 17(1): 68-74.
- Nigam PS, Singh A (2011) Production of liquid biofuels from renewable resources. *Progress in energy and combustion science*, 37(1): 52-68.
- Nimlos MR, Beckham GT, Matthews JF, Bu L, Himmel ME, Crowley MF (2012) Binding preferences, surface attachment, diffusivity, and orientation of a family 1 carbohydrate-binding module on cellulose. *Journal of biological chemistry*, 287(24): 20603-20612.
- Niven RK (2005). Ethanol in gasoline: environmental impacts and sustainability review article. *Renewable and sustainable energy reviews*, 9(6): 535-555.
- Nogawa M, Goto M, Okada H, Morikawa Y (2001) L-Sorbose induces cellulase gene transcription in the cellulolytic fungus *Trichoderma reesei*. *Current genetics*, 38(6): 329-334.
- Ooshima H, Burns DS, Converse AO (1990) Adsorption of cellulase from *Trichoderma reesei* on cellulose and lignocellulosic residue in wood pretreated by dilute sulfuric acid with explosive decompression. *Biotechnology and bioengineering*, 36(5): 446-452.
- Palonen H, Tjerneld F, Zacchi G, Tenkanen M (2004). Adsorption of *Trichoderma reesei* CBH1 and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *Journal of biotechnology*, 107(1): 65-72.

Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant journal*, 54(4): 559-568.

Peterson R, Nevalainen H (2012). *Trichoderma reesei* RUT-C30—thirty years of strain improvement. *Microbiology*, 158(1): 58-68.

Phillips CM, Beeson IV WT, Cate JH, Marletta MA (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS chemical biology*, 6(12): 1399-1406.

Pingali SV, Urban VS, Heller WT, McGaughey J, O'Neill H, Foston M et al. (2010). Breakdown of cell wall nanostructure in dilute acid pretreated biomass. *Biomacromolecules*, 11(9): 2329-2335.

Poidevin L, Feliu J, Doan A, Berrin JG, Bey M, Coutinho PM et al. (2013) Insights into exo- and endoglucanase activities of family 6 glycoside hydrolases from *Podospora anserina*. *Applied and environmental microbiology*, 79(14): 4220-4229.

Poidevin L, Berrin JG, Bennati-Granier C, Levasseur A, Herpoël-Gimbert I, Chevret D et al. (2014). Comparative analyses of *Podospora anserina* secretomes reveal a large array of lignocellulose-active enzymes. *Applied microbiology and biotechnology*, 98(17): 7457-7469.

Popper ZA, Fry SC (2003) Primary cell wall composition of bryophytes and charophytes. *Annals of botany*, 91(1): 1-12.

Popper ZA (2008) Evolution and diversity of green plant cell walls. *Current opinion in plant biology*, 11(3): 286-292.

Qing Q, Yang B, Wyman CE (2010) Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes. *Bioresource technology*, 101(24): 9624-9630.

Qing Q, Wyman CE (2011). Supplementation with xylanase and β -xylosidase to reduce xylo-oligomer and xylan inhibition of enzymatic hydrolysis of cellulose and pretreated corn stover. *Biotechnology for biofuels*, 4(1): 18.

Quinlan RJ, Sweeney MD, Leggio LL, Otten H, Poulsen JCN, Johansen KS et al. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the national academy of sciences*, 108(37): 15079-15084.

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA et al. (2006). The path forward for biofuels and biomaterials. *Science*, 311(5760): 484-489.

Rahikainen JL, Mikander S, Marjamaa K, Tamminen T, Lappas A, Viikari L, Kruus K (2011) Inhibition of enzymatic hydrolysis by residual lignins from softwood—study of enzyme binding and inactivation on lignin-rich surface. *Biotechnology and bioengineering*, 108(12): 2823-2834.

