Enzyme cocktail development for the conversion of pretreated wood biomass

Sylvanus Enongene Ekwe

A Thesis
In the Department
of
Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Biochemistry) at
Concordia University
Montreal, Quebec, Canada

March 2014

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Sylvanus Enongene Ekwe

Entitled: Enzyme cocktail development for the conversion of pretreated wood biomass

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

Doctor of Philosophy (Chemistry)

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

Signed by the final examining committee:

Dr. K. Li — Chair

Dr. M. Beauregard — External Examiner

Dr. P. Wood-Adams — External to Program

Dr. P. Pawelek — Examiner

Dr. R. Storms — Examiner

Dr. J. Powlowski — Thesis Supervisor

Approved by

Chair of Department or Graduate Program Director

Dean of Faculty

Abstract

Enzyme cocktail development for the conversion of pretreated wood biomass

Sylvanus Enongene Ekwe, Ph.D.
Concordia University, 2014

A steady rise in global consumption of fossil-based energy has led to a surge in prices of petroleum-derived fuels, chemicals and materials in recent years. The abundantly available lignocellulosic biomass is a renewable, potential substitute for petroleum that can be biotransformed into biofuels, chemicals and materials. Commercial lignocellulose-degrading enzyme cocktails used in this biotransformation still require improvement. In this thesis, I report the biochemical characterization of 34 plant cell wall degrading proteins from *Aspergillus niger* expressed in one or more of four host systems (*Aspergillus niger, Pichia pastoris, Escherichia coli* and *Nicotiana benthamiana*), in order to identify industrially relevant differences in their catalytic function. The results show that *N. benthamiana* is as effective a production host as *A. niger* itself and the recombinant proteins produced in the four host systems show similar biochemical properties. I also report the development of a medium- to high-throughput-adaptable screening method for evaluating the hydrolytic capability of lignocellulolytic enzymes at biomass loadings greater than 15% dry w/v. The results show that the method is suitable for the screening of cell wall degrading enzymes with superior properties and for evaluating biomass hydrolysability. I also report the identification of biomass-liquefying enzymes from the secretome of thermophilic saprotroph, *Myceliophthora thermophila* grown on various pretreated wood biomass types. Results reveal an arsenal of 47 secreted proteins which concertedly liquefy pretreated wood biomass at 15% dry solids and boost glucose release by a commercial cellulase system. A highly-expressed GH7 cellobiohydrolase, *Mt*CBH7 was found to liquefy 15% black spruce kraft pulp, acting optimally at 55 °C. The overall finding are: (i) *N. benthamiana* holds potential as a production host for eukaryotic biomass-degrading enzymes; (ii) a screening method for evaluating the hydrolytic potential of enzymes on natural biomass has been developed; and (iii) *Mt*CBH7 is a promising enzyme candidate for application in biorefineries where enhanced liquefaction of lignocellulose is required.

Acknowledgements

I express special appreciation to my supervisor Dr Justin Powlowski, for giving me the opportunity to be trained in your lab, and for supporting and guiding me through this challenging path. Your insightful discussions and suggestions about my research and the freedom you have given me to pursue various parts of the projects as I see it have improved me as a scientist and person. Dr Adrian Tsang, I also thank you for providing a weekly critical analysis of my results and for persistently reminding me to get things right first time. To members of my supervisory committee, Dr Reginald Storms and Dr Peter Pawelek, thank you for your constructive critical analysis of my research and for supporting me throughout.

Thank you all members of the Powlowski Lab. Dr Lena Sahlman, thank you for your ideas to improve on my work. To my special colleagues Racha Sheikh-Ibrahim, Mark-Anthony Timeliadis, Nathalie Khor and co., your fun-loving nature and care for one another in the team made my work feel light. Thank you for being the most lively, enthusiastic and energetic people to work with.

To my colleagues of the Tsang Lab at the Centre of Structural and Functional Genomics, you have been amazing. You made my walk to the finishing line feel easy. To all my colleagues who supported me in writing, thank you.

Words cannot express how grateful I am to my parents for all you have sacrificed for my sake. Your persistent prayers have sustained me throughout. To all my siblings and extended family, especially, Sixtus Ekane Masango, thank you for your relentless support.

I express special appreciation to my wife and companion, Juliette Ntube Ekwe for being very supportive and for sharing my burden during the difficult moments. To my daughters, Yoella Nzelle Enongene Ekwe and Nsoli Enongene Ekwe, your coming has brought and sustained a steady beam of happiness in my life. You both are my inspiration.

Ultimately, my life and all I do are dedicated to God almighty, the Cause of all causes, Creator of that which is seen and unseen, the Infinite and Infinitesimal, the Beginning and the End. Through Him, I can do all things (Philippians 4:13).

Authors' Contribution

Enongene Ekwe characterized the biochemical properties of the enzymes produced in *P. pastoris* and *E. coli* in Manuscripts I and II, conducted all experiments in Manuscripts III and IV (except mass-spectrometry), reviewed Manuscript I and prepared the Manuscripts II – V. Xiao Zhang characterized the biochemical properties of the enzymes produced in *A. niger* and *N. benthamiana* and reviewed Manuscript I. Sarah Zerbs amplified the target genes, cloned and transformed them into *E. coli* and *P. pastoris*. Andrew J. Conley constructed the genes in the plant expression vector, conducted the transformation in *N. benthamiana* and generated the plant extracts. Angelo Kaldis, Frank Collart, Jon Magnuson, Scott E. Baker and Adrian Tsang conceived and designed the initial study reported in Manuscripts I and II. Frank Collart, Jon Magnuson and Scott E. Baker oversaw the cloning and expression of the target genes in *E. coli* and *P. pastoris*. Rima Menassa oversaw the transformations and characterization in *N. benthamiana*. Adrian Tsang supervised the transformation in *A. niger,* wrote Manuscript I and reviewed Manuscript V. Manuscript I was reviewed by all listed authors therein and critically edited by Andrew J. Conley, Frank Collart, Justin Powlowski, Rima Menassa, Scott E. Baker and Sarah Zerbs. Marcos Di Falco analysed the proteins by mass spectrometry and reviewed Manuscript IV. Adrian Tsang and Justin Powlowski conceived, designed and supervised the studies in Manuscripts III and IV. Ingo Morgenstern and Reginald Storms reviewed and critically edited Manuscript V. Justin Powlowski supervised the thesis and reviewed and critically edited all manuscripts.

Table of contents

List of Manuscripts and Papers

I) Zhang X, Ekwe E, Conley AJ, Kaldis A, Zerbs S, Collart F, Magnuson J, Baker SE, Menassa R, Powlowski J, Tsang A, 2014. A comprehensive comparison of fungal glycoside hydrolases produced in bacterial, yeast, filamentous fungal and plant expression hosts. *Manuscript*

II) Ekwe E, Tsang A, Powlowski J, 2014. Cloning, Expression and characterization of *Aspergillus niger* glycoside hydrolases. *Manuscript*

III) Ekwe E, Tsang A, Powlowski J, 2014. A microplate-based method for screening fungal enzymes on high consistency biomass. *Manuscript*

IV) Ekwe E, Di Falco M, Tsang A, Powlowski J, 2014. *Myceliophthora thermophila* as a source of proteins to enhance liquefaction of high consistency pretreated wood biomass. *Manuscript*

V) Ekwe E, Morgenstern I, Tsang A, Storms R, Powlowski J, 2013. Non-hydrolytic cellulose active proteins: research progress and potential application in biorefineries. *Ind biotechnol* 9(3): 123-131

List of Figures

List of Tables

List of Appendices

Abbreviations and symbols

°C: degree Celsius
Abf: arabinofuranosidase
Abn: arabinanase
Ace: acetylxylanesterase
AFEX: ammonia fibre expansion
Agl: α-glucosidase
Agu: α-glucuronidase
AKP: aspen kraft pulp
AMP: aspen mechanically pretreated pulp
Ani: *Aspergillus niger*
Axh: arabinoxylan arabinofuranohydrolase
BC: Before Christ
BGL: β-glucosidase
BSKP: black spruce kraft pulp
C1: carbon 1
C4: carbon 4
C6: carbon 6
CAZy database: carbohydrate active enzyme database
CAZymes: carbohydrate active enzymes
CBH: cellobiohydrolase
CBHI: reducing end active cellobiohydrolase
CBHII: non-reducing end active cellobiohydrolase
CBM33: carbohydrate binding module 33
CDH: cellobiose dehydrogenase
CIP: cellulose induced protein
Cip: cellulosome scaffolding protein
DM: dry matter
DP: degree of polymerization
Eco: *Escheriachia coli*
EGL: endoglucanase
Exg: exoglucanase
EXLA: expansin-like family A
EXLB: expansin-like family B
EXLX: expansin-like family X
EXPA: α-expansins
EXPB: β-expansins
Fae: feruloyl esterase
GH: glycoside hydrolase
GHG: greenhouse gas

GRAS: generally recognized as safe
H_2: hydrogen gas
Inu: endo-inulinase
Inx: exo-inulinase
Lac: β-galactosidase
kDa: kilo Dalton
LIC: ligation independent cloning
Lic: licheninase
PMO: polysaccharide monoozygenase
Man: mannanase
Mel: α-galactosidase
Mnd: β-manosidase
Mog: mannosyl-oligosaccharide glucosidase
MPa: mega Pascal
MT: metric ton
Nbe: *Nicotiana benthamiana*
NSERC: Natural Sciences and Engineering Research Council
OH: hydroxyl group
OLP: organosolv pretreated lodgepole pine
OPP: organosolv pretreated poplar
PASC: phosphoric acid swollen cellulose
Pga: endo-polygalacturonase
Pgx: exo-polygalacturonase
pH: negative logarithm of hydrogen ion concentration
PNGaseF: peptide-N-glycosidase F
Ppa: *Pichia pastoris*
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPORL: Sulphur pretreatment to overcome recalcitrance of lignin
SPP: steam pretreated poplar
TSP: total soluble proteins
Ugl: unsaturated glucoronyl hydrolase
US$: United States dollar
USFDA: United States Food and Drug Administration
Xeg: xyloglucanase
Xyl: xylosidase
Xyn: xylanase

Chapter 1 Introduction

1.1 The petroleum problem

Modern society depends increasingly on petroleum-based fuels and materials. A steady rise in global consumer habits has enabled a 17-fold increase in global energy consumption in the past century [1]. At least 22% of global primary energy, which is predominantly fossil-derived, is used in the transportation sector [2]. In addition to liquid transportation fuels, petroleum oil is used for the production of a broad spectrum of other products. For instance, one barrel (42 gallons) of crude petroleum oil yields 19.4 gallons of liquid transportation gasoline; the rest (53%) is used for the production of plastics, a variety of household consumables, pharmaceuticals, cosmetics, construction materials, automobile and aircraft components, bulk and fine chemicals, raw materials for various industries, etc. Although petroleum-derived fuels and products are integral to human lifestyles in modern society, fossil fuel reserves are finite [3] and the costs of petroleum derived energy, liquid transportation fuels and related high value products have been on a steady rise in recent years [4]. Moreover, fossil fuel use is known to emit pollutants and green house gases (GHG) which contribute to climate change [4,5,6,7]. Consequently, alternatives to petroleum are being sought after to provide cleaner and more sustainable sources of energy and materials.

Among the available renewable alternatives to fossil derived energy are biofuels, hydrogen, wind energy, hydro energy, tidal energy, solar energy and nuclear fission, all of which, relative to petroleum, are considered more environmentally friendly [8]. Collectively, these energy sources can significantly reduce global dependence on fossil-derived energy, chemicals and materials, but none of them can completely replace fossil-derived energy on their own [9,10,11]. Moreover, not all of these alternative forms of energy can actually be used to produce liquid transportation fuels. Nonetheless, lignocellulosic biomass is the only realistic, sustainable substitute or

supplement of fossilized petroleum oil as a source of liquid transportation fuels, in addition to its potential to provide bulk or fine chemicals and materials [5].

1.2 Lignocellulosic Biomass

Lignocellulosic biomass refers to the non-food, carbohydrate polymers and lignin produced by terrestrial plants. As shown in Figure 1, lignocellulose is localised in plant cell walls and is mainly made up of crystalline or semi-crystalline cellulose, a meshwork of hemicellulosic polysaccharides and heterogeneous, polyaromatic lignin accounting for 40-50%, 25-35% and 15-20% dry mass, respectively [12,13]. Also present are small amounts of pectin, proteins, extractives and ash [14].

The cellulose, hemicellulose and lignin contents vary (Table 1) depending on the biomass type, source and plant age, stage of growth, lineage and its surrounding environmental conditions [15,16,17,18].

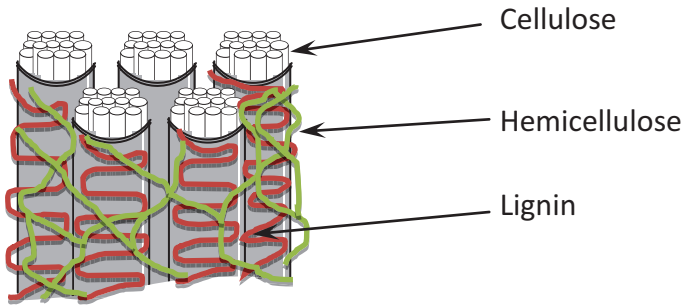


Figure 1. Cartoon representation of the structure of lignocellulosic biomass depicting cellulose, hemicellulose and lignin

Table 1 Cellulose, hemicellulose and lignin composition of selected lignocellulosic biomass (adapted from [19,14,16,20] and references cited therein)

Biomass type	% Biochemical composition		
	Cellulose	Hemicellulose	Lignin
Poplar (hard) wood	45 – 51	25 – 28	10 – 21
Hardwood stems	40 – 55	24 – 40	18 – 25
Eucalyptus	45 – 51	11 – 18	29
Softwood stems	45 – 50	24 – 40	18 – 25
Douglas fir	35 – 48	20 – 22	15 – 21
Waste papers from chemical pulps	60 – 70	10 – 20	5 – 10
Grasses	25 – 40	25 – 50	10 – 30
Corn stover	35 – 40	22 – 25	11 – 19
Corn cob	32 – 46	40	7 – 14
Rice straw	29 – 35	23 – 26	17 – 19
Barley straw	36 – 43	24 – 33	6 – 10
Wheat straw	35 – 39	22 – 30	12 – 16
Sugarcane bagasse	25 – 45	28 – 32	15 – 25
Sugarcane tops	35	32	14
Sorghum straw	32 – 35	24 – 27	15 – 21
Coffee pulp	34 – 37	44 – 48	16 – 19
Cotton stalk	31	11	30
Cotton	85 – 95	5 – 15	0
Leaves	15 – 20	80 – 85	0
Newspaper	40 – 55	24 – 39	18 – 30
Paper	85 – 99	0	0 – 15
Solid cattle manure	2 – 5	1 – 3	3 – 6

Lignocellulosic biomass is renewable and abundant and can be sourced from non-food forestry wastes and dedicated energy crops, agricultural waste, industrial waste, animal residue, sewage and municipal solid waste. Global annual production is estimated at 10^{10} MT, which makes it the most reliable resource for the sustainable production of renewable fuels and bio-based chemicals [21,22] that can potentially substitute a significant proportion of fossil fuels and petroleum based products.

Depending on the type of feedstock, liquid transportation biofuels are classified as either “first-generation”, which are derived from food-based feedstocks (e.g. cane sugar, sugar beet, sweet sorghum, corn, rice, wheat, potato, cassava, rape oil seed and oil palm) or “second-generation”, which are derived from non-food lignocellulosic biomass. Liquid transportation biofuels are currently produced at commercial scale

almost solely as first generation biofuels [23,24]. The fact that feedstocks for first generation biofuels are also used as food has raised ethical concerns against the production of first generation biofuels and chemicals. Concerns raised include the risk of a shift in land use away from food and feed production, which leads to shortages and price hikes of staple foods such as corn [24,25]. The impact is likely to weigh more on developing countries where food crops like corn are a major staple. Concerns on the depletion of soil nutrients have also been cited, which can potentially compromise food production in the future [26]. In addition production of first-generation biofuels have been linked to an increase in green house gas emission [27,28] and concerns have been raised about pesticide and water use, nutrient run-off, eutrophication and a wide range of other environmental problems [29,30,31,32,33,34].

Lignocellulosic material, however, does not compete with food [35], is abundant and as a source of energy and its use is generally thought to reduce GHG emission. For instance, cellulosic ethanol is thought to potentially offset greenhouse gas emission by up to 86% [36]. Lignocellulose therefore represents a huge potential resource for the sustainable production of liquid transportation fuels, commodity chemicals, materials and power in biorefineries [37].

1.2.1 Cellulose

Cellulose, a major product of plant photosynthesis, is the predominant structural component of plant cell walls and the most abundant biopolymer on earth [38]. Cellulose is a linear homopolymer of anhydroglucose subunits linked to each other by β-(1,4) glycosidic bonds (Figure 2a). The degree of polymerization (DP) of the anhydroglucose units varies widely: 100 – 20 000. For instance, the DP cellulose from wood pulp is in the range 500 – 1500, while that of cotton and bacterial cellulose is in the range 100 – 10,000. Cellulose with DP < 6 is generally soluble and DP > 12 renders cellulose insoluble [39,40,41].

Cellulose exists in two forms: crystalline and amorphous. In nature, individual cellulose linear chains interact with each other via hydrogen bonding and van der Waals forces to form tightly packed crystalline microfibrils (Figure 2b). This crystalline form constitutes the predominant form in which cellulose exists in nature. The less compact amorphous form represents the random, less-organised, non-crystalline regions within cellulose microfibrils [42], which constitutes a very small percentage of cellulose in nature. Unlike the compact crystalline cellulose that is more resistant to depolymerisation, amorphous cellulose is readily hydrolysable into the constituent glucose subunits by cellulase enzymes [43]. In plant cell walls, cellulose is wrapped in a sheath of hemicelluloses within a lignin matrix [38].

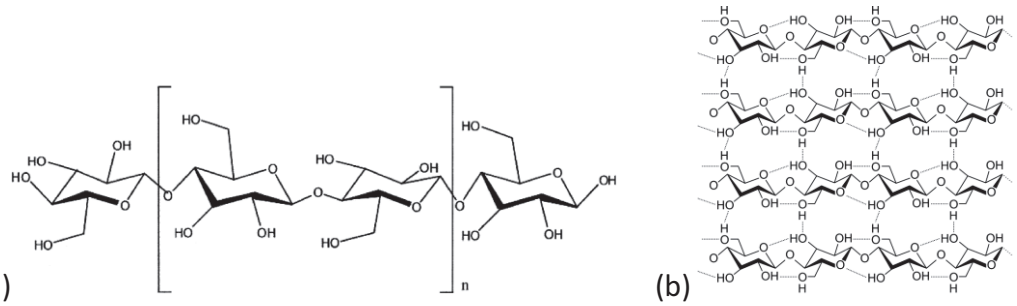


Figure 2. Structure of cellulose. (a) Linear chain and [40] (b) crystalline cellulose with multiple linear chains held tightly together by intra-molecular and inter-molecular hydrogen bonds and van der Waals forces. Images adapted from reference [44]. 'n' indicates the number of repeating unit that makes up a cellulose polymer

1.2.2 Hemicellulose

Hemicelluloses are branched carbohydrate homo- or hetero-polysaccharides, located in plant cell walls, which are built on equatorial β-(1-4)-linked backbone structures of pentoses (xylose, rhamnose, and arabinose), hexoses (glucose, mannose and galactose) and uronic acids (such as 4-0-methylglucuronic, D-glucuronic, and D-galactouronic acids) [45,46]. The DP range for hemicelluloses is 50 – 300, a much smaller value than for cellulose. The short lateral chain branches are usually made of different sugars and are predominantly linked to the polysaccharide backbone by β-(1,4)-glycosidic bonds, although a few β-(1,3)-glycosidic bonds may occur [45]. Acetylation of hemicellulose side chains is very common, especially with branched xylan. Some examples of hemicelluloses are shown in Figure 3. It is still unclear whether hemicelluloses are covalently linked to cellulose [47].