- Rahikainen JL, Martin-Sampedro R, Heikkinen H, Rovio S, Marjamaa K et al. (2013a) Inhibitory effect of lignin during cellulose bioconversion: the effect of lignin chemistry on non-productive enzyme adsorption. *Bioresource technology*, 133: 270-278.
- Rahikainen JL, Evans JD, Mikander S, Kalliola A, Puranen T, Tamminen T et al. (2013b). Cellulase–lignin interactions — The role of carbohydrate-binding module and pH in non-productive binding. *Enzyme and microbial technology*, 53(5): 315-321.
- Rouvinen J, Bergfors T, Teeri T, Knowles JKC, Jones TA (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science*, 249(4967): 380-386.
- Rowland SP (1977) Cellulose: pores, internal surfaces, and the water interface. In: Arthur Jr JC (eds.) *Textile and Paper Chemistry and Technology* (pp. 20-45). American Chemical Society.
- Saloheimo M, Lehtovaara P, Penttilä M, Teeri TT, Stahlberg J, Johansson G et al. (1988) EGIII, a new endoglucanase from *Trichoderma reesei*: the characterization of both gene and enzyme. *Gene*, 63(1): 11-21.
- Sarkar N, Ghosh SK, Bannerjee S, Aikat K (2012). Bioethanol production from agricultural wastes: an overview. *Renewable energy*, 37(1): 19-27.
- Scheller HV, Ulvskov P (2010) Hemicelluloses. *Annual review of plant biology*, 61: 263-289.
- Schülein M (1988) Cellulases of *Trichoderma reesei*. *Methods in enzymology*, 160: 234-242.
- Seidl V, Seiboth B (2010). *Trichoderma reesei*: genetic approaches to improving strain efficiency. *Biofuels*, 1(2): 343-354.
- Shanmughapriya S, Kiran GS, Selvin J, Thomas TA, Rani C (2010) Optimization, purification, and characterization of extracellular mesophilic alkaline cellulase from sponge-associated *Marinobacter sp.* MSI032. *Applied biochemistry and biotechnology*, 162(3): 625-640.
- Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical chemistry*, 68(5): 850-858.
- Sojar HT, Bahl OP (1987) Chemical deglycosylation of glycoproteins. *Methods in enzymology*, 138: 341-350.
- Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J et al. (2004) Toward a systems approach to understanding plant cell walls. *Science*, 306(5705): 2206-2211.
- Sorrell S, Speirs J, Bentley R, Brandt A, Miller R (2010) Global oil depletion: A review of the evidence. *Energy policy*, 38(9): 5290-5295.

Srisodsuk M, Reinikainen T, Penttilä M, Teeri TT (1993). Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *Journal of biological chemistry*, 268(28): 20756-20761.

Stricker AR, Grosstessner-Hain K, Würleitner E, Mach RL (2006) Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. *Eukaryotic cell*, 5(12): 2128-2137.

Tambor JH, Ren H, Ushinsky S, Zheng Y, Riemens A, St-Francois C et al. (2012) Recombinant expression, activity screening and functional characterization identifies three novel endo-1,4- β -glucanases that efficiently hydrolyse cellulosic substrates. *Applied microbiology and biotechnology*, 93(1): 203-214.

Takashima S, Nakamura A, Masaki H, Uozumi T (1997) Cloning, sequencing, and expression of a thermostable cellulase gene of *Humicola grisea*. *Bioscience, biotechnology, and biochemistry*, 61(2): 245-250.

Teeri TT (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends in biotechnology*, 15(5): 160-167.

Uzbas F, Sezerman U, Hartl L, Kubicek CP, Seiboth B (2012) A homologous production system for *Trichoderma reesei* secreted proteins in a cellulase-free background. *Applied microbiology and biotechnology*, 93(4): 1601-1608.

Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørli M et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, 330(6001): 219-222.

Van Dyk JS, Pletschke BI (2012) A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes—factors affecting enzymes, conversion and synergy. *Biotechnology advances*, 30(6): 1458-1480.

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin biosynthesis and structure. *Plant physiology*, 153(3): 895-905.

Villares A, Moreau C, Bennati-Granier C, Garajova S, Foucat L, Falourd X et al. (2017) Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. *Scientific reports*, 7: 40262.

Vohra M, Manwar J, Manmode R, Padgilwar S, Patil S (2014) Bioethanol production: feedstock and current technologies. *Journal of environmental chemical engineering*, 2(1): 573-584.

Vu VV, Beeson WT, Phillips CM, Cate JHD, Marletta MA (2013) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *Journal of the american chemical society*, 136(2): 562-565.

- Walton J, Banerjee G, Car S (2011) GENPLAT: an automated platform for biomass enzyme discovery and cocktail optimization. *Journal of visualized experiments*, pii: 3314. doi: 10.3791/3314.
- Wang M, Du J, Zhang D, Li X, Zhao J (2017) Modification of different pulps by homologous overexpression alkali-tolerant endoglucanase in *Bacillus subtilis* Y106. *Scientific reports*, 7: 3321. doi: 10.1038/s41598-017-03215-9.
- Werpy T, Petersen G, Aden A, Bozell J, Holladay J, White J et al. (2004) Top value added chemicals from biomass. Volume 1-Results of screening for potential candidates from sugars and synthesis gas (No. DOE/GO-102004-1992). Department of Energy Washington DC. Retrieved from <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA436528>
- Westereng B, Ishida T, Vaaje-Kolstad G, Wu M, Eijsink VG, Igarashi K et al. (2011) The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One*, 6(11): e27807.
- Westereng B, Cannella D, Agger JW, Jørgensen H, Andersen ML, Eijsink VG et al. (2015) Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Scientific reports*, 5: 18561.
- White AR, Brown RM (1981) Enzymatic hydrolysis of cellulose: visual characterization of the process. *Proceedings of the national academy of sciences*, 78(2): 1047-1051.
- Wilkerson CG, Mansfield SD, Lu F, Withers S, Park JY, Karlen SD et al. (2014) Monoglucuronol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science*, 344(6179): 90-93.
- Wilkie KCB (1979) The hemicelluloses of grasses and cereals. *Advances in carbohydrate chemistry and biochemistry*, 36: 215-264.
- Wilson DB (2009) Cellulases and biofuels. *Current opinion in biotechnology*, 20(3): 295-299.
- Woodward J, Wiseman A (1982) Fungal and other β -D-glucosidases—their properties and applications. *Enzyme and microbial technology*, 4(2): 73-79.
- Wu TH, Huang CH, Ko TP, Lai HL, Ma Y, Chen CC et al. (2011) Diverse substrate recognition mechanism revealed by *Thermotoga maritima* Cel5A structures in complex with cellotetraose, cellobiose and mannotriose. *Biochimica et biophysica acta (BBA)-proteins and proteomics*, 1814(12): 1832-1840.
- Wyman CE, Dale BE, Elander RT, Holtzapfle M, Ladisch MR, Lee YY (2005) Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover. *Bioresource technology*, 96(18): 2026-2032.

- Xiao Z, Storms R, Tsang A (2004) Microplate-based filter paper assay to measure total cellulase activity. *Biotechnology and bioengineering*, 88(7): 832-837.
- Xin D, Ge X, Sun Z, Viikari L, Zhang J (2014) Competitive inhibition of cellobiohydrolase I by manno-oligosaccharides. *Enzyme and microbial technology*, 68: 62-68.
- Xu Q, Singh A, Himmel ME (2009) Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose. *Current opinion in biotechnology*, 20(3): 364-371.
- Yang B, Willies DM, Wyman CE (2006) Changes in the enzymatic hydrolysis rate of Avicel cellulose with conversion. *Biotechnology and bioengineering*, 94(6): 1122-1128.
- Ye Z, Zheng Y, Li B, Borrusch MS, Storms R, Walton JD (2014) Enhancement of synthetic *Trichoderma*-based enzyme mixtures for biomass conversion with an alternative family 5 glycosyl hydrolase from *Sporotrichum thermophile*. *PloS One*, 9(10): e109885.
- Yeh AI, Huang YC, Chen SH (2010) Effect of particle size on the rate of enzymatic hydrolysis of cellulose. *Carbohydrate polymers*, 79(1): 192-199.
- Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic acids research*, 40(W1), W445-W451.
- Yoshida M, Liu Y, Uchida S, Kawarada K, Ukagami Y et al. (2008) Effects of cellulose crystallinity, hemicellulose, and lignin on the enzymatic hydrolysis of *Miscanthus sinensis* to monosaccharides. *Bioscience, biotechnology, and biochemistry*, 72(3): 805-810.
- Yuan SF, Wu TH, Lee HL, Hsieh HY, Lin WL, Yang B et al. (2015) Biochemical characterization and structural analysis of a bifunctional cellulase/xylanase from *Clostridium thermocellum*. *Journal of biological chemistry*, 290(9): 5739-5748.
- Zabed H, Sahu JN, Suely A, Boyce AN, Faruq G (2017) Bioethanol production from renewable sources: Current perspectives and technological progress. *Renewable and sustainable energy reviews*, 71: 475-501.
- Zhang J, Tang M, Viikari L (2012) Xylans inhibit enzymatic hydrolysis of lignocellulosic materials by cellulases. *Bioresource technology*, 121: 8-12.
- Zhang YHP, Lynd LR (2004) Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnology and bioengineering*, 88(7): 797-824.
- Zhao X, Zhang L, Liu D (2012) Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels, bioproducts and biorefining*, 6(4): 465-482.

Zheng F, Ding S (2013) Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from *Volvariella volvacea*. *Applied and environmental microbiology*, 79(3): 989-996.

Zheng Y, Lee C, Yu C, Cheng YS, Simmons CW, Zhang R et al. (2012) Ensilage and bioconversion of grape pomace into fuel ethanol. *Journal of agricultural and food chemistry*, 60(44): 11128-11134.