The main hemicelluloses present in cell walls of terrestrial plants include xyloglucans, xylans, mannans and glucomannans. The β-(1-3,1-4)-glucans are restricted to the Poales order of flowering plants that includes grasses and a few other groups. The inclusion of galactans, arabinans and arabinogalactans in the group of hemicelluloses is debatable since these polysaccharides appear to be part of pectin, at least in their initial synthesis, and do not share the equatorial β-(1-4)-linked backbone structure [46].

The lack of homogeneity in hemicelluloses does not permit the H-bonding, facilitated packing, crystallization or aggregation of polymers seen in cellulose. Consequently, hemicelluloses are essentially amorphous, with OH groups that tend to be more reactive to physical or chemical agents such as steam and dilute acid than they are in cellulose, rendering hemicelluloses readily hydrolysable [48].

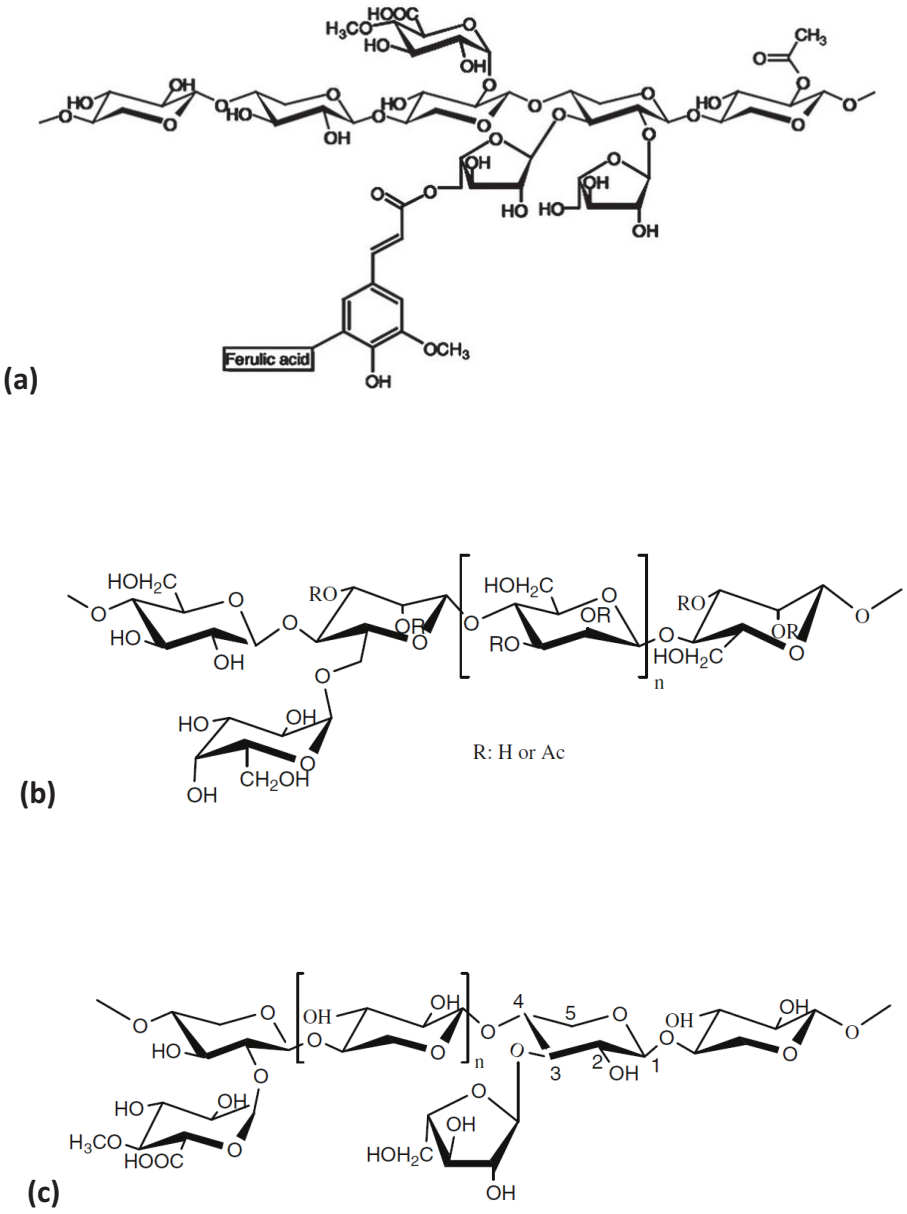


Figure 3. Structural representation of some hemicelluloses and their substituents. (a) xylan [49], (b) softwood galactoglucomannan, (c) softwood arabinoglucuronoxylan [50]. 'Ac' denotes acetate group

1.2.3 Lignin

Lignin is a complex, heterogeneous, polyaromatic macromolecular structure (Figure 4a) in plant cell walls that is built of cross-linked phenolic monomers [16]. Lignin is the second most abundant biopolymer on earth, surpassed only by cellulose. It glues cellulose, hemicellulose and other plant cell wall structural components together and is covalently linked to hemicellulose via ester/ether linkages to form lignin-carbohydrate complexes [51,52], which are still poorly characterized. In the plant cell wall lignin is predominantly found in the middle lamella where it binds adjacent cells and maintains cell structural integrity [53]. It also renders the plant cell impermeable and resistant to microbial invasion [54].

The monomeric building blocks of lignin are three phenyl-propionic alcohols, which are linked together by alkyl-aryl, alkyl-alkyl and aryl-aryl ether bonds. These monomers include: coniferyl alcohol (guaiacyl propanol), coumaryl alcohol (*p*-hydroxyphenyl propanol) and sinapyl alcohol (syringyl alcohol) (Figure 4b). Due to the random nature of polymerization of these building blocks, lignin possesses no definitive structure and has large, difficult-to-estimate molecular mass. Nonetheless, the frequency of each of the individual monolignol group types has been established [55]. Lignin is highly amorphous and is known to start melting (glass transition) at 135°C, a process which is facilitated by moisture [56]. As shown in Table 1, lignin content is dependent on biomass type. Generally, grasses and other herbaceous plants have the lowest contents of lignin, whereas softwoods have very high lignin contents. The structure of lignin is shown in Figure 4a.

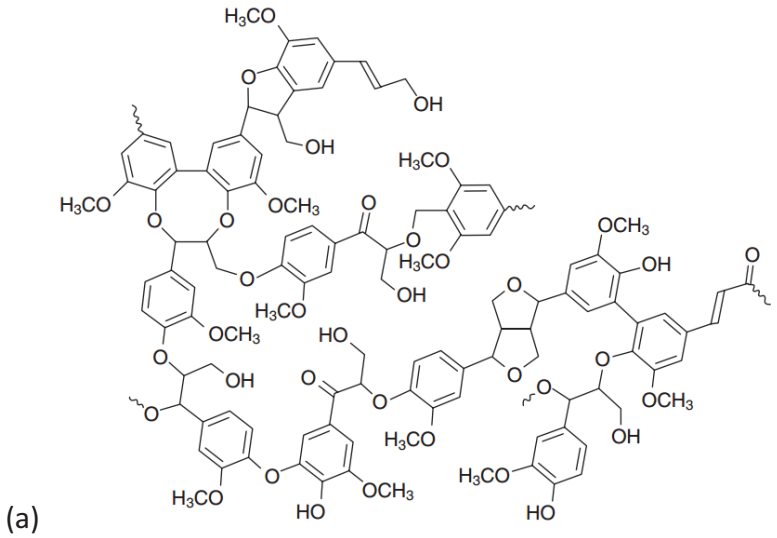


Figure 4. Chemical structure of lignin showing (a) the partial constitution of the large molecular complex and (b) the major monolignol building blocks in plant lignin (adapted from references [57] and [55], respectively)

1.3 The biorefinery concept for bio-based fuels and chemicals

The biorefinery concept is an integrated, environmentally friendly and cost-effective approach to maximize biomass utilization through a combination of steps including the supply and preparation of biomass, the physicochemical, thermochemical and biochemical conversions processes and all the necessary downstream processing steps for the delivery of energy, high value chemicals and materials [19]. Co-production of fuels, materials and high-value chemicals ultimately reduces the cost of secondary energy carriers such as transport fuels, heat and power, which increases their market competitiveness. A schematic illustration of the biorefinery concept is shown in Figure 5. This integrated concept is based on the recognition that lignocellulosic biomass is spread over wide areas of land which posses a challenge in terms of logistics, and is heterogeneous, consisting of pentose and hexose polysaccharides and polyaromatic lignin, all of which must be fully utilized in order to maximize the energy potential of the feedstock. The heterogeneity of biomass inherently poses a severe challenge in biomass conversion, such that a single conversion approach may not be sufficient to fully harness the energy potential of the biomass. For instance, the biochemical transformation stream does not utilize lignin, which constitutes up to 30% of biomass. Besides its use as fuel source, lignin has several other possible applications such as the production of bioplastic, dispersants, surfactants, activated carbon, carbon fibre, binders, phenol, sorbents, vanillin, mineral pelleting and granulating production aid, oil drilling additives, animal feed additive, road dust control, etc [58,59]. Lignin and other residuals from the biochemical conversion processes must therefore be transformed via physicochemical or thermochemical processes to maximize biomass utilization.

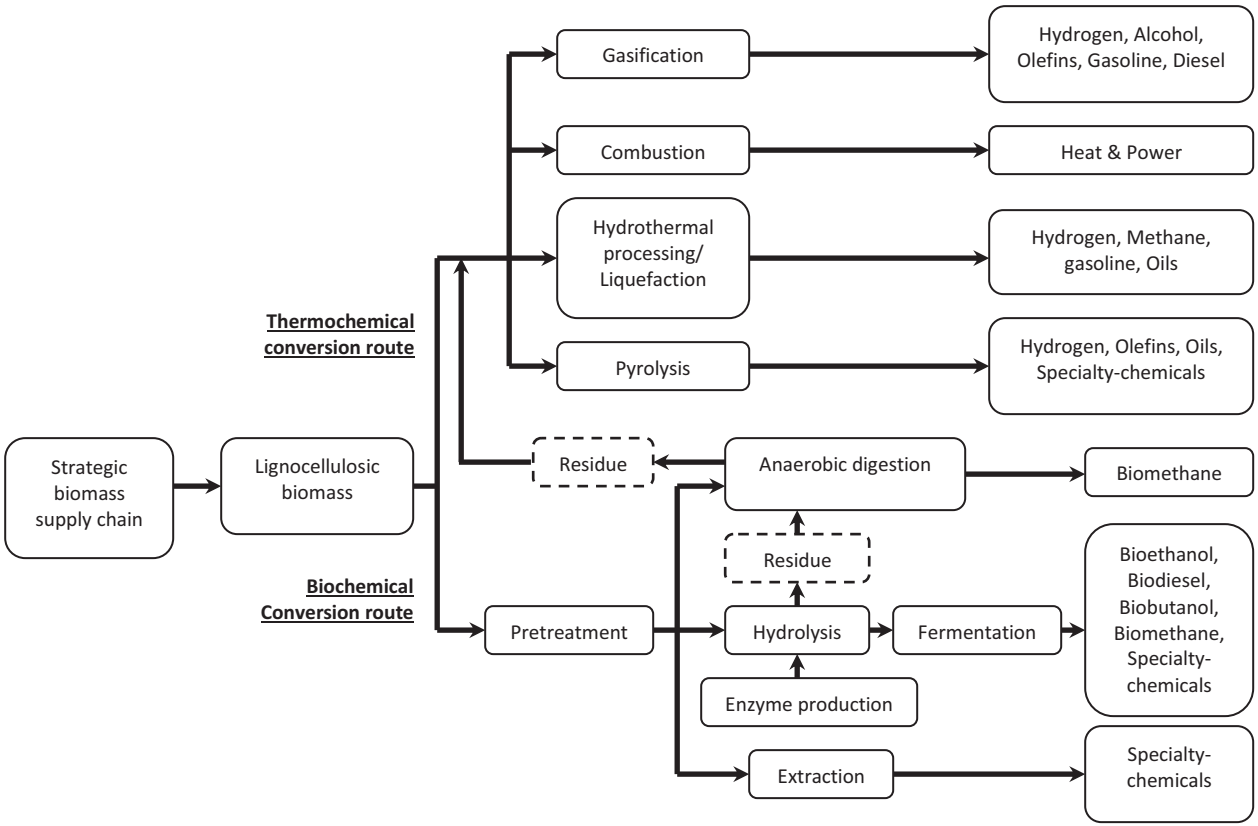


Figure 5. The biorefinery concept illustrating the thermochemical and biochemical transformation of biomass into heat, power, fuels and high value chemicals

1.3.1 Thermochemical process stream

The thermochemical process stream generally consists of a number of interlinked unit operations including biomass pre-treatment, feeding, conversion and product clean up and conditioning. Feedstock pretreatment mainly involves sorting, drying and size reduction. As shown in Figure 5, the main thermochemical conversion technologies include combustion, gasification, pyrolysis and hydrothermal liquefaction, which convert biomass and its residues to fuels, chemicals and power [50,60,61,62].

Combustion involves the direct burning of biomass at temperatures > 850 °C to produce heat and power [63,64]. This technique, although commonly used in pulp and paper mills and sugarcane processing plants, releases significant amounts of toxic gases and needs further improvement for efficiency. In gasification, biomass is heated under pressure with limited amount of oxygen (about one-third the amount necessary for complete combustion) to produce syngas – a mixture of CO and H_2 [50,60,61,65]. Flash pyrolysis involves heating biomass in the absence of oxygen to produce liquid bio-oil [66]. The operation temperature in gasification is generally lower than for pyrolysis (400 – 600 °C) [50,67]. Both syngas and pyrolysis oil can be directly used as fuel for heat and power generation or can be catalytically converted into a variety of fuels such as methanol, ethanol, short chain alkanes, dimethyl ether, H_2 and a variety of high value chemicals [50,67,68]. Biomass gasification is the most attractive thermochemical conversion technology because of its greater efficiency compared to combustion and that it is at a relatively advanced stage compared to flash pyrolysis, which is still in the developmental stage [66,69].

Hydrothermal liquefaction involves the catalytic depolymerization and conversion of carbohydrates and lignin in lignocellulosic material into oil that can be refined into various fuels. In this process, the polymers in lignocellulose react with water at >200 °C and high pressure (2 – 20 MPa) [70,71]. This technology mimics the natural process of fossil fuel formation from biomass, which makes it a very promising route for renewable

liquid fuels. However, this biomass conversion pathway requires high energy cost compared with biochemical conversion.

1.3.2 Biochemical process stream

In the biochemical transformation stream, polysaccharides in lignocellulose must be broken down by enzymes into their monomeric pentose and hexose sugar building blocks [72], which can be recovered or fermented by micro-organisms [73,74]. Sustainable biomass conversion via the biochemical transformation stream involves five critical steps: identification of suitable biomass type(s) and design of a reliable supply chain; effective pretreatment of the biomass to increase enzyme accessibility during hydrolysis; production of biomass degrading enzymes; fermentation of hexoses and pentoses and downstream processing [19]. For the scope of this thesis, only the pretreatment, hydrolysis and fermentation steps will be discussed.

1.3.2.1 Pretreatment

The structural complexity of lignocellulose renders it inherently recalcitrant to enzymatic degradation. Lignocellulolytic enzymes face a challenge in accessing the carbohydrates, particularly cellulose, in biomass due to structural barriers of hemicellulose and lignin and crystallinity of cellulose [75]. To improve enzymatic hydrolysis, lignocellulose is physically or chemically pretreated to hydrolyse the surrounding hemicellulose, remove or modify lignin, reduce cellulose crystallinity by amorphogenesis and increase enzyme accessibility to cellulose [16,76]. Several physical, chemical and biological pretreatment methods have been developed which improve the hydrolysis of lignocellulose. These include dilute acid hydrolysis [77,78,79], steam explosion [80,81,79], ammonia fibre expansion (AFEX) [82,83,84,85], liquid hot water [86,87,88], lime/alkali [79,89,90], organosolv [91], SPORL [92,93], thermo-mechanical comminution [94], kraft pulping [95], ionic liquids [96] and biological [97,98,99] treatment methods.

Each of the available pretreatment methods possesses inherent advantages and shortcomings. The differences in biomass complexity or biochemical composition may require that pre-treatment be biomass type-specific. Ultimately, pretreatment must be efficient and cost-effective.

Depending on the type and severity (temperature and concentration of chemical agents), pretreatment may modify lignin and carbohydrates in biomass to produce compounds that inhibit enzymes [100,101,102] and show toxicity to fermentative yeast [103,104]. For instance, although steam explosion is generally regarded as the most cost effective pretreatment method, it degrades hemicellulose and transforms lignin. This reduces yields of recoverable hemicelluloses and lignin and generates microbial and lignocellulolytic enzyme inhibitors [16,19,105]. Kraft pulping is known to efficiently remove lignin and hemicelluloses from biomass, but it is harsh and remains very expensive for use as a pretreatment method for biomass meant for the production of fuels and chemicals. However, it proves cost-effective in cases where paper and pulp facilities are integrated into biorefineries for bioproducts delivery [106].

Although efficient in hydrolysing lignin and hemicellulose with minimal formation of enzyme and microbial inhibitors, organosolv pretreatment is expensive due to high solvent cost and high process cost to drain, evaporate, condense and recycle the solvents after every batch [16].Thermo-mechanical comminution reduces biomass particle size and to a limited extent, cellulose crystallinity and increases accessible surface area to facilitate enzymatic hydrolysis, but requires energy which is usually higher than the inherent biomass energy [16]. Biological pretreatment does not require chemicals, consumes minimal amounts of energy, and utilizes living organisms or their products in mild, environmentally friendly conditions [98,99]. However, this method is slow, requires stringent process conditions and large amount of space, and leads to loss of cellulose and hemicellulose [19]. Table 2 summarizes the features of most of the available biomass pretreatment methods.

Table 2. Summary of the features of various lignocellulosic biomass pretreatment methods (adapted from references [19,16])

Pretreatment method	Sugar yield	Inhibitor formation	Bioproducts generation	Chemical reuse	Success at pilot scale	Advantages	Limitations and disadvantages
Mechanical comminution	Low	None	None	No	Yes	Reduce cellulose crystallinity	Expensive equipment, power consumption higher than inherent biomass energy
Dilute mineral acid	High	High	High	Yes	Yes	Cellulose and hemicellulose Hydrolysis alters lignin structure	Hazardous, toxic and corrosive
Alkali	High	Low	High	Yes	Yes	Hemicellulose & lignin removal, low equipment cost	Long residence time, irrecoverable salt formation
Liquid hot water	High	High	Low	No	Yes	Hemicellulose removal, increased cellulose accessibility	Long residence time, limited or no lignin removal/modification
Organosolv	High	High	High	Yes	Yes	Hydrolyzes lignin & hemicellulose	Solvent drained, evaporated & condensed before reuse
Wet oxidation	High / Low	None	Low	No	-	Removes lignin, dissolves hemicellulose & decrystallizes cellulose	-
Ozonolysis	High	Low	High	No	Yes	Removes Lignin, no toxic or inhibitory compounds produced	Large amounts of ozone required, expensive
CO_2 Explosion	High	Low	High	No	-	Removes hemicellulose, decrystallizes cellulose, increases accessible surface area, no inhibitory compounds formed, inexpensive	No lignin removal or modification
Steam Explosion	High	High	Low	-	Yes	Removes hemicellulose, modifies lignin	Incomplete dissociation of lignin-carbohydrate matrix, destruction of a portion of the xylan fraction; generates microbial inhibitors
AFEX	High	Low	Low	Yes	-	Removes some amount of lignin and hemicellulose, increases accessible surface area, no inhibitors that affect downstream processes	Inefficient for biomass with high lignin content
Ionic liquids	High / Low	Low	-	Yes	-	Dissolves cellulose, increases amendability to cellulose	Still in initial phase of development
Biological	Low	None	None	Possible	-	Degrades lignin & hemicelluloses; low energy requirements, mild process conditions	Slow process due to very low rate of hydrolysis, stringent conditions, large space requirement, non-economical

Dilute-acid hydrolysis and steam explosion are the most commonly-used biomass pretreatment techniques because they have been developed and demonstrated industrially [107,108,109].

1.3.2.2 Enzymatic hydrolysis of lignocellulose

The structural complexity and heterogeneity of lignocellulose [110] in plant cell walls necessitates a diverse enzyme repertoire for complete hydrolysis of the constituent carbohydrates and lignin polymers into reducing sugars [14,111] and aromatics [112,113,114]. Co-evolution of saprotrophic or plant parasitic fungi and photosynthetic plants has ensured that the genomes of saprotrophic and plant parasitic fungi are endowed with genes that encode a variety of lignocellulose degrading enzymes to meet their nutritional needs or guarantee survival via pathogenesis [115,116]. Fungal saprotrophs therefore secrete complex protein cocktails containing many lignocellulolytic enzymes that deconstruct lignocellulose to fermentable sugars [117,118,119,120,121,122]. Genomes of several fungal saprotrophs such as *Trichoderma reesei*, *Thielavia terrestris, Myceliophthora thermophila* and *Aspergillus niger* have been sequenced, and are publicly available [123,124]. This facilitates further enzyme development for biomass deconstruction.

Cell wall degrading protein cocktails from fungal saprotrophs typically include carbohydrate-active enzymes (CAZy) [125,126] and oxidative lignin-active enzymes [127] which synergistically interact in the deconstruction of lignocellulose to sugars and phenolics. Protein cocktails secreted by fungal saprotrophs can contain more than 80 enzyme components [121,122,128]. The most-studied CAZymes are the glycoside hydrolases (GHs), which are extensively classified into families in the CAZy database [126,129,130].

Enzymatic hydrolysis of glycosidic linkages by GHs occurs through acid catalysis that requires a proton donor and a nucleophile/base [131,132], and involves either an inversion or retention mechanism. The catalytic mechanisms are distinguished

depending on the configuration of the anomeric carbon in the product. In the retention mechanism, the stereochemical configuration of the anomeric carbon is unchanged from substrate to product, whereas this configuration is reversed in the inversion mechanism [132].

The main GH enzymes involved in the hydrolysis of lignocellulose to sugars include the cellulose-degrading enzymes and hemicellulases, especially xylanases. Both groups of enzymes act synergistically [133]. Accessory enzymes including other hemicellululases, pectinases [134,135], non-hydrolytic cellulose active proteins [136] and fungal oxidative lignin enzymes [137] also play important roles in biomass deconstruction. The role of each of these enzyme categories in the hydrolysis of cellulose or hemicellulose is described below.

1.3.2.2.1 Cellulose degradation

The enzymes required for the complete hydrolysis of cellulose are generally referred to as cellulases. Cellulase activities are distributed across eight GH families, and include: β-1,4-endoglucanases (EGL), cellobiohydrolases (CBH), and β-glucosidase (BGL) [125,138,139]. These proteins act concertedly such that the hydrolytic function of one potentiates the other [140,141]. For instance, EGLs decrease the DP of cellulose chains by random hydrolysis at internal amorphous positions to expose new points of attack for CBHs, which then processively hydrolyse crystalline cellulose chains releasing cellobiose [142]. Cellobiose is further hydrolysed into glucose by BGLs [143]. EGLs and CBHs are therefore thought to increase and decrease, respectively, the crystallinity index of cellulose [144]. CBHs which attack cellulose from reducing ends of the cellulose chain are referred to as CBHI, whereas those that prefer the non-reducing ends are CBHII [145]. Cellobiohydrolases are often the most abundant enzyme component in fungal cellulase secretomes. For instance, the cellobiohydrolases Cel6A and Cel7A account for up to 80% of the total *T. reesei* secreted cellulases (with Cel6A and Cel7A typically account for 15-20% and 60% of the total, respectively [146]). Cel7B and Cel5A, the other major components, make up to 10% each. The higher proportion by mass of

cellobiohydrolases in secreted cellulase cocktails from filamentous fungi appears to be a natural compensation for their low turnover numbers, which are typically less than 4 s^{-1} [147].

Studies on the hydrolysis of untreated cellulose suggest that random internal breakdown of cellulose chains by EGL enzymes into cello-oligomers is the overall rate-limiting step. This explains why CBH activity decreases with increasing average chain length of cellulose: a corresponding decrease in the number of points of attack [148,149,150,151]. Meanwhile, the rate-limiting step in cellobiohydrolase activity is the low rate of product dissociation from the active site [152,153]. Furthermore, the CBHs are very sensitive to product inhibition by cellobiose, which is another factor that is thought to account for the high amount of CBHs in fungal cellulase mixtures [154]. BGLs are also affected by product inhibition from glucose [155].

Other cellulose-active accessory proteins have recently been found to contribute to cellulase hydrolysis of cellulose via non-hydrolytic mechanisms. These proteins include the polysaccharide monooxygenases (PMOs, previously referred to as GH61s), the fungal flavocytochrome – cellobiose dehydrogenases (CDHs), plant expansins and their fungal equivalent, swollenins [137,136]. PMOs which share structural and functional homology with CBM33 [156,157,158,159] oxidatively disrupt crystalline cellulose in the presence of molecular oxygen and reducing agents such as ascorbate and gallate [136,160]. Unlike the canonical hydrolytic cellulases, PMOs release a variety of native and aldonic C1-, C4- or C6-oxidized cellodextrins from cellulose [136,161,162,163,164] with varying DP [160]. By generating new non-oxidized ends for cellulase attack, this action increases cellulose susceptibility to enzymatic degradation [136]. The catalytic activity of PMOs is synergistically enhanced by CDH. Together, PMOs and CDHs promote cellulose degradation by cellulases [165,166]. Two groups of PMOs have been distinguished: Type I and Type II PMOs, which yield products oxidized at C1 (reducing end) and C4 (non-reducing end), respectively [156,162,167].

Expansins are non-hydrolytic proteins that are thought to reduce cellulose crystallinity and loosen cell wall matrix compactness by disrupting intra-molecular and inter-molecular non-covalent bonds between cellulose polymers or cellulose and other matrix polymers. This facilitates slippage of cellulose microfibrils past each other within the cell wall [136,168,169,170,171,172]. This increases the proportion of amorphous regions in cellulose, and reduces the hemicellulose protective barrier around it, both of which favour degradation of lignocellulose by classical cellulases. Swollenins are fungal, expansin-like proteins that loosen, partially disrupt and cause swelling of plant cell walls, without releasing any reducing sugars [173,174]. Both expansins and swollenins have been reported to enhance cellulase hydrolysis of lignocellulose [136]. A summary of cellulose hydrolysis is shown in Figure 6 [167,136].

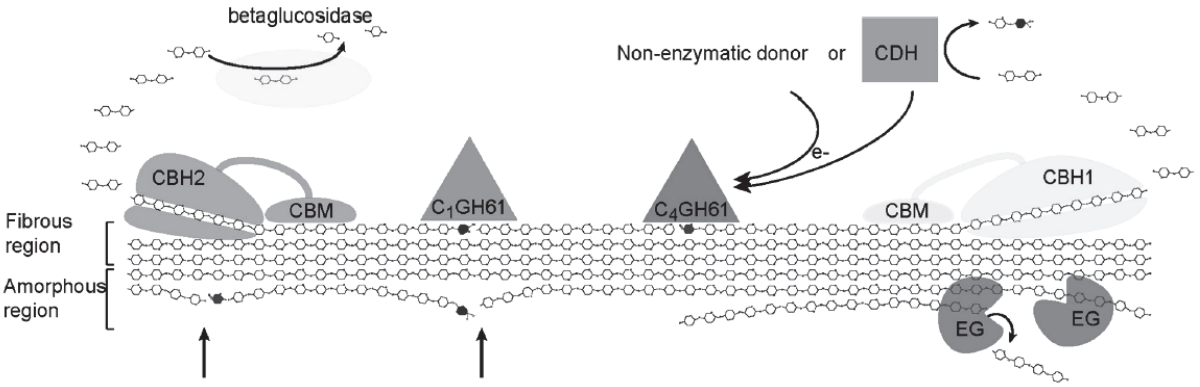


Figure 6. Illustration of the current understanding of fungal enzymatic degradation of cellulose by Horn *et al.* [167]. Abbreviations: EG, endoglucanase; CBH, cellobiohydrolase; CDH, cellobiose-dehydrogenase; CBM, carbohydrate-binding module. The picture shows a C1 and a C4 oxidizing GH61, which generates non-oxidized ends that serve as new points of attack for the CBHII (arrow on the right) and the CBHI (arrow on the left), respectively (oxidized sugars are colored black). Like non-enzymatic electron donors such as ascorbic acid, gallic acid and reduced glutathione, CDH plays the role of reductant in GH61 activity

1.3.2.2.2 Hemicellulose degradation

Enzymatic degradation of hemicellulose in lignocellulosic biomass ultimately yields monomeric pentose and hexose sugars, and increases cellulase accessibility to cellulose. Hemicellulases therefore synergize with cellulases in lignocellulose hydrolysis [175, 176, 177]. The complexity and heterogeneity of hemicelluloses necessitates a specialized arsenal of enzyme activities for their complete hydrolysis. The three main hemicellulose backbones – xylan, xyloglucan and (galacto[gluco]-) mannan are degraded by the following groups of GH enzymes: β-1,4-endoxylanase and β-1,4-xylosidase for xylan; xyloglucan-specific β-1,4-endoglucanase and β-1,4-glucosidase for xyloglucan; and β-1,4-endomannanase and β-1,4-mannosidase for mannan. In the case of galactomannan, α-1,4-galactosidase is also required in addition to the mannanases. Similarly, galactoglucomannan requires β-1,4-glucosidase in addition to the listed mannan-active enzymes [154,178,179]. Xylanases, the most important hemicellulases (xylan being the most abundant hemicellulose), belong to two main GH families based on amino acid sequence similarities [180]: GH10 xylanases which show broader substrate specificity than the GH11 xylanases, enabling the former to better degrade highly substituted xylans [181,182,183]. Several debranching enzymes play an important role in the hydrolysis of hemicellulose. Figure 7 illustrates the involvement of debranching enzymes in hemicellulose (xylan) degradation.

Oxidative lignin-degrading enzymes such laccases, manganese peroxidase and lignin peroxidase, typically originating from white rot fungi, delignify biomass to expose the carbohydrates buried within the matrix [112,113,114]. Both the hemicellulases and ligninases synergistically enhance cellulase hydrolysis of the cellulose component of lignocellulose by increasing accessibility to cellulose.

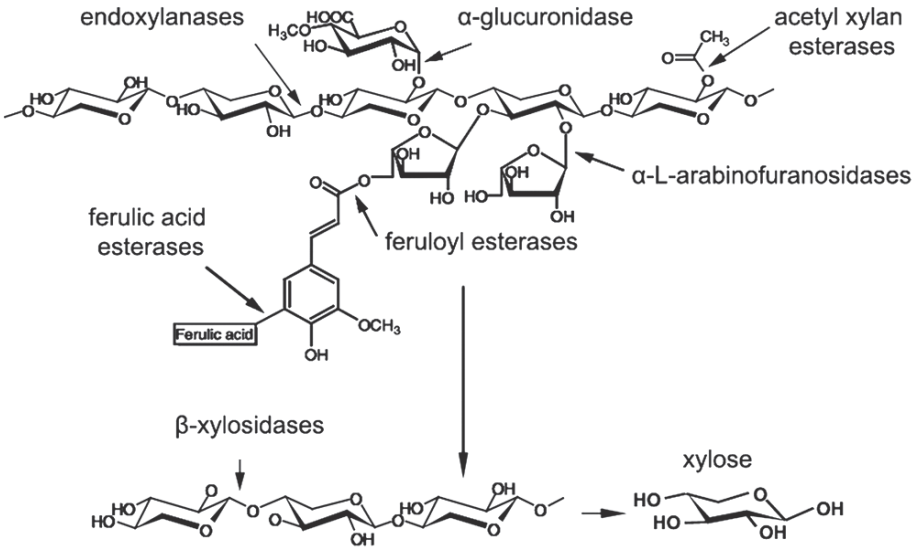


Figure 7. Schematic representation of xylan degradation illustrating a variety of enzymes involved (adapted from references [184,185])

1.3.2.3 Fermentation

Humankind from diverse cultural backgrounds has been practicing fermentation as far back as the 1700 BC [186]. The role of microorganisms, especially yeast, in fermentation was demonstrated by Louis Pasteur and contemporaries by 1860 [187,188,189]. Unlike the fermentation of hexoses, which has been a well-established bioprocess since the 1800s, pentose fermentation is a relatively new area of knowledge in which pentose sugars such as xylose and arabinose are converted into fuels and valuable chemicals by microorganisms. Cellulose and xylan are the main sources of sugars from lignocellulosic biomass hydrolysis. The released sugars (mainly glucose, xylose and to a lower proportion, arabinose) may either be recovered by various

downstream processing steps or be biologically converted into liquid fuels and high value chemicals such as methanol, ethanol, xylitol, butanol, 2,3-butanediol, acetone, isopropanol, propanol, lactic acid, hydrogen, etc. by fermentative microorganisms such as *Saccharomyces cereviceae* [190,191,192]. Unlike enzymatic hydrolysis of lignocellulose, fermentation technologies are relatively advanced. Metabolic engineering strategies have been used to develop various robust industrial cell factories which possess increased tolerance to inhibitors from biomass pretreatment, efficient and specific pentose transporters, improved substrate channeling for higher pentose fermentation rates and ability to co-ferment hexoses and pentoses [193,194,195]. Furthermore, consolidated bioprocessing technologies are being used to develop strains with capabilities of cellulase production, cellulose hydrolysis and sugar fermentation in a single step [196,197,198]. This strategy is intended to reduce process cost of lignocellulose derived fuels and chemicals. However, it is still an unmet challenge to obtain a single organism that delivers high lignocellulolytic enzyme productivity, sugar yield and ethanol titer directly from lignocellulose [199]. Other proposed approaches include microbial co-cultures for simultaneous saccharification and fermentation [200], as well as the development of consortia of cellulolytic and fermentative organisms that utilize scalable symbiosis mechanisms to provide stable bioprocesses [201].

1.4 Enzymatic biomass conversion technologies and challenges

The catalytic efficiencies of enzyme cocktails from fungal saprotrophs are adequate in meeting the nutritional needs of fungal species in the natural environment. However, these cocktails are not necessarily efficient in commercial-scale degradation of lignocellulose for advanced bio-based products, and it is generally agreed that a lot of improvement is required for them to be directly used as catalysts on commercial scale projects utilizing lignocellulosic biomass for the production of fuels and chemicals [202].

Many commercial cellulase cocktails for lignocellulose hydrolysis have been developed. Although a broad spectrum of biomass pretreatment methods has been developed to improve on the hydrolytic efficiency of cellulase mixtures [203,204] commercial cellulases are still not cost-effective in the hydrolysis of pretreated feedstock at commercial scale [205]. Among the contributing factors are the residual crystallinity of cellulose, and the physical barrier of hemicellulose and lignin surrounding cellulose (e.g. for biomass pretreated by steam explosion), which reduce susceptibility of the feedstock to enzymatic hydrolysis [206]. Furthermore, lignin and other structural components non-productively adsorb the enzymes, preventing their bioavailability for lignocellulose hydrolysis. The presence of naturally occurring and pretreatment-generated enzyme inhibitors (e.g. phenolics, furfural and hydroxymethylfurfural) is unfavourable to enzyme function [207,206,208]. All these factors increase processing costs [206]. For instance, recent cost estimates for cellulase enzymes for the bioconversion of corn stover range between US$ 0.68 and US$1.47 per gallon of ethanol, making enzymes the second largest contributor (behind feedstock) to operating costs in the lignocellulose to ethanol bioprocess [205].

Several approaches have been and are being used to improve on the catalytic efficiency of cellulase cocktails for biomass saccharification. These strategies are focused on three main areas (Figure 8): modification of plants to produce biomass which is easy to hydrolyze; development of efficient pretreatment methods (as discussed above); and development of more efficient enzymes or enzyme mixtures for hydrolysis [209].

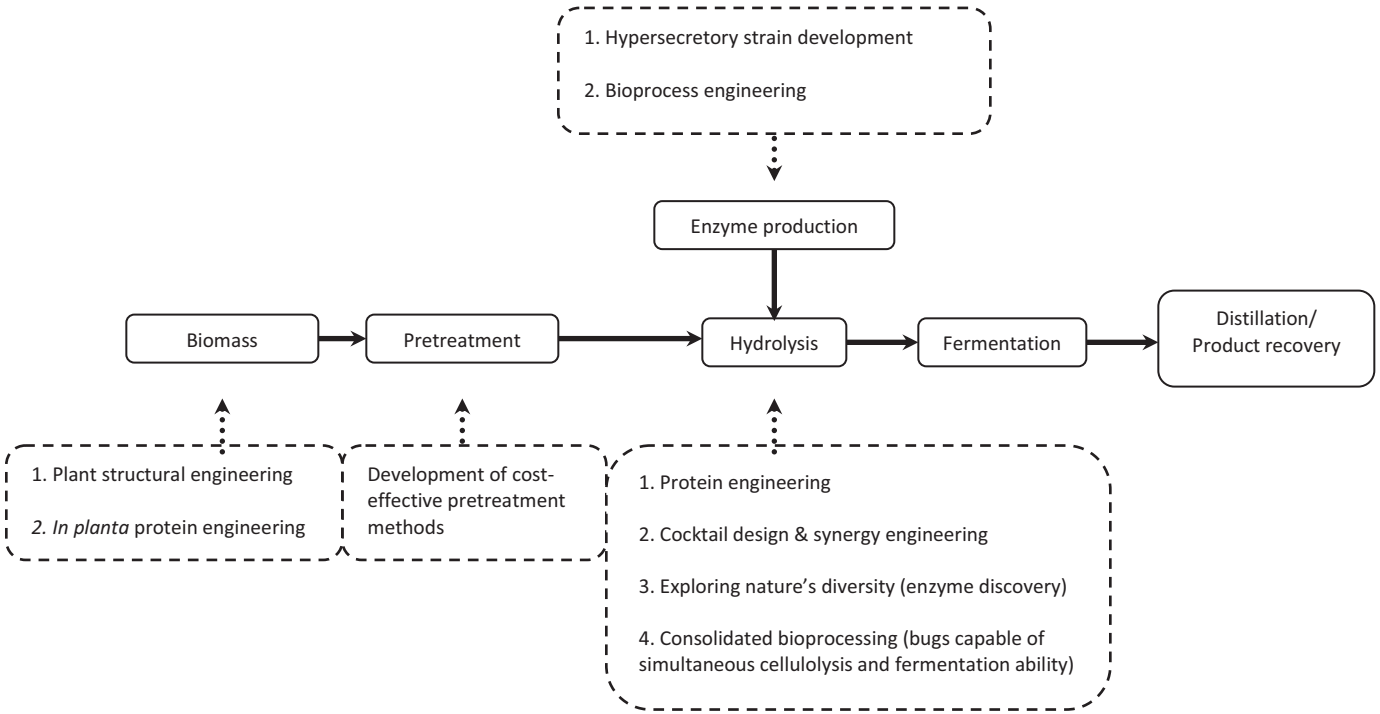


Figure 8 Schematic representation of the strategies (in dotted lines) used to overcome challenges associated with the enzymatic conversion of lignocellulosic biomass [209]

Strategies that modify plants to ease hydrolysis include the engineering of plants to lower the lignin content [210,211,212], enrich the polysaccharide composition [213,214,215] and co-express cell wall degrading enzymes that are inactive at physiological temperatures but can be induced by post-harvest temperature conditioning [216,217,218,219]. However, the reduction of plant lignin content has also been accompanied by major challenges such as reduction in plant growth, biomass output, stalk strength, or pathogen resistance [220]. Similarly, the *in planta* constitutive expression of cell wall degrading enzymes has been shown to reduce plant growth, change leaf morphology and reduce stress tolerance [221,222,218]. Furthermore, many of these techniques have been used almost exclusively in non-wood model plant systems. Their effects on woody biomass is as yet unknown.

The strategies employed in the development of efficient lignocellulolytic enzymes (Figure 8) can be grouped under the following categories: protein engineering; exploration of nature's diversity for efficient microbes and enzymes; and cellulase cocktail design and synergy engineering.

1.4.1 Protein engineering

Protein engineering strategies commonly used to improve the catalytic efficiency of plant cell degrading enzymes [223] are site-directed mutagenesis [224], directed evolution [225], construction of protein chimeras or cellulosomes [226] and synergy engineering [209]. The industrial microbial factory, *T. reesei* [227] has played a central role in understanding the biochemistry of enzymes that deconstruct plant cell wall material. Since its discovery six decades ago [228], *T. reesei* QM6 has been upgraded through cycles of classical random mutagenesis with mutant selection strategies to produce the publicly available hyper-secretory strain, *T. reesei* RUT-C30, which secrets cell wall-degrading enzymes at levels that exceed 30 g/L [229,230,231,232]. Improved industrial strains derived from this mutant, such as *T. reesei* CL487 [233] have been reported to secrete cellulase proteins in excess of 100 g/L

[227,234]. In addition to its GRAS (Generally Recognized as Safe by the USFDA) status, this hyper-secretory prowess makes *T. reesei* the benchmark of industrial microbial systems for the regulation, biochemistry and scalable production of cell wall-degrading enzymes [235,236,206].

Despite the progress made in increasing the enzyme production capability of *T. reesei*, cell wall degrading enzyme cocktails for the production of fuels and chemicals from biomass still requires optimization [237,238,239,240]. Understanding the genetic basis of the hyper-secretory capabilities of *T. reesei* RUT-C30 and *T. reesei* CL487 can help identify key differences in other species which may be instrumental in engineering new hyper-secretory strains from *T. reesei* or other cell wall-degrading filamentous fungal species such as *A. niger, T. terrestris* and *M. thermophila*.

Additional limitations of commercial cellulases are: they are sourced mainly from mesophilic fungi (*T. reesei* and *A. niger*) and are most effective at 50°C or below [246,241,242] and labile at higher temperatures; and that they are less efficient at high feedstock loadings. Conversion of high concentrations of biomass at temperatures above 50°C is known to enhance substrate solubilization, improve mass transfer, increase the reaction rates of thermotolerant biocatalysts, shorten reaction time [246,118], minimize water use and ultimately reduce process cost [243]. Thermostabilization of fungal proteins which accelerate liquefaction of high loadings of feedstock at elevated temperatures could significantly impact the commercialization of the lignocellulose to fuels bioprocess [244,245,246,247]. Protein engineering strategies are being used to developed thermostable cellulases [224,225,248,249].

1.4.2 Exploring nature's diversity for efficient microbes and enzymes

The diversity of cellulolytic species, especially saproptrophic and phytoparasitic fungi and bacteria, is extensive. These organisms have evolved over millions of years to

occupy and adapt to diverse ecological niches. Their adaptation is accompanied by enormous genetic resources that can be mined for cellulases and accessory proteins with improved lignocellulolytic efficiency [250]. However, the existing commercial cellulases are sourced from a very small number of fungal species. Bioprospecting now offers the opportunity to tap from the extensive already-known microbial culture collections as well as explore the vast microbial communities in extreme ecosystems for efficient enzymes. Examples of approaches in use are: the isolation of microbes that grow better on biomass substrates; sequencing the genomes/metagenomes of diverse microbial species/communities; mining databases of sequenced genomes/metagenomes to find new genes; and cloning variants of known enzyme genes [21,251].

The Joint Genome Institute has provided a platform for the sequencing of microbial genomes and metagenomes, especially those of lignocellulos-degrading extremophiles [123]. Many of these genomes have been sequenced and the sequencing of many metagenomes is on-going. Although the sequencing projects of genome targets have been on a steady rise in recent years, they still remain relatively few compared to the phylogenetic and ecological diversity of plant cell wall degrading fungi and bacteria available. This necessitates further exploration. For instance, the sequenced genomes of *T. reesei* QM6 [117] (http:// genome.jgi-psf.org/) and its hyper-secretory mutant RUT-C30 have provided opportunities to better understand the genomic basis of the high protein production and secretion capabilities of the latter through comparative genomic studies [252]. In addition, comparative genomic analysis of the *T. reesei* genome and those of other cell wall degrading fungal species reveal that the *T. reesei* genome harbours 200 GH-encoding genes, the lowest number of all the lignocellulose-degrading fungal species which have had their genomes sequenced [117,227]. It is thus worthy to explore the genomes of other organisms for the cellular traits and biosynthetic characteristics required for the production of renewable fuels and chemicals: e.g. the ability to efficiently degrade lignocellulosic materials, demonstrate resistance to

substrate and product inhibition and the capability to biosynthesize specific biofuels with high yield [253].

Fungal species, especially thermophiles such as *M. thermophila* and *T. terrestris,* which are known to secrete lignocellulolytic enzymes, may harbour genes which encode proteins with superior hydrolytic potential [118]. Many thermostable cellulose active hydrolases from thermophiles have been characterized [254] and their potential application in lignocellulose bioconversion discussed [246]. The thermophilic species of the genius *Myceliophthora* are known to hydrolyse and grow efficiently on lignocellulosic biomass, secreting enzymes with thermal stability up to 70°C [241,255,256,257,258]. *M. thermophile* can grow on cellulose at the same rate as it grows on glucose, is known to secret an endoglucanase that liquefies 18% biomass [259]. This species potentially harbours other proteins that could liquefy high consistency biomass. The genomes of *M. thermophila* and *T. terrestris* have been sequenced and are publicly available for the research community to exploit for the development of efficient industrial enzymes for lignocellulose degradation. Bioprospecting for enzymes that show prolonged tolerance to temperatures > 50°C, inhibitors and high biomass solid-loadings as well as the development of hyper-secretory strains can be random or guided by evolutionary or ecological principles [21,251].

1.4.3 Cellulase cocktail design and synergy engineering

Synergistic catalysis in the degradation of lignocellulosic substrates by enzymes secreted by individual or consortia of microorganisms occurs in the natural environment to meet the nutritional needs of the species [253]. Many bacteria and fungi are known to produce efficient plant cell wall degrading enzymes that can be used to supplement commercial cellulase systems, such as the *T. reesei* system [21,176,260,261,262]. Optimization of commercial cellulase cocktails by such enzyme additives and the design of novel efficient enzyme cocktails using the minimal number, amount and optimal

assembly of the best-performing enzyme(s) for improved cellulose hydrolysis have attracted considerable interest in recent years [263,21,264].

The identification of the best-performing enzymes necessitates the mining and expression of new genes to find those encoding individual cell wall degrading enzymes with superior catalytic properties. Biochemical characterization helps identify industrially relevant properties of enzymes with respect to specific activity, pH and temperature optima, substrate specificity, thermal stability, salt and inhibitor tolerance, etc. The development of screening tools for the assessment of the hydrolytic capabilities of such enzymes on natural lignocellulosic substrates is very important because activities on model substrates may not necessarily correlate with those on natural biomass [134].

Cell wall degrading enzymes with improved catalytic properties have been identified and characterized using such a strategy, examples of which include thermostable cellulases [246,265,266], cellulases with ability to liquefy lignocellulosic biomass at high solid loadings [259,267], cellulases with tolerance to product inhibition [268,269], non-hydrolytic cellulose active proteins [136] and proteins sourced from extremophiles [270]. Significant improvements in lignocellulose hydrolysis have been obtained by supplementing cellulase cocktails with crude fungal enzyme preparations [260] and auxiliary enzymes such as xylanases, acetylxylan esterases, arabinofuranosidases, xylosidases, mannanases and pectinases [209,263,271,272]. High-level synergism is possible in such cases because auxiliary proteins can access insoluble biomass and disrupt highly ordered polymer packaging surrounding cellulose, thereby facilitating attack by cellulases.

Non-hydrolytic cellulose active proteins including the PMOs, expansins and swollenins have been used to supplement cellulase cocktails for improved catalysis of lignocellulose degradation [136]. The recently identified cellulose induced protein (CIP1 and CIP2) was found to show synergy with PMOs and swollenins. CIPs are thought to play a role in the cleavage of hemicellulose-lignin crosslinks [273,272].

Multi-enzyme complexes from anaerobic bacteria called cellulosomes are known to possess multiple catalytic sites (cellulolytic and hemicellulolytic activities) and CBMs linked by dockerins to the same scaffoldin (CipA) which is anchored to the bacterial cell by cohesin. In nature, these activities synergistically degrade lignocellulose into sugars [274,275,276]. Synergy engineering strategies are being used to design synthetic cellulosomes with improved cellulolytic activity. This strategy involves the docking of enzyme catalytic domains from different microorganisms to synthetic hybrid scaffoldin molecules which are anchored via cohesins onto bacterial cells [274,277]. Using this technique, four different types of synthetic cellulosomes have been described: cellulase-cellulase, hemicellulase-hemicellulase, hemicellulase-cellulase and hemicellulase-carbohydrate esterase systems [278].

1.5 Approach employed in this study

The work carried out during this PhD project was part of Theme 2 (Enzymatic hydrolysis of lignocellulosic biomass) of NSERC Bioconversion Network, a unique Canadian R&D network that is aimed at developing energy efficient, commercially viable and environmentally sustainable biomass conversion processes that generate ethanol and high value co-products. The work was focused on the development of lignocellulolytic enzymes from saprophytic fungi for the efficient hydrolysis of a variety of pretreated woody biomass types from Canadian forests. The project involved the following three parts: (1) mining of genes of cell wall degrading enzymes from the sequenced genomes of filamentous fungi and expressing them in different host systems to assess host-related improvements in activities (Manuscripts I and II discussed in Chapters 2 and 3, respectively); (2) development of a high throughput-adaptable screening method for the evaluation of the hydrolytic capability of biomass degrading enzymes from filamentous fungal species under conditions of high biomass loadings (Manuscript III discussed in Chapter 4); and (3) identification of efficient biomass liquefying enzymes from thermophilic fungal saprophytes (Manuscript IV discussed in Chapter 5). Finally, a literature review of the potential role of non-hydrolytic cellulose

active proteins in the bioconversion of lignocellulosic biomass for renewable fuels and chemicals is presented in Manuscript V (discussed in Chapter 6).

In part 1, genes of plant cell wall degrading glycoside hydrolases were mined from the publicly available genome of *A. niger* strain ATCC 1015, cloned and expressed in four expression hosts: *A. niger*, *P.* pastoris, *E. coli* and *N. benthamiana* (tobacco). The recombinant proteins were purified and biochemically characterized in terms of specific activity, pH and temperature optima and glycosylation differences. The data were included in a library of characterized fungal enzymes, which can potentially be useful in the identification of enzyme candidates for supplementing commercial cellulases or for the design of efficient minimal enzyme cocktails meant for the hydrolysis of high-consistency lignocellulosic biomass.

In the work in part 2 (reported in Chapter 4, Manuscript III), a customized, high-throughput compatible 96-well microfilter plate method was developed for screening recombinant lignocellulolytic enzymes with various pretreated biomass types at solids loading of 15% w/v or higher. This simplified medium- to high through-put adaptable system can be used to assess the hydrolytic potential of enzymes on natural substrates at high solids loading.

Part 3, reported in Manuscript IV (Chapter 5), was focused on developing proteins which liquefy high consistency biomass from the thermophilic saprotroph, *M. thermophila* strain ATCC 42464. The strain was grown on a variety of pretreated lignocellulosic biomass as sole carbon source, namely: organosolv pretreated lodgepole pine (OLP), organosolv pretreated poplar (OPP), acid catalysed steam exploded poplar (SPP), aspen kraft pulp (AKP), black spruce kraft pulp (BSKP) and aspen thermomechanical pulp (AMP). The *M. thermophila* secreted protein cocktails were analysed by mass spectrometry and tested for the liquefaction and hydrolysis of various pretreated biomass types at 15% solid-loading. A pair of custom designed parallel plates was used for rheology measurement of high consistency biomass on a parallel plate rheometer following enzyme treatment at different temperatures.

Chapter 2 Manuscript I

A comprehensive comparison of fungal glycoside hydrolases produced in bacterial, yeast, filamentous fungal and plant expression hosts

Xiao Zhang[1], Enongene Ekwe[1], Andrew J. Conley[2†], Angelo Kaldis[2†], Sarah Zerbs[3], Frank Collart[3], Jon Magnuson[4], Scott E. Baker[4], Rima Menassa[2], Justin Powlowski[1], and Adrian Tsang[1]*

[1]Centre for Structural and Functional Genomics, Concordia University, Montreal, Quebec; [2]Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, London, Ontario; [3]Argonne National Laboratory, Lemont, Illinois; [4]Pacific Northwest National Laboratory, Richland, Washington

†Current address: VTT Technical Research Centre of Finland, Espoo, Finland
*Corresponding author: Adrian Tsang adrian.tsang@concordia.ca

Running title: Host Expression Comparison of GH Proteins

2.1 Abstract

Background

Glycoside hydrolases play an important role in biofuel production through their ability to convert biomass into sugars and oligosaccharides. The published data on properties of glycoside hydrolases are typically obtained after expression in heterologous hosts, usually bacteria, yeasts and filamentous fungi. In almost all studies, a single host is used to produce the protein encoded by a target gene. Where a target gene is expressed in more than one host, the levels of protein production in the different hosts are usually reported rather than their biochemical properties. To date,

there has been no systematic evaluation of the efficiency or suitability of these hosts in producing optimally active enzymes.

Results

We selected 24 *Aspergillus niger* genes encoding glycoside hydrolases, attempted to express them in four hosts, and compared the biochemical properties of the resulting recombinant proteins. Biochemically active enzymes were obtained from *A. niger*, *Escherichia coli*, *Pichia pastoris* and *Nicotiana benthamiana* at a rate of 88%, 21%, 50%, and 91%, respectively. The pH optimum, temperature optimum and specific activity of enzymes produced in *N. benthamiana* are nearly identical to those produced in the native *A. niger* host. Biochemical properties of two xylanases, two polygalacturonases and a cellobiohydrolase differed significantly when produced in *E. coli* or *P. pastoris*. Protein sequence analysis and experimental evidence showed that the two xylanases produced in the four hosts are not glycosylated, suggesting that the difference in biochemical properties observed in proteins produced in *E. coli* and *P. pastoris* is not caused by absence of, or aberrant, protein glycosylation.

Conclusion

Our results indicate that *N. benthamiana* is an effective production host for fungal glycoside hydrolases. Some recombinant glycoside hydrolases produced in *E. coli* and *P. pastoris* have biochemical properties that are different from those produced in *A. niger* and *N. benthamiana*.

Keywords: fungal glycoside hydrolases; enzyme production hosts; *Aspergillus niger*; *Escherichia coli*; *Nicotiana benthamiana*; *Pichia pastoris*; pH optimum; temperature optimum; specific activity

2.2 Background

The glycoside hydrolase (GH) family of enzymes is a widespread, diverse group of biological catalysts that are responsible for hydrolysis of polysaccharides (reviewed in [279]). Based on sequence similarity, over 120 GH families have been categorized so far and they include cellulases, hemicellulases, and pectinases [280]. Glycoside hydrolases are used extensively in detergent, food, feed, textile and many other industries. However, the most rapidly-growing sector of the industrial enzymes market is expected to be in cellulosic biofuels production where GH family enzymes are critical for the efficient saccharification of polysaccharides for fermentation to alcohol or infrastructure compatible hydrocarbon biofuels or biofuel precursors. Owing to the complexity and recalcitrance of lignocellulose, efficient enzymes and combinations of enzymes, especially GHs, will be critical to ensure that the production of fuels derived from lignocellulosic biomass is economically competitive [206].

A rapid approach to expanding the repertoire of plant cell-wall-degrading GHs, which in most cases are secreted enzymes, is microbial genome mining [281,124]. Fungi are especially rich reservoirs of these enzymes, and the genomes of an increasing number of lignocellulosic biomass decomposers are being sequenced [117,118,282,283,284]. On average, each of these genomes harbours ~200 GH-encoding genes. Although existing bioinformatics tools can identify homologues of GHs and assign function based on similarity, they are unable to predict the relevant catalytic properties. To date, only direct biochemical assays of the gene products can determine properties such as specific activity, temperature and pH optima; and direct application assays are needed to determine industrial utility. To link the functionality of the enzymes to their encoding genes, two general approaches have been deployed: 1) for novel enzyme activities, native proteins are purified and their encoding genes are cloned based on partial protein sequence information [285,286,287]; and 2) for homologues of GHs, recombinant enzymes are typically used for biochemical determinations [288,289].

We have curated a comprehensive set of fungal genes encoding GH proteins for which biochemical characterization has been reported in the peer-reviewed literature [290]. Of the 675 enzymes included in this database, 527 of them utilized recombinant proteins for their characterization (July 2013 version; https://mycoCLAP.fungalgenomics.ca). The commonly-used hosts in the production of recombinant GHs of fungal origin are the bacterium *Escherichia coli*, the yeasts *Pichia pastoris* and *Saccharomyces cerevisiae*, and the filamentous fungi *Aspergillus niger*, *A. oryzae* and *Trichoderma reesei*. While early studies used *E. coli* and *S. cerevisiae* as hosts for recombinant protein production because they are well-established laboratory organisms [291,292,293,294], *P. pastoris* has lately become a popular host for producing fungal GHs, particularly in medium-throughput platforms [286,287]. Nonetheless, *E. coli* remains the host of choice for studies in protein engineering and structure analysis because it is highly amenable to high-throughput screening [295,296,297]. Genes encoding GHs continue to be introduced into *S. cerevisiae*, often in the construction of consolidated bioprocessing strains for simultaneous saccharification and fermentation [298,299]. *Aspergillus niger* is widely used for the industrial production of extracellular enzymes [300,301,302]. Phylogenetically, it is closer to the biomass-degrading fungi than the other commonly used protein-production hosts. For these reasons, *A. niger* is used by the major enzyme manufacturers and many research groups to produce recombinant GHs for biochemical analysis [289,300,301,303,304,305,306] . Moreover, there is currently a significant surge of interest in producing GHs in plants to promote autohydrolysis of biomass [307,308,309].

In almost all studies involving recombinant GHs, a single host is used to produce the protein encoded by a target gene. Where a target gene is expressed in more than one host, the levels of protein production in the different hosts are usually reported rather than their biochemical properties [310,311,312,313,314,315,316]. In the few studies where comparative information is available, the recombinant enzymes produced in different hosts displayed dissimilar temperature and pH optima [317,308]. If different

hosts produce recombinant GHs that are biochemically dissimilar, it may not be possible to reliably extrapolate biochemical and structural data obtained from recombinant proteins to those of the native enzymes or to other expression hosts.

To compare the production efficiency of recombinant GH proteins and their biochemical properties in commonly used expression hosts, we have expressed 24 genes encoding extracellular GHs of *A. niger* in *A. niger*, *E. coli*, *P. pastoris*, and in a *Nicotiana benthamiana,* transient system. The recombinant GHs from these four hosts were fully or partially purified, and their pH, temperature optima and specific activity were determined. The results show that the active enzymes production rates in *E. coli* and *P. pastoris* are lower than those for *A. niger* and *N. benthamiana*. In general, proteins produced in the four production hosts exhibit similar biochemical properties, with some notable exceptions.

2.3 Results

Success rate of producing active, recombinant enzymes varies with expression hosts

We attempted to express 24 *A. niger* genes encoding previously characterized GHs in each of four protein production host organisms. The enzymes encoded by these 24 genes encompass the major activities involved in the decomposition of plant-derived biomass. We focused on characterizing lignocellulolytic enzymes, hence some of the genes encoding starch-degrading enzymes were not transformed into all four hosts. Further, genes that we could not in the first attempt amplify, clone or transform were not pursued further. We used SDS-PAGE followed by Coomassie Blue staining as the method of screening for protein production in *E. coli*, *P. pastoris* and *A. niger*. Therefore, only recombinant proteins produced at levels that could be visualized unambiguously were further characterized. Immunoblotting was used to determine protein production in *N. benthamiana*, which also allowed us to quantify the levels of protein production

(data not shown). Figure 9 shows that recombinant proteins were produced in *N. benthamiana* at 0.2 % to 13 % of total soluble proteins.

Of the successful transformations, there are four basic outcomes: 1) recombinant protein is produced and the enzyme is biochemically active; 2) recombinant protein is produced but it is not enzymatically active; 3) no recombinant protein is produced; and 4) in the case of *E. coli*, the recombinant protein produced is insoluble. Table 3 shows the results of protein production and enzymatic activity screening of the positive transformants of the four hosts. Of the 19 GH genes transformed into *E. coli*, 16 of them produced proteins. However 12 of these recombinant proteins were insoluble, resulting in four active enzymes; a success rate of 21%. Nine out of 16 genes transformed into *P. pastoris* yielded recombinant proteins. However one of them, Xeg12A xyloglucanase, did not exhibit enzyme activity. The overall success rate of obtaining active enzymes for *P. pastoris* is 50%. In contrast, about 88% of the genes transformed into *A. niger* (14 out of 16) and 91% of the genes transformed into *N. benthamiana* (20 out of 22) yielded active enzymes.

The four active enzymes produced in *E. coli* have molecular weights lower than 49 kDa. For *A. niger* and *N. benthamiana*, enzymes of different sizes were actively produced. Further, for the proteins produced in *N. benthamiana*, there does not appear to be a correlation between levels of production and length of the polypeptides (Table 3 and Figure 9). The levels of production for different recombinant enzymes in *A. niger*, *E. coli* and *P. pastoris* also vary (data not shown). Levels of protein production in *A. niger* and *P. pastoris* can vary significantly depending on culture conditions and genetic background of the host organisms. Since the main purpose of this study is to compare biochemical properties, we made no attempt to optimize protein production levels.

Table 3. Production and activity screening of recombinant proteins

Protein	Enzyme Activity	Predicted molecular weights (kDa)	Production host			
			Ani	Eco	Ppa	Nbe
Abf54A	arabinofuranosidase	50.7				
Abn43A	arabinanase	32.5				
Agl31A	alpha-glucosidase	106.4				
Agu67A	alpha-glucuronidase	91.6				
Axh62A	arabinoxylan-arabinofuranhydrolase	33.1				
Bgl3A	beta-glucosidase	91.3				
Cbh7A	cellobiohydrolase	48.3				
Cbh7B	cellobiohydrolase	56.2				
Egl12A	endoglucanase	24.3				
Egl5B	endoglucanase	34.8				
Inu32A	endo-inulinase	53.1				
Inx32E	exo-inulinase	57.3				
Lac35A	beta-galactosidase	107.5				
Mel27A	alpha-galactosidase	60.1				
Mel27B	alpha-galactosidase	48.8				
Mnd2A	beta-mannosidase	102.3				
Pga28A	endo-polygalacturonase	35.5				
Pga28E	endo-polygalacturonase	35.6				
Pga28II	endo-polygalacturonase	35.0				
Pgx28C	exo-polygalacturonase	45.8				
Xeg12A	xyloglucanase	23.3				
Xyl3D	xylosidase	84.7				
Xyn10A	xylanase	32.7				
Xyn11B	xylanase	20.1				

Blank cells indicate occasions where gene expression was not attempted; solid grey, active enzyme produced; horizontal strips, protein was produced but with no enzyme activity; vertical stripes, no protein production; trellis, the protein produced was insoluble. Ani, *Aspergillus niger*; Eco, *Escherichia coli*; Ppa, *Pichia pastoris*; Nbe, *Nicotiana benthamiana*.

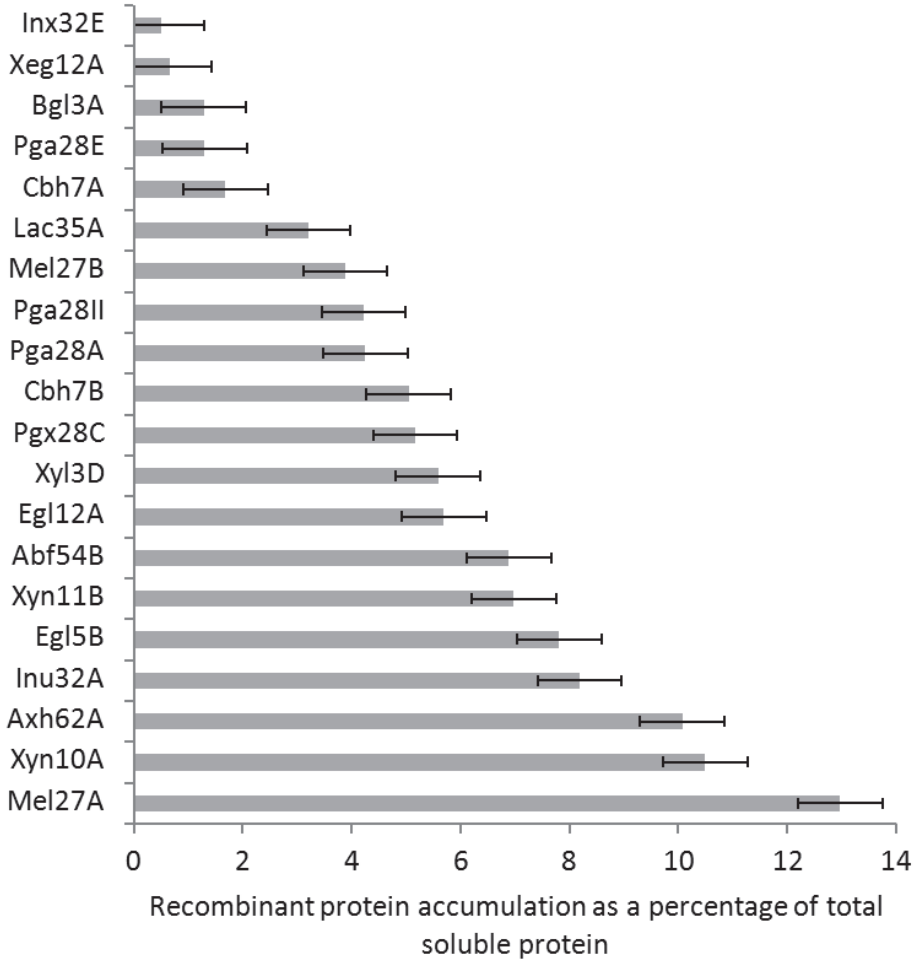


Figure 9. Levels of recombinant protein production in *N. benthamiana*. The levels were quantified in 10 replicate plants by immunoblotting using an anti-c-myc monoclonal antibody. Error bars represent the standard error of the mean

Comparison of biochemical properties

Proteins showing their expected activities in the enzyme screens were further characterized for pH and temperature profiles, Figure 10 and Figure 11 respectively. Specific activity was also examined for those proteins which we were able to purify sufficiently for unambiguous quantification (Figure 12). Table 4 summarizes the specific activities, temperature optima and pH optima for the active enzymes. Since we did not obtain active enzymes from *agl31A* and *agu67A*, the properties of only 22 enzymes are shown.

The pH optimum, temperature optimum and specific activity of recombinant enzymes produced in *N. benthamiana* are nearly identical to those produced in the endogenous *A. niger* host, although Xyn11B expressed in *N. benthamiana* had approximately 3-fold lower specific activity than the enzyme expressed in *A. niger.* On the other hand, Xyn10A, Egl12A and Pgx28C produced in *N. benthamiana* have about 2-fold higher specific activity than the enzymes produced in *A. niger* (Table 4). It should be noted that the specific activity was determined for most enzymes at pH 5.0, not necessarily the pH optimum of the recombinant enzymes. Slight differences in pH optimum can change the specific activity by 2 – 3 folds.

Two xylanases, Xyn10A and Xyn11B, produced in *E. coli* have pH optima higher than those produced in the other three hosts (Table 4; Figure 10). Despite measuring the activity at their optimal pH, these two enzymes displayed 5 – 10 times lower specific activity than the enzymes produced in *A. niger* (Table 4).

Enzymes produced in *P. pastoris* have pH profiles essentially indistinguishable from those produced in *A. niger* and *N. benthamiana* (Figure 10). However, the endo-polygalacturonase Pga28II has lower specific activity, as do the xylanases Xyn10A and Xyn11B. Pga28A and Pga28II produced in *P. pastoris* are temperature labile in the standard assay of pH 5.0 (data not shown). The temperature profiles for these two enzymes shown in Table 4 and Figure 11 were determined at the optimal pH of 4.0.

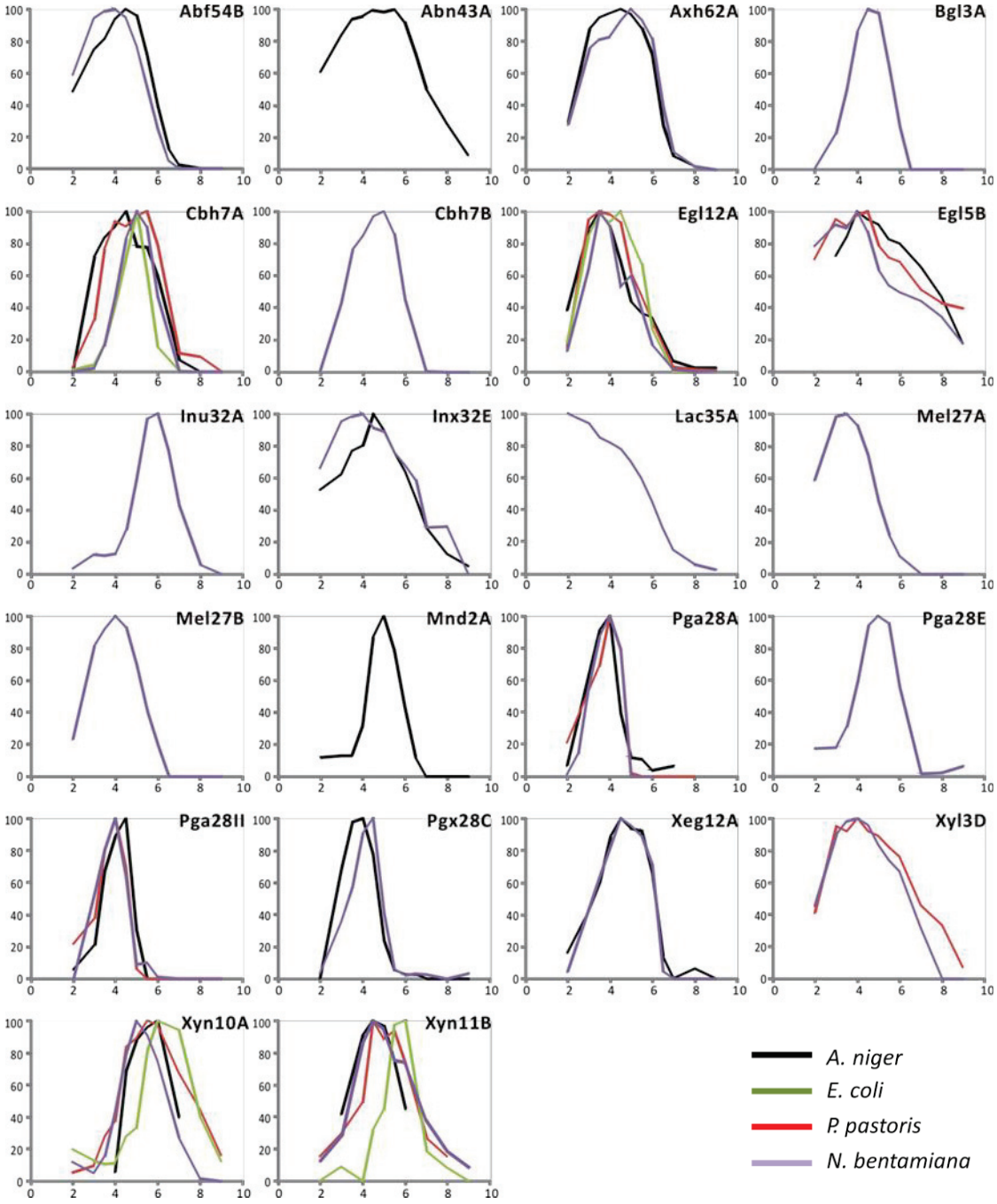


Figure 10. pH profiles of recombinant enzymes. The enzymes produced in the four expression hosts were assayed at different pH at 40 °C: x-axis, pH values; y-axis, relative activity as percentage of peak activity

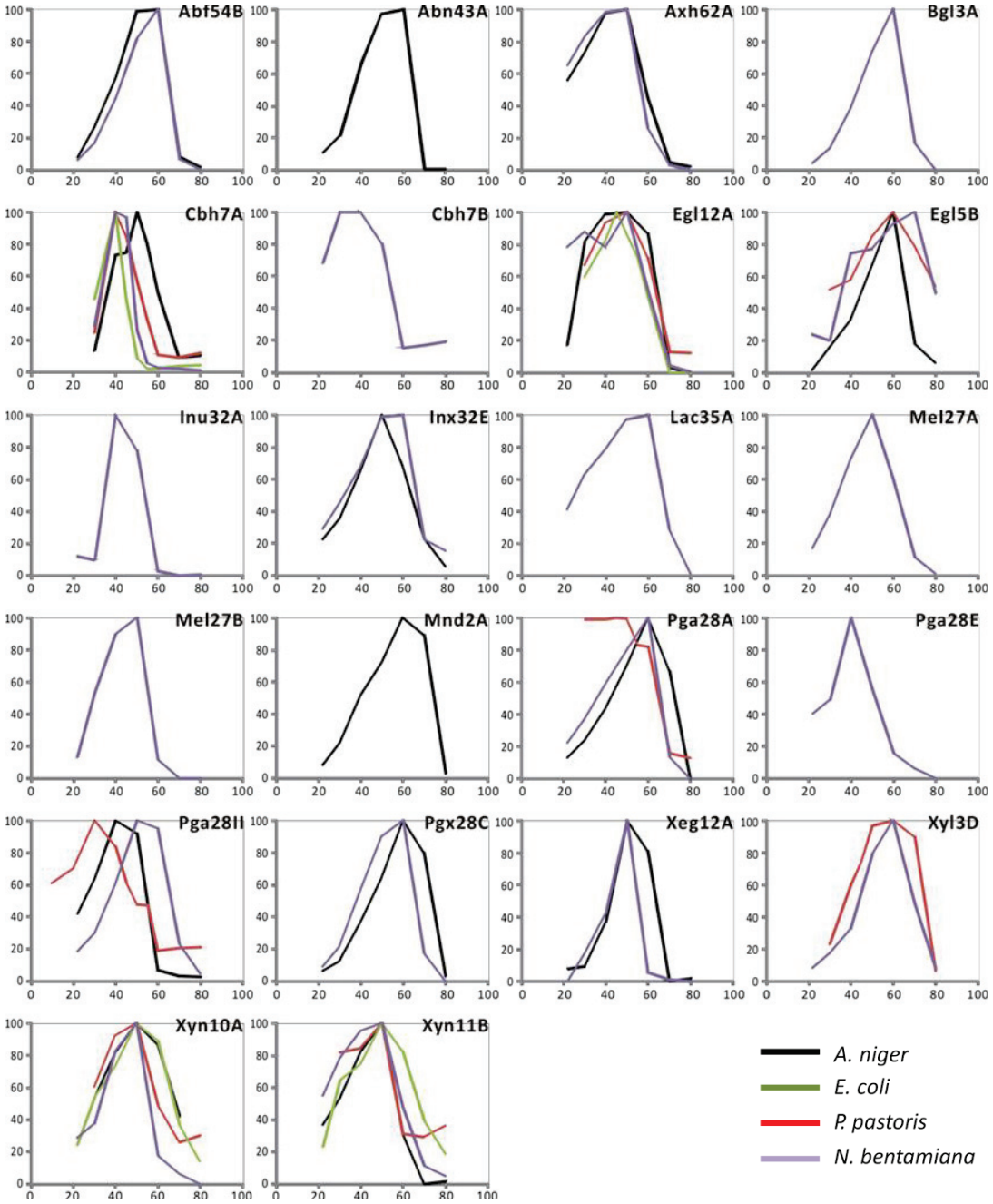


Figure 11. Temperature profiles of recombinant enzymes. The enzymes produced in the four expression hosts were assayed at different temperatures at pH 5: x-axis, temperatures at °C; y-axis, relative activity as percentage of peak activity

Table 4. Biochemical properties of recombinant glycoside hydrolases produced by four expression hosts

		Temperature optimum, °C				pH optimum				Specific activity, U/mg protein			
Protein	Enzyme Activity	Ani	Eco	Ppa	Nbe	Ani	Eco	Ppa	Nbe	Ani	Eco	Ppa	Nbe
Abf54B	arabinofuranosidase	60			60	3.5-5.0			3.0-4.5	88±1			47±2
Abn43A	arabinanase	60				3.0-6.0				77±2			
Axh62A	arabinoxylan-arabinofuranhydrolase	50			50	3.0-5.5			3.5-6.0	53±3			
Bgl3A	β-glucosidase				60				4.0-5.0				50±3
Cbh7A	cellobiohydrolase	50	40	40	40	3.5-5.5	5.5	4.0-5.5	4.5-5.5			33±2	
Cbh7B	cellobiohydrolase				40				4.0-5.5				
Egl12A	endoglucanase	50	50	50	50	3.0-4.0	3.0-5.0	3.0-4.5	3.5-4.0	44±2	62±1	45±2	100±4
Egl5B	endoglucanase	60		60	70	3.5-6.0		3.0-4.5	3.0-4.5	180±5		30±1	150±6
Inu32A	endo-inulinase				40				5.5-6.0				68±3
Inx32E	exo-inulinase	50			50	4.5-5.0			3.0-5.0	98±4			
Lac35A	β-galactosidase				60				≤2.0				
Mel27A	α-galactosidase				50				3.0-4.0				
Mel27B	α -galactosidase				50				3.0-4.5				
Mnd2A	β-mannosidase	60				4.5-5.0							
Pga28A	endo-polygalacturonase	60		60[a]	60	3.5-4.0		4.0-4.5	4.0-4.5	2800±23[a]		2600±31[a]	
Pga28E	endo-polygalacturonase				40				4.5-5.5				
Pga28II	endo-polygalacturonase	40		40[a]	50	4.0-4.5		3.5-4.0	3.5-4.0	11000±57[b]		2300±36[a]	
Pgx28C	exo-polygalacturonase	60			60	3.5-4.0			4.0-4.5	120±4			215±12
Xeg12A	xyloglucanase	50			50	4.0-5.5			4.0-5.5	185±5			
Xyl3D	xylosidase			60	60			3.0-5.5	3.0-5.0			31±1	
Xyn10A	xylanase	50	50	50	50	5.0-6.0	5.5-7.0	4.5-6.0	4.5-5.5	230±6	21±2[c]	76±3	400±15
Xyn11B	xylanase	50	50	50	50	4.0-5.0	5.5-6.0	4.5-5.5	4.0-5.0	1600±24	300±11[c]	170±7	450±13

Optimal pH is presented in a range of pH where the enzyme activity is ≥80% of peak activity. Some of the enzymes displayed low activity at the standard assay conditions of pH 5.0. Therefore their specific activities were determined at their optimal pH: [a], assays performed at pH 4.0; [b], assays performed at pH 4.5; [c], assays performed at pH 6.0.

Protein glycosylation

Secreted proteins of eukaryotes are often modified post-translationally, with glycosylation being the dominant modification. The glycosylation sites predicted by sequence analysis and those determined by mass spectrometric analyses [306,318] are summarized in Table 5. To determine whether or not differences in protein glycosylation can explain the dissimilar biochemical properties among the recombinant proteins, we have examined, in addition prediction and previous mass spectrometric characterization, glycosylation of the target enzymes by staining and enzymatic cleavage. The results are summarized in Table 5 and Figure 12. In almost all cases, the results from glyco-staining and PNGaseF digestion correspond positively with prediction and previous characterization. The enzyme Pga28II possesses one predicted N-glycosylation site. The Pga28II protein produced in *P. pastoris* stained positively with the glycol-stain, but the one produced in *A. niger* did not. This result suggests that the Pga28II produced in *A. niger* was either not glycosylated or glycosylated at such low level that it cannot be detected by glycol-staining. Importantly, the xylanases Xyn10A and Xyn11B have no predicted glycosylation sites and the proteins produced in the four hosts are not glycosylated. Therefore, the glycosylation pattern cannot be used to explain the differences in biochemical properties for the xylanases produced in *E. coli* and *P. pastoris*.

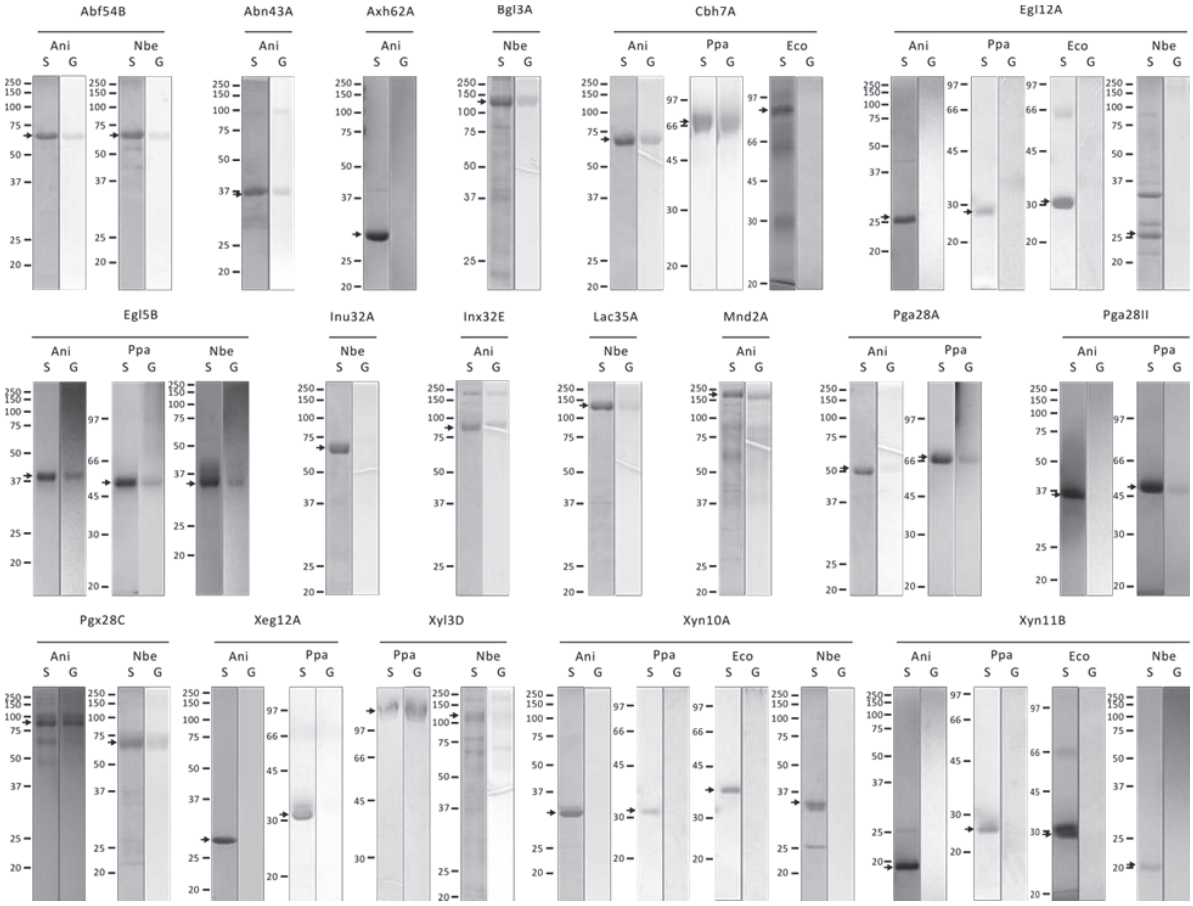


Figure 12. SDS-PAGE and glyco-staining of partially purified recombinant enzymes. Partially purified proteins were stained with Coomassie Blue (left lanes) with the staining intensities used to determine protein concentrations. Proteins in duplicate SDS gels were transferred onto nitrocellulose membranes and stained with Pierce® Glycoprotein staining kit. S, Coomassie Blue staining of protein; G, staining of glycoprotein

Table 5. Summary of protein glycosylation

	N-glycosylation sites		Production host			
Protein	**Predicted**	**Determined**	**Ani**	**Eco**	**Ppa**	**Nbe**
Abf54A	2		■			■
Abn43A	1		■			
Axh62A	0		▨			
Bgl3A	11	9[a]				■
Cbh7A	1		■	▨	■	■
Egl12A	0		▨	▨	▨	▨
Egl5B	3		■		■	■
Inu32A	4					■
Inx32E	5		■			
Lac35A	11					■
Mel27B	7	3[a]				■
Mnd2A	11		■			
Pga28A	1		■		■	
Pga28II	1	1[a]	▨		■	
Pgx28C	7		■			■
Xeg12A	1	0[b]	▨		▨	▨
Xyl3D	10	1[a]			■	■
Xyn10A	0		▨	▨	▨	▨
Xyn11B	0		▨	▨	▨	▨

For Cbh7A and Mel27B produced in *N. benthamiana*, evidence for glycosylation is based on results of PNGaseF cleavage. Blank cells indicate instances where the enzymes were not tested; solid grey, protein glycosylated; diagonal stripes, glycosylation was not detected. Prediction of N-glycosylation sites was based on the detection of the sequence motif Asn-Xaa-Ser/Thr. The N-glycosylation sites of some of the target enzymes had previously been investigated by mass spectrometry. The number of sites experimentally determined by these studies is shown: [a], Wang et al. [318]; and [b], Master et al. [306]. Ani, *Aspergillus niger*; Eco, *Escherichia coli*; Ppa, *Pichia pastoris*; Nbe, *Nicotiana benthamiana*.

2.4 Discussion

Prospecting efficient GHs for the sustainable development of the future bioeconomy is a prime motivation of sequencing the genomes of biomass-degrading fungi [117,118,282,283]. With rapidly-growing numbers of sequenced fungal genomes, there is an unprecedented opportunity to identify new fungal catalysts for industrial and environmental processes. To capture this opportunity effectively, medium- to high-throughput platforms that can produce active proteins for biochemical characterization are required to screen tens of thousands of enzyme-encoding genes. Furthermore the properties of enzymes produced by the screening platforms should be similar to those of the native enzymes or those produced in industrial cell factories.

Our initial goal was to evaluate the efficiency and suitability of commonly used hosts for expressing fungal genes encoding biomass-degrading enzymes. For comparative analysis we chose genes encoding GHs from *A. niger* because: 1) a high-quality genome sequence is available [282], providing reliable gene sequences as templates; and 2) more GHs have been characterized from *A. niger* than from any other fungus [290], offering a spectrum of enzymes for comparative analyses. Although all the genes chosen for this study have been characterized previously, very limited biochemical characterization had been performed on most of the enzymes that they encode (Supplementary Table 1). In this study, we have provided additional biochemical data for 17 of the enzymes (Table 4).

Comparing biochemical properties of enzymes conducted in different studies is extremely difficult because many parameters in the assay conditions can have dramatic effects on activity. Factors that can influence biochemical properties include temperature, pH, substrate, buffer, method of detecting the enzyme products, and reaction time. To facilitate comparison, we attempted to assay the enzymes at the same pH of 5.0 and at 40 °C. These conditions provided reasonably meaningful comparison for most, but not all, enzymes produced in different host organisms. For example, some of the enzymes (Pga28A, Pga28II, Xyn10A and Xyn11B) produced in *E. coli* and *P. pastoris*

are sensitive to minor changes in pH, and they needed to be assayed at their optimal pHs to obtain specific activity data that are several-fold higher than at pH 5.0. Where data are available, the pH optima of the enzymes produced in *A. niger* reported in this study are essentially the same as those reported in other studies (Table 4 and Supplementary Table 1). Comparing the enzymes produced in *A. niger* with those of previous studies, the specific activity can vary by 3-fold (Pga28A, Pga28II and Xyn11B) and temperature optimum by 10 °C (Inx32E, Mnd2A and Xyn11B). We think that most of these differences are caused by the variations in the conditions used to determine the biochemical properties.

The four expression hosts that we chose are evolutionarily diverse, commonly used, and they have the potential of producing heterologous proteins in medium- and high-throughput platforms. *Aspergillus niger* was chosen as one of the expression hosts to provide a native environment to generate the recombinant enzymes as references. As well, *A. niger* is an industrial cell factory widely used for the production of GHs. The inability to secrete proteins efficiently and to glycosylate proteins should rule out *E. coli* as a suitable host for the production of extracellular GHs. However, protein structures including those of GHs are often determined using proteins produced in *E. coli* [319,320]. Owing to the relative ease of genetic manipulation, *P. pastoris* has gained a reputation as a suitable host for producing recombinant, extracellular proteins [321]. It has been used for the medium-throughput production of GHs from several filamentous fungi [176,288,322]. *Nicotiana benthamiana,* a close relative of tobacco, and tobacco have been routinely used for the production of heterologous proteins including GHs [308,309,323,324]. The four production platforms compared in this study have the potential for large-scale and inexpensive production of industrial enzymes.

Variations in biochemical properties for secreted enzymes produced in heterologous hosts are often attributed to alterations in protein glycosylation [325]. The xylanases produced in *E. coli* and *P. pastoris* displayed lower specific activity, and in the case of the *E. coli*-produced enzymes, higher pH optimum. Since there is neither

predicted nor detected glycosylation, the dissimilar properties are caused by other unknown factors. We cannot rule out the possibility that the presence of the various tags fused to these recombinant proteins might have altered the properties. Another possibility is that there are differences in post-translational processing. Previously, when expressed in *A. niger*, Xyn11B was reported to have 19 amino acids removed downstream of the predicted signal peptide, resulting in an N-terminal sequence of STPSST [326]. Although also produced in *P. pastoris*, the amino terminal sequence of the protein was not reported. An almost identical Xyn10 orthologue from *A. kawachii* was previously produced by the native host and shown to have an N-terminal of QASVSIDS, where the N-terminal glutamine was converted to pyroglutamate: thus the six residue sequence, EPIEPR, after the signal protease cleavage site was removed post-translationally [327]. It is possible that the same types of processing do not occur in *P. pastoris*, and very unlikely that they do in *E. coli*.

Almost all previous studies on protein production report only successful attempts. A review of protein production in *E. coli*, however, showed that this organism is poor at producing large proteins in a soluble form [328]. Our finding that all four active enzymes produced in *E. coli* have molecular weights lower than 39 kDa (Table 3) is consistent with previous observations. The 88% success rate of obtaining active enzymes in *A. niger* is very high by comparison, but this is likely because we were expressing *A. niger* genes in their native host. In the past three years, we have attempted to express over 4,000 genes predicted to encode biomass-degrading enzymes from 40 fungal species in *A. niger*. The success rate varies from 10 % to 40 % with an overall rate of ~30 % (unpublished data). Thus the 50 % success rate of producing active heterologous proteins in *P. pastoris* shown in this study (Table 3) is comparable to, if not higher than, that of *A. niger*. The surprising finding is the ~90 % success rate of producing active, heterologous enzymes in *N. benthamiana*. With the recent interest in producing heterologous biomass-degrading enzymes in plants [221],

the successes in producing active enzymes in *N. benthamiana* reported in this study should stoke further investigations.

2.5 Conclusions

The data presented here indicate that *N. benthamiana* is an effective production host for fungal glycoside hydrolases. Based on the present study and our unpublished results, *P. pastoris* and *A. niger* have similar success rates of producing heterologous GHs. Xylanases produced in *E. coli* and *P. pastoris* as well as polygalacturonases produced in *P. pastoris* display biochemical properties that are substantially different from those produced in *A. niger*. The differences in biochemical properties in the xylanases are unlikely to be due to differences in protein glycosylation as the native enzymes and the recombinant proteins are not glycosylated.

2.6 Methods

Chemicals and materials

Substrates used for enzyme assays were purchased from Sigma-Aldrich (Oakville, Canada), Megazyme International Ireland Co. (Wicklow, Ireland), and Glycosynth (Warrington, UK). Restriction and DNA modification enzymes were purchased from New England BioLabs (Ipswich, MA, U.S.A.). MonoQ and HisTrapFF chromatography columns were from GE Healthcare Lifesciences (Baie d'Urfe, QC, Canada). Molecular weight markers were obtained from GE Healthcare Lifesciences or BioRad (Mississauga, ON, Canada). Vivaspin ultrafiltration devices (10kDa MWCO) with polyethersulfone membranes were obtained from Sartorius Stedim Biotech *(*Goettingen, Germany).

Genes used in this study

The genes used in this study have previously been identified, and in some cases the corresponding enzyme has been produced in a microbial host and characterized, as summarized in Supplementary Table 1. Sequences of the target genes were obtained from the *A. niger* genome sequencing portal (genome.jgi-psf.org/Aspni5/Aspni5.home.html). Protein sequences were analyzed using the SignalP [329] algorithm to identify the signal peptides.

Source of DNA for cloning

PCR products were generated from cDNA libraries of *A. niger* strain N400 or from *A. niger* strain ATCC 11414 genomic DNA. Genes with multiple exons were obtained by amplification and reassembly of the exon fragments using the SPLICE technique [330]. DNA manipulations were carried out using standard techniques [331].

Cloning and expression in *E. coli*

Primers were designed [332] to exclude the N-terminal signal peptide sequences predicted for the native proteins and with the ligation-independent cloning sequences appended to the 5' ends of forward and reverse primers (Supplementary Table 2). Amplified coding regions were cloned into expression vectors as described previously [333]. Figure 13 illustrates the target proteins and their flanking regions of the constructs used in the four expression hosts to produce the target proteins. Expression plasmids were transformed into BL21(DE3) Magic or T7 Shuffle Express (NEB) *E. coli* expression strains. Positive clones were selected from small-scale expression cultures induced with 1 mM IPTG. Whole cell samples and cleared lysates were assayed for fusion protein production and solubility on SDS-PAGE gels stained with Coomassie Brilliant Blue. Positively cloned DNA was amplified using Templiphi Sequencing

Amplification kits (GE Healthcare) and sequenced to confirm that the coding regions were properly integrated.

For larger scale protein production, 10 mL of overnight cultures in LB containing ampicillin (100 µg/mL) or chloramphenicol (34 µg/mL) were diluted in 1 L of fresh medium in 2.8 L Fernbach flasks. Cells were grown at 30 °C until the OD_{600} reached between 0.6 and 0.9. Cells were cooled on ice for 10 minutes, induced with 0.5 M IPTG, and incubated at 20 °C overnight. Cells were pelleted (4000 × *g*, 10 min), lysed by ultrasonication, centrifuged (4000 × *g*, 30 min), and the soluble fraction was recovered for further use.

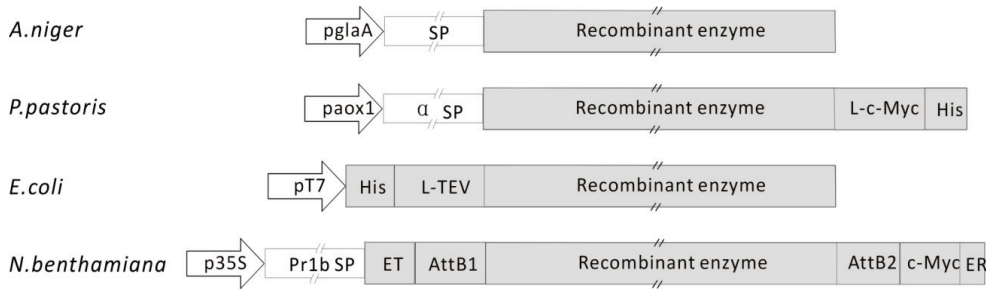


Figure 13. Schematic representation of constructs for recombinant protein production. The mature protein regions are shown as shaded boxes. The component parts of the expression constructs in *A. niger* are: pglaA, glucoamylase gene promoter; SP, signal peptide of target enzymes. For *P. pastoris*, the parts are: paox1, alcohol oxidase 1 promoter; α SP, α-factor signal; L-c-Myc, linker containing c-Myc tag; His, 6xHistidine tag. For *E. coli*, the parts are: pT7, RNA polymerase promoter from Bacteriophage T7; His, 6xHistidine tag; L-TEV, linker containing TEV cleavage site. For *N. benthamiana*, the parts are: p35S, double enhanced cauliflower mosaic virus 35S promoter; Pr1b SP, secretory signal peptide; ET, express tag; AttB1 and AttB2, Gateway recombination attachment sites; c-Myc, c-Myc tag; ER, ER-retrieval tetrapeptide

Cloning and expression in *P. pastoris*

Cloning into *P. pastoris* expression vectors was accomplished as described in the *Pichia* EasySelect Manual (Invitrogen). Coding regions of the target genes were analyzed for restriction sites to determine the suitable restriction enzymes to be used for cloning. Primers were designed using Clone Manager (Sci-Ed Software) and appended with *Xho*I or *Sal*I sites at the 5' end and *Xba*I or *Nhe*I sites at the 3' end (Supplementary Table 3). Predicted N-terminal signal peptide sequences were excluded. All targets were designed to be expressed in-frame with both the N-terminal α-mating factor secretion signal and the C-terminal C-myc and hexahistidine tags contained in the expression vector (Figure 13). A Kex2 protease site for signal peptide cleavage was also added at the N-terminus of the target. Amplified fragments were digested with appropriate restriction enzymes and ligated into the pPICZαC *Pichia* expression vector (Invitrogen) digested with *Xho*I and *Xba*I restriction enzymes. Ligation reactions were transformed into NEB Turbo *E. coli* and positive clones were identified by colony PCR and sequence-verified to confirm gene integration. Plasmid DNA was purified from bacterial cells using Qiagen Miniprep kits and subsequenlty linearized with *Sac*I, *BstX*I, or *Pme*I restriction enzymes. One microgram of linearized DNA was transformed into wild-type X-33 *Pichia* cells (Invitrogen) as previously described [334]. Colonies were screened for multicopy integrants by patch-plating transformants onto media containing 1 µg/mL Zeocin (Invitrogen). Strains that grew on high-selection plates were assayed for expression in 5 mL of buffered rich media cultures induced with 0.5 % methanol (BMMY media) for 2 days at 30 °C. Culture supernatants were run on SDS-PAGE gels and stained with Coomassie Brilliant Blue to identify clones producing fusion proteins.

Clones that secreted fusion proteins in small-scale cultures were used for large-scale protein production. Single colonies were used to inoculate 50 – 200 mL starter cultures in BMYG medium which were grown overnight. Cultures were then pelleted and resuspended in different volumes of BMMY medium, as described in either of the two methods in the *Pichia* EasySelect Manual (Invitrogen), and grown at 30 °C for an

additional 1 – 3 days with daily additions of methanol as described. After an appropriate induction time, cell cultures were clarified by centrifugation and proteins were either precipitated using 70% ammonium sulfate prior to purification or applied directly to chromatography columns.

Cloning and expression in *A. niger*

For expression in *A. niger*, genes were amplified by PCR using the primers shown in Supplementary Table 4. The native signal peptide was included in the amplified sequence (Figure 13). One-step Gateway recombinational cloning was performed to transfer the amplified fragment to ANIp7G, a Gateway-compatible integrative expression vector [306,335]. Protoplasts of *A. niger* strain N593 *glaA*::*hisG* were transformed with recombinant plasmids by following the methods described by Debets [336]. Transformants were selected on minimal medium without uracil and uridine [337]. Supernatants from individual transformants were screened for recombinant protein production after growth in MMJ [306] medium, and those exhibiting the highest level were selected for further study.

Cloning and expression in *N. benthamiana*

A binary vector for protein production in plant was constructed by routine cloning procedures [331]. This vector was modified from pCaMterX [338] to include a tCUP translational enhancer [339], the Pr1b signal peptide [340], and the Xpress tag (DLYDDDK) at the N-terminus, and the c-Myc tag (EQKLISEEDL) and the ER retention signal KDEL at the C-terminus (Figure 13). A Gateway destination vector was then produced with the 'Gateway Vector Conversion System' (Invitrogen) according to the manufacturer's instructions. The entry clones were generated by a Gateway BP recombination reaction between the pDONR™/Zeo vector (Invitrogen) and the targeted genes which were PCR-amplified with oligonucleotide primers containing attB1 and attB2 recombination sites (Supplementary Table 5). The desired expression clones were

then generated by performing a Gateway LR recombination reaction between the entry clones and the plant binary Gateway destination vector. The expression clones were electroporated into *Agrobacterium tumefaciens* strain EHA105 [341] and then used for plant transformation.

For transient expression, the *Agrobacterium* suspensions were infiltrated into the intact leaves of 5- to 6-week-old *N. benthamiana* plants as previously described [342]. Briefly, the induced *Agrobacterium* suspensions were adjusted to a final OD_{600} of 1.0 and then co-infiltrated with equal amounts of an *Agrobacterium* suspension carrying a p19 suppressor of posttranscriptional gene silencing [343] into the intercellular spaces of leaves using a 1-mL syringe. To account for the variability between plants, leaves, and the position on the leaf, comparably sized leaves from ten different plants of similar age were systematically agro-infiltrated for each expression construct. After infiltration, the plants were maintained for 4 days in a controlled growth chamber at 22 °C with a 16 h photoperiod. Tissue samples from the individually infiltrated leaf panels served as biological replicates and were analyzed separately, with the average of the ten panels used to represent the concentration of a given recombinant protein.

For each leaf sample, total soluble protein (TSP) was extracted from four 7 mm leaf discs (approximate fresh weight of 25 mg) by homogenization with a Mixer Mill MM 300 (Retsch, Haan, Germany). The resulting frozen powdered leaves were resuspended at 4 °C in 300 µL of extraction buffer [phosphate-buffered saline (PBS), pH 7.4, 0.1% Tween-20, 2% polyvinylpolypyrrolidone (PVPP), 1 mM ethylenediaminetetraacetic acid (EDTA), 100 mM ascorbic acid, 1 mM phenylmethylsulphonylfluoride (PMSF) and 1 µg/mL leupeptin]. The homogenate was clarified twice by centrifugation at 20 000 x*g* for 10 min at 4 °C. The TSP concentration was measured according to the method of Bradford using the Bio-Rad reagent (Bio-Rad, Hercules, CA, USA) with bovine serum albumin as a standard [344].

For enzymatic activity determination and biochemical characterization, ten grams of leaf material were agro-infiltrated with a syringe, harvested 4 days post

infiltration and flash frozen in liquid nitrogen. Samples were extracted in 3 volumes of extraction buffer (v/w) and treated as described above.

Purification of recombinant proteins

Recombinant His-tagged proteins from *P. pastoris* or *E. coli* were purified using nickel affinity column chromatography with 1 mL HisTrapFF columns as specified by the manufacturer. Briefly, the column was first equilibrated with the binding buffer (20 mM sodium phosphate, 30 mM imidazole, 0.5 M sodium chloride, pH 7.4). Filtered (0.22 µm membrane) crude samples adjusted with equilibration buffer were loaded onto the column. After extensive washing with equilibration buffer, the bound proteins were eluted with elution buffer (20 mM sodium phosphate, 200 mM imidazole, 0.5 M sodium chloride, pH 7.4) and concentrated using Vivaspin ultrafiltration devices (10 kDa MWCO). Purified proteins were stored in 50 mM sodium acetate buffer, pH 5.0.

Recombinant proteins produced in *A. niger* or *N. benthamiana* were purified by anion exchange and gel filtration chromatography using an ÄKTA chromatography system (Amersham Biosciences) at room temperature. Owing to their low concentrations, crude plant extracts containing recombinant proteins derived from *axh62A*, *cbh7A* and *pga28II* were used for biochemical characterization directly after buffer exchange. Prior to purification by column chromatography, supernatants from *A. niger* cultures were concentrated using a Vivaspin ultrafiltration device with a 10 kDa cut-off. Leaf extracts of *N. benthamiana* were centrifuged for 20 min at 20 000 x *g* and the supernatants were concentrated by ultrafiltration. The concentrated proteins were buffer-exchanged into 20 mM Tris-HCl buffer, pH 8.0, by repeated dilution and ultrafiltration, and applied to a MonoQ HR 5/10 anion exchange column equilibrated with the same buffer. Bound proteins were eluted using a linear 0 to 1 M KCl gradient in 20 mM Tris-HCl buffer, pH 8.0. Fractions (1 mL) were collected at a flow rate of 1 mL/min and stored on ice until assayed for enzyme activity. Recombinant proteins were

furthered purified by gel filtration column chromatography. Purified proteins were characterized immediately after purification or stored at -80 °C.

Enzyme activity assays

All reactions were carried out in triplicate. One unit of activity was defined as the amount of enzyme used to release 1 µmole of product per min. The substrates used for the enzyme assays are listed in Supplementary Table 6.

Reducing sugar assays were performed in 96-well microplate format as described elsewhere [345,346,347] with slight modifications in either 50 or 100 µL assay volumes. For the 50 µL assay, 10 µl of substrate (1 % stock concentration) was added to 30 µL of 50 mM Britton-Robinson buffer [348] (50 mM boric acid, 50 mM acetic acid and 50 mM phosphoric acid), pH 5.0, and the reaction initiated by the addition of 10 µL of appropriately diluted (in 10 mM Britton-Robinson buffer) enzyme. All volumes were doubled for the 100 µL assays. The reaction mixture was immediately incubated at 40 °C for 30 min. The reaction was stopped by incubating on ice. Ten microlitres of the reaction mixture were withdrawn and added to 190 µl of ice-cold BCA reagent and incubated at 80 °C for 40 min for colour development. Following incubation, 160 µl of the mixture were transferred to a flat bottom-microplate and the optical density was read at 562 nm. The corresponding monosaccharides making up the substrates were used to prepare standard curves.

Activities against *p*NP-derivatives were carried out as described previously [348] in 50 mM Britton-Robinson buffer, pH 5.0, in either 50 or 100 µL assay volumes. For the 50 µL assays, 10 µl of 5 mM substrate (stock concentration) was added to 30 µL of buffer, the reaction was initiated by addition of 10 µL of appropriately diluted enzyme followed by incubation at 40 °C for 30 min. All volumes were doubled in the 100 µL assays. The reaction was stopped by the addition of 50 µL of 1 M sodium carbonate, then 80 µL of reaction mixture were transferred to flat-bottom 96-well plates and the optical density was read at 410 nm and compared to a standard curve of *p*-nitrophenol.

To determine the pH profile of enzyme activity, reactions were performed at 40 °C in 50 mM Britton-Robinson buffer (pH 2.0 – 9.0) unless noted otherwise. The effect of temperature on enzyme activity was determined using 50 mM citrate buffer (pH 5.0) and incubating reaction mixtures at 22, 30, 40, 50, 60, 70 and 80 °C unless noted otherwise. Reactions were initiated by adding an appropriate dilution of enzyme and incubated for 30 min. Reactions were terminated and analyzed for reducing sugar or *p*-nitrophenol release, as appropriate.

Quantitative Western blot analysis of plant leaf extracts

Extracted plant proteins were resolved by SDS-PAGE and then transferred to nitrocellulose membranes by semi-dry electroblotting. Membranes were blocked with 1% Western Blocking Reagent (Roche, Mannheim, Germany) in Tris-buffered saline (TBS, 50 mM Tris, 150 mM NaCl, pH 7.5) overnight at 4 °C and incubated with a 1:2000 dilution of mouse anti-c-Myc[Biotin] monoclonal antibody (A00864; GenScript, Piscataway, NJ, USA) for 1 h at room temperature with gentle shaking. The primary antibody was detected with a 1:5000 dilution of HRP-conjugated goat anti-mouse IgG (170-6516; Bio-Rad) and visualized using an ECL kit (GE Healthcare), according to the manufacturer's instructions. Membranes were washed four times between each step with TBS containing 0.05% Tween-20, and all antibodies were diluted in TBS with 0.5 % Western Blocking Reagent. The concentration of recombinant protein was determined by using image densitometry with TotalLab TL100 software (Nonlinear Dynamics, Durham, USA). Band intensities were compared to lanes containing known amounts of a purified custom-made control protein (GenScript), which consisted of a cellulose-binding domain containing a c-myc tag.

Glycosylation analysis – staining and enzymatic deglycosylation

The presence of glycosylation on proteins produced in *A. niger*, *P. pastoris, E. coli* and *N. benthamiana* was detected using the Pierce Glycoprotein staining kit (Thermo Scientific) as directed by the manufacturer.

Since the plant-produced proteins were tagged and could be detected by immunoblotting, we also used enzymatic deglycosylation to detect protein glycosylation. Total plant protein extracts were deglycosylated with PNGaseF (New England Biolabs, Ipswich, MA, USA) for 24 h at 37 °C, according to the manufacturer's instructions. Control samples were treated the same, except that no PNGaseF was added. Finally, the samples were analyzed by SDS-PAGE and immunoblotted as described in the previous section.

Estimation of enzyme concentration

Enzymes used for characterization were quantified after electrophoresis on 12 % SDS-PAGE gels followed by Coomassie Blue staining [349]. Gels were scanned on a Syngene G: BOX Chemi system and the gel images were processed and analyzed using Syngene Gene Tools. The intensities of protein bands were compared with a series of BSA standards run on the same gel whose intensity values were used to generate a standard curve as described previously [350].

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was funded by Genome Canada, Génome Québec, Agriculture and Agri-Food Canada, the Bioconversion Network of the Natural Sciences and Engineering Research Council of Canada and the United States Department of Energy. We thank Annie Bellemare, Tieling Zhang and Marianne Desautels for producing the *A. niger* transformants.

A set of 15 *A. niger* biomass degrading proteins expressed in *P. pastoris* and *E. coli,* which were not reported in Manuscript I were biochemically characterized as reported in Manuscript II below.

Chapter 3 Manuscript II

Cloning, Expression and characterization of *Aspergillus niger* glycoside hydrolases in *P. pastoris* and *E. coli*

Enongene Ekwe and Justin Powlowski
Centre for Structural and Functional Genomics, Concordia University, Montreal, Quebec

Manuscript

3.1 Abstract

Plant cell wall degrading proteins, especially glycoside hydrolases, are key players in the biochemical conversion of biomass to fuels and high value chemicals. Efforts to exploit glycoside hydrolases from fungal saprotrophs have sometimes been challenging due to the fact that some of these proteins are not easily expressed in the parent host. The use of recombinant glycoside hydrolases as performance-enhancing supplements in cellulase mixtures has been reported. In some cases, the expression host has been found to affect the level of production and biochemical properties of the protein. Heterologous host systems such as *Pichia pastoris* and *Escherichia coli* can be used as industrial cell factories for the production of recombianant fungal glycoside hydrolases.

In this study, we report the expression and biochemical characterization of 15 *Aspergillus niger* glycoside hydrolases in *P. pastoris* and *E. coli*, most of which showed low level expression or were not expressed at all using the native gene or recombinant genes engineered for expression in the native host. All 15 proteins were successfully expressed in *P. pastoris* at higher production levels than the parent host, with only 38% of the biochemically characterized proteins showing the predicted activity. Egl5C and Man26A, which were not expressed in the parent host, were expressed in active form in

P. pastoris. Two proteins which were tested in *E. coli* were successfully expressed, one of which showed the predicted activity. The pH optimum, temperature optimum and specific activity of Xyn11G produced in both *P. pastoris* and *E. coli* were identical. All the five active proteins from *P. pastoris* showed pH and temperature optima of 3.5 – 5 and 40 – 70 °C, respectively. The endoglucanase Egl5C had a temperature optimum of 70 °C, which was at least 10°C higher than that of the four active proteins. There was no evidence to suggest that the activity of the proteins were influenced by glycosylation. With a low percentage of active proteins, the *P. pastoris* expression system used does not appear suitable for the expression of *A. niger* proteins.

Keywords: fungal glycoside hydrolases; enzyme production hosts; *Aspergillus niger*; *Escherichia coli*; *Pichia pastoris*; pH optimum; temperature optimum; specific activity

3.2 Introduction

Proteins that deconstruct plant cell walls are important in the development of renewable fuels and chemicals from lignocellulosic biomass. Lignocellulose is a renewable and abundant natural resource that can be sourced from forestry wastes and non-food dedicated energy crops, agricultural waste, industrial waste, sewage and municipal solid waste. Global annual production is estimated at 10^{10} MT, which makes it a credible resource for the sustainable production of renewable fuels and bio-based chemicals [21,22]. Increasing fossil fuel prices, and associated pollution and climate change concerns [4,5] have prompted the development of bio-based substitutes of fossil-based energy, fuels and products [351, 352].

For the production of bio-based fuels and high value chemicals through the biochemical transformation stream in biorefineries, the polysaccharides in lignocellulose must be broken down by enzymes into their monomeric sugars [72], which can be recovered or fermented by micro-organisms [73,74]. Lignocellulose in plant cell walls is

made up of the glucose homopolymer, cellulose, a meshwork of hemicellulosic polysaccharides and heterogeneous, polyaromatic lignin [12,13]. The structural complexity and heterogeneity [75,110] of plant cell walls necessitates a diverse enzyme repertoire for complete hydrolysis of its constituent carbohydrates to sugars [14,111]. Fungal saprotrophs are armed with genes that encode a variety of glycoside hydrolase (GH) proteins which hydrolyze glycosidic bonds, particularly the O-glycosyl hydrolases (EC 3.2.1.x) [353]. The core GHs required to convert cell wall carbohydrates to sugars include endoglucanases, cellobiohydrolases, β-glucosidase and xylanases [140,141,133].

Several physical and chemical pretreatment methods have been developed to modify or remove lignin, hydrolyse hemicellulose and increase enzyme accessibility to cellulose [16]. Challenges remain however, which include the non-productive adsorption of enzymes to lignin and other structural components [268] and the low hydrolysis rates imposed by the high crystallinity of cellulose [21,208].

Commercial lignocellulolytic enzymes cocktails are almost exclusively sourced from *T. reesei,* whose genome (http://genome.jgi.doe.gov/Trire2/Trire2.home.html) harbours the smallest number of genes encoding cell wall degrading proteins of all the sequenced lignocellulolytic fungi [117]. Other fungal saprotrophs such as *Aspergillus niger* [120,121,122] which carry a relatively higher number of glycoside hydrolase genes in their genome and secrete complex cocktails of plant cell wall degrading enzymes (genome.jgi psf.org/Aspni5/Aspni5.home.html) are a potential sources of new feedstock-specific enzymes for biomass conversion.

The use of specific enzyme components to optimize the core lignocellulolytic GHs present in commercial cellulase cocktails or to design novel efficient minimal enzyme cocktails has attracted significant interest in recent years [263]. Supplementation of lignocellulose degrading enzymes cocktails with accessory enzymes, mainly hemicellulases, pectinases [134,135], non-hydrolytic cellulase active proteins [136], as well fungal oxidative enzymes [137] has been found to ease cellulose accessibility via degradation of the surrounding non-cellulosic polymers [175,177,176].

Production and characterization of cell wall degrading enzymes from saprotrophs such as *A. niger* could identify enzymes which can be used in the optimization of cell wall degrading enzyme cocktails.

A large set of biochemically characterized fungal GH encoding genes, including those from *A. niger*, have been curated (https://mycoCLAP.fungalgenomics.ca) [290]. For many filamentous fungi, including *A. niger*, a significant proportion of the native proteins are poorly expressed or are not expressed at all using the parent host as the expression system. Recombinant fungal GH proteins have therefore been expressed in other host systems such as *E. coli* [295,296]. *P. pastoris* has emerged as a suitable host system for the production of eukaryotic proteins due to its ability to grow on minimal medium at very high cell densities (> 100 g dry weight per litre), secrete proteins into the growth medium for simplified downstream processing and perform eukaryotic post-translational modifications including protein folding, proteolytic processing, disulfide bridge formation and glycosylation [354]. *P. pastoris* has consequently been used in the production of several GH proteins from fungal saprotrophs [176,288,321,322,355]. In some cases, the production host has been found to affect the biochemical properties of the recombinant protein [325,355].

In this study, genes of cell wall degrading proteins were mined from the publicly available genome of *A. niger* strain ATCC 1015, cloned and heterologously expressed in *P. pastoris* and *E. coli*. In order to evaluate the effects of the expression host on the biochemical properties of the proteins, recombinant proteins were purified and biochemically characterized in terms of specific activity, pH and temperature optima and glycosylation differences. These enzymes were included in a library of purified and biochemically characterized fungal enzymes [355], which can potentially be useful in the design of efficient minimal or enhanced enzyme cocktails for the hydrolysis of high-consistency lignocellulosic biomass.

3.3 Materials and methods

Chemicals and materials

Substrates used for enzyme assays were purchased from Sigma-Aldrich (Oakville, Canada), Megazyme International Ireland Co. (Wicklow, Ireland) and Glycosynth (Warrington, UK). Restriction and DNA modification enzymes were purchased from New England BioLabs (Ipswich, MA, USA). MonoQ and HisTrapFF chromatography columns were from GE Healthcare Lifesciences (Baie d'Urfe, QC, Canada). Molecular weight markers were obtained from GE Healthcare Lifesciences or BioRad (Mississauga, ON, Canada). Vivaspin ultrafiltration devices (10kDa MWCO) with polyethersulfone membranes were obtained from Sartorius Stedim Biotech *(*Goettingen, Germany).

Purification of recombinant proteins

The gene constructs for proteins used in this study were produced as part of the study reported in reference [355]. Gene cloning and cell growth were as described therein. Recombinant His-tagged proteins from *P. pastoris* or *E. coli* were purified using nickel affinity column chromatography with 1 mL HisTrapFF columns as specified by the manufacturer (GE Healthcare). The protocols and buffers used in both cases were as described by Zhang *et al.* [355].

Enzyme activity assays

Reducing sugar assays and activities assays against *p*NP-derivatives were performed using a 96-well microplate format at 40 °C in 50 mM Britton-Robinson buffer as described by Zhang *et al.* [355]. The effect of temperature was assessed at pH 5.0 between 30 and 80 °C. The effect of pH was assessed at 40 °C between pH 2.0 – 9.0 [355]. All reactions were carried out in triplicate. One unit of activity was defined as the

amount of enzyme used to release 1 µmole of product per min. The substrates used for the enzyme assays are listed in Table 7.

Glycosylation analysis

Glycosylation was predicted by NetNGlyc v1.0 [356] and NetOGlyc v4.0 [357] tools, respectively of the Centre for Biological Sequence Analysis. The presence of glycosylation on proteins produced in *P. pastoris* and *E. coli* was detected using the Pierce Glycoprotein staining kit (Thermo Scientific) as directed by the manufacturer.

3.4 Results and Discussion

Glycosylation

Ten of the 15 proteins were previously tested for recombinant expression in the parent host, *A. niger* (Table 6 and Table 7). Of these, 5 (Bgl3N, Exg5, Xeg12, Xyn11G and Xyl43B) were successfully expressed and all but Exg5 showed the predicted activity. However, all fifteen proteins were successfully expressed in *P. pastoris* with expression levels exceeding 0.5 mg/ml for ten proteins. Of the four proteins (Egl5C, Exg5A, Man26A and Mog63) which did not express in *A. niger*, two (Egl5C and Man26A) showed the predicted activity when expressed in *P. pastoris* (Table 7).

Only two proteins, Xeg12 and Xyn11G, were attempted for expression in *E. coli*. Both were successfully expressed in soluble form, with the Xeg12 and Xyn11G being inactive and active, respectively.

Eukaryotic proteins that are destined for extracellular environments often undergo post-translational modification, especially N-glycosylation [358,318]. Thirteen of the 15 proteins were analysed for N-glycosylation by glycol-staining on SDS-PAGE gel (Figure 14 and Table 6). Eight proteins were found to be N-glycosylated, and 5 did not stain positive on SDS-PAGE with glycoprotein staining kit. With the exception of Bgl3G, Exg5A (which were predicted to be glycosylated but did not stain positive on glycoprotein

staining), Exg17B and hypothetical Lic16A (which stained positively on glycosylation, though without any predicted N-glycosyltation site), glycosylation was correctly predicted for the rest of the proteins. Although N-glycosylation is known to be prevalent in proteins destined for extracellular environments [358], Lic16A was predicted to possess up to 24 *O*-glycosylation sites (results not shown). Its SDS-PAGE molecular weight was 63.1 kDa, which is about 26 kDa higher than the predicted mass.

Table 6. Summary of glycosylation of *A. niger* glycoside hydrolases produced in *Pichia pastoris* host system

				N-glycosylation		
Protein	JGI ID	GH Family	Predicted MW (kDa)	Predicted N-sites	Glycosylated	Expression level in *P. pastoris (mg/ml)*
Bgl3N	210981	3	87.3	5	-	<0.2
Cbh6C	54490	6	48.1	0	•	≥0.5
Egl5C	205580	5	43.4	2	+	≥0.5
Exg5	52811	5	43.1	3	+	≥0.5
Exg5A	202490	5	45.7	1	-	≥0.5
Exg17B	191172	17	33.4	0	+	≥0.5
Lic16A	199085	16	37.1	0	+	<0.2
Mog63	56298	63	93.0	0	-	< 0.2
Man76	134687	76	38.0	4	+	≥0.5
Man26A	40875	26	73.4	1	+	<0.2
Xeg12	191511	12	27.0	0	-	≥0.5
Xyn11G	171269	11	24.9	0	-	≥0.5
Xyn11	183088	11	25.8	2	•	≥0.5
Xyl43B	38924	43	35.4	3	+	<0.2
Ugl88	41877	88	48.2	6	+	≥0.5

≥ 0.5 mg/ml secreted protein: high protein expression; <0.2 mg/ml secreted protein: low protein expression; •, experimental determination of N-glycosylation was not attempted; +, positive staining of protein on SDS-PAGE with glycoprotein staining kit; -, negative staining with glycoprotein staining kit; JGI ID, Joint Genome Institute gene ID number.

Similarly, Exg17B which was found to stain positively with the glycol-stain, was predicted to contain one *O*-glycosylation site and no *N*-glycosylation site in its amino acid sequence. As expected, proteins produced in *E. coli* did not show glycosylation upon glycol-staining (Figure 14).

The SDS-PAGE molecular weights of all proteins were around the predicted value, with the exception of Lic16A, Xyl43B and Exg5 which showed SDS-PAGE molecular weights more than 20 kDa higher than the predicted value (Figure 14). Glycosylation did not appear to significantly influence the molecular mass of the rest of the proteins.

It is not possible to compare level of protein production and biochemical properties for proteins expressed in *P. pastoris* with those from *E. coli* and *A. niger* because only 2 of the 15 proteins were tested for expression in *A. niger* and *E. coli*. Furthermore, none of these proteins have been previously expressed or characterized, limiting the possibilities of comparative analysis of results obtained in this study. Similarly, the effect of the various tags on the biochemical properties of the fusion proteins cannot be assessed.

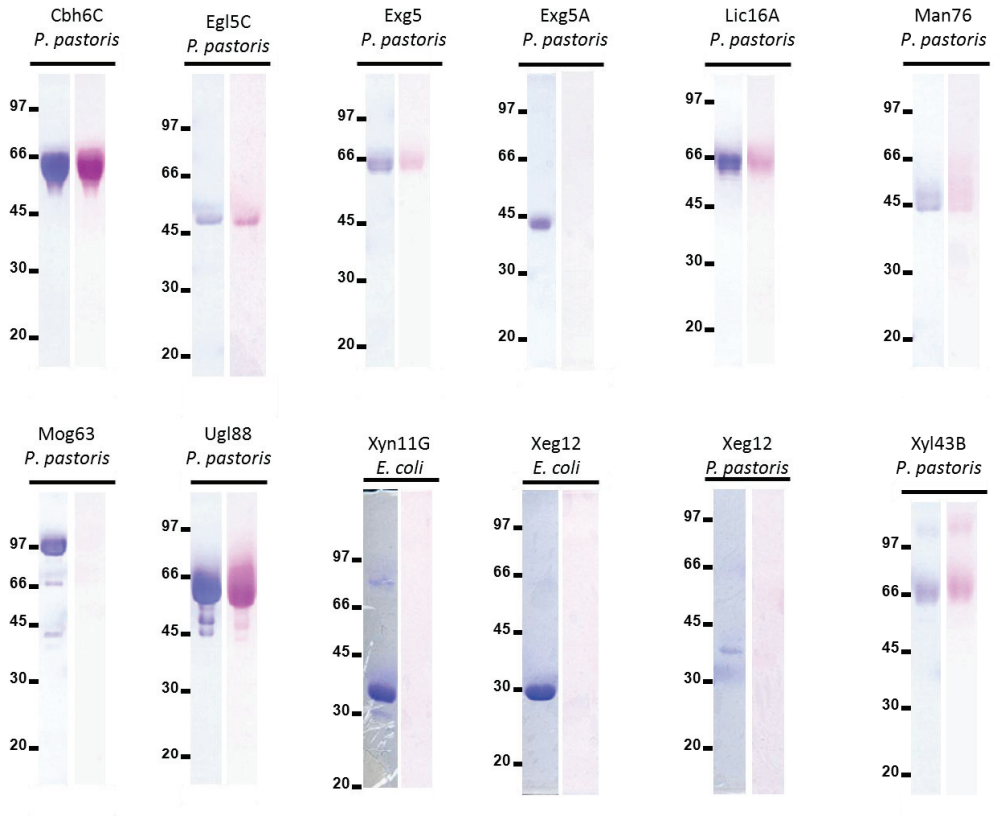


Figure 14. SDS-PAGE and glyco-staining of some of the purified recombinant enzymes. Proteins were stained with Coomassie Blue (left lanes). Proteins in duplicate SDS gels were stained with Pierce® Glycoprotein staining kit (right lanes).

Biochemical properties

All produced proteins were characterized for specific activity and pH and temperature profiles as shown in Table 7 and Figure 15 respectively. All the biochemical characterization data is summarized in Table 7. The biochemical properties of Xyn11G were identical for both the *P. pastoris* and *E. coli* expressed protein (Figure 15). However, Xeg12 was active when produced in *P. Pastoris* and the parent host, but was inactive when produced in *E. coli.* Eight of the 15 proteins expressed in *P. pastoris* did not show the predicted activity. Out of 11 proteins tested for expression in *A. niger,* 5 were successfully expressed, four of which were active. The proteins were further tested for a variety of other activities on different substrates, but they were all inactive (Table 7). Although the hypothetical GH63 Mannosyl-oligosaccharide glucosidase did not show activity on any of the manno-carbohydrate substrates tested, it was not assayed for its predicted activity using the appropriate model substrate, 4-methylumbelliferyl-α-D-glucoside or α-N-acetyl hexamannosyl glucooligosaccharide.

For the five proteins that were expressed in *P. pastoris* and that showed activity, Bgl3N, Man26A, Xeg12, Xyn11G and Egl5C, the optimal temperature ranged between 40 – 70 °C, Xyn11G and Egl5C showing the lowest and highest values respectively. The optimal temperature of Egl5C was at least 10°C higher than that of the other proteins. The optimal pH ranged between 3.5 – 5.5. Meanwhile, Xyn11G was found to have a temperature optimum which was 10°C lower than that of three of the other active enzymes, suggesting that this enzyme may be relatively thermo-labile.

Table 7. Biochemical properties of recombinant *A. niger* glycoside hydrolases produce in *P. pastoris* host system*

Protein	Predicted enzyme activity	Assay substrates	Temperature optimum (°C)	pH optimum	Specific activity (Units / mg protein)	Expressed in *A. niger*
Bgl3N	β-glucosidase	pNPG	50	3.5 – 5.0	123	Yes/Active
Cbh6C	Cellobiohydrolase	pNPC, pNPL, Avicel	NA	NA	NA	NT
Egl5C	Endoglucanase	CMC-4M	70	3.5 – 4.5	46	No
Exg5	Exo-β-1,3-glucanase	EG glucan, CMC-4M	NA	NA	NA	Yes/NA
Exg5A	Exo-β-1,3-glucanase	EG glucan, CMC-4M	NA	NA	NA	No
Exg17B	Exo-1,3-β-glucosidase	EG glucan, CMC-4M	NA	NA	NA	NT
Lic16A	Licheninase	Lichenan,CMC-4M,Avicel	NA	NA	NA	NT
Man26A	Mannanase	LBG	45	4.0 – 5.0	50	No
Man76	α-1,6-mannanase	Yeast mannan, LBG	NA	NA	NA	No
Mog63	Mannosyl-oligosaccharide glucosidase	LBG, yeast mannan, pNPM	NA	NA	NA	No
		4-MU-α-D-glucoside, Glc3Man6GlcNAc	NT	NT	NT	
Xeg12	Xyloglucanase	TS xyloglucan	50 (NA)	3.5 – 5.5 (NA)	38 (NA)	Yes /Active
Xyn11G	Xylanase	BW xylan	40 (40)	3.5 – 5.0 (3.5 – 5.0)	146 (137)	Yes/Active
Xyn11	Xylanase	BW xylan, pNPX	NA	NA	NA	NT
Xyl43B	Xylosidase	pNPX	NA	NA	NA	Yes/Active
Ugl88	Unsaturated glucoronyl hydrolase	unsaturated gellan tetrasaccharide (ΔGlcA-Glc-Rha-Glc)	NT	NT	NT	NT

*Optimal pH is presented in a range of pH where the enzyme activity is ≥80% of peak activity. Values in parenthesis are for the *E. coli* expressed protein. NA denotes, not active; NT, not tested; pNPG, 4-nitrophenyl glucoside; pNPC, 4-nitrophenyl cellobioside; pNPL, 4-nitrophenyl lactoside; pNPM, 4-nitrophenyl mannoside; pNPX, 4-nitrophenyl xyloside; CMC-4M, carboxymethyl cellulose-4M; EG glucan, *Euglena gracilis* glucan; BW xylan, birchwood xylan; LBG, locust bean gum; 4-MU-α-D-glucoside, 4-methylumbelliferyl-α-D-glucoside; and Glc3Man6GlcNAc, α-N-acetyl hexamannosyl glucooligosaccharide.

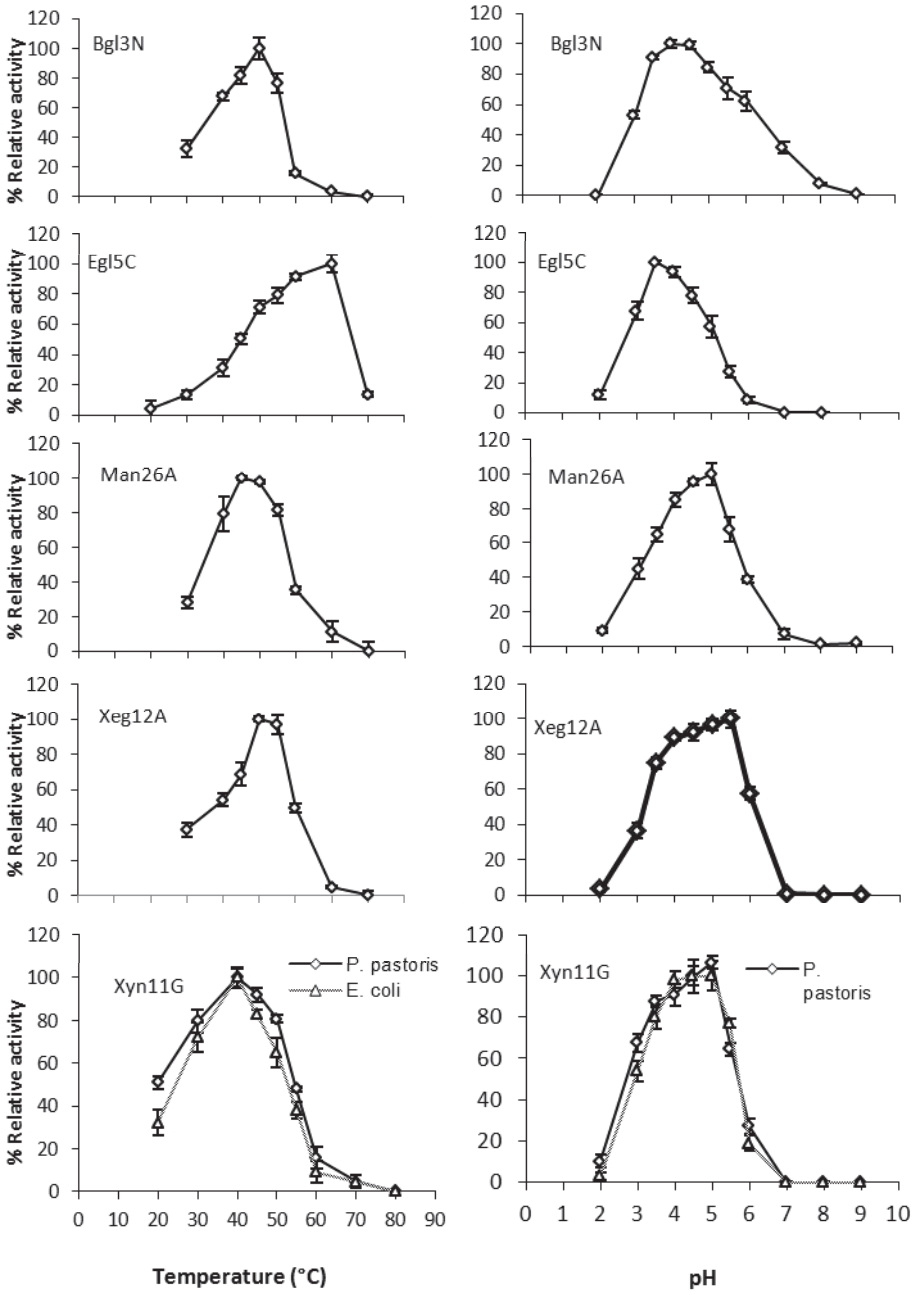


Figure 15. Temperature (left column) and pH (right column) profiles of recombinant *A. niger* enzymes produced in *P. pastoris* and *E. coli* hosts. For the determination of temperature activity profile, the enzymes were assayed at different temperatures at pH 5: x-axis, temperatures in °C; y-axis, relative activity as percentage of peak activity. For the determination of pH activity profile, enzymes were assayed at 40 °C: x-axis, pH values; y-axis, relative activity as percentage of peak activity.

Sequencing of genomes of lignocellulose degrading species provides a huge resource and opportunity for the development of novel biocatalysts for the prospective lignocellulose-based bioeconomy [117,282,282,118]. Mining of the lignocellulolytic genes and their expression in heterologous systems will improve the scientific community's understanding of the encoded proteins and enable production of novel protein targets which can be used as additives in existing cellulase mixtures or as components of synthetic minimal enzyme cocktails for lignocellulose degradation. Enzymes from a library of such characterized recombinant protein targets must be screened on natural lignocellulosic biomass to assess their efficiency in cell wall degradation.

3.5 Conclusions

All the fifteen *A. niger* proteins were successfully expressed heterologously in *P. pastoris,* ten of which did not show the predicted activity and five were biochemically characterized. With about two-thirds of the expressed proteins inactive, *P. pastoris* did not appear to be a suitable expression host for these proteins. It was not possible to know whether the heterologous expression of these in *P. pastoris* or *E. coli* conferred any catalytic benefit which could be taken advantage of in the development of lignocellulolytic enzyme cocktails. There was no evidence to suggest that the activity of the proteins was influenced by glycosylation or the lack of it.

In order to assess the hydrolytic potential of lignocellulolytic proteins such as those reported in Manuscript I and II on natural biomass, a medium- to high-throughput-adaptable screening method was developed as reported in Manuscript III.

Chapter 4 Manuscript III

A microplate-based method for screening fungal enzymes on high consistency biomass

Enongene Ekwe, Adrian Tsang, Justin Powlowski

Centre of Structural and Functional Genomics, Concordia University, Montreal, QC, Canada

Manuscript

4.1 Abstract

For economy of scale, bioconversion of lignocellulose to fuels and chemicals must operate at solids loadings in excess of 15% in commercial scale. Proteins developed for use in such industrial processes consequently require tolerance for high solids loadings. We have developed a customized, high-throughput compatible 96-well microfilter plate method for screening lignocellulolytic enzymes and pretreated biomasses at solids loading of 15% w/v or higher. The method was able to maintain residual moisture content of ≥ 90% after 72h, indicating limited effect from evaporative loss. It was reproducible and showed a dose-dependent response in sugar yields at protein loadings up to 32 mg/g dry biomass (≤60 mg/g glucan). This method has shown differences in digestibility of biomass as a function of biomass type and pretreatment method. This method was able to identify a GH7 cellobiohydrolase from *A. niger* (Cbh7A-1), a GH10 xylanase from *T. terrestris* (Xyn10-2) and a putative acetylesterase (Ace-1) from *A. niger* as promising additives for enhancement of cellulase-driven hydrolysis of 15% thermomechanically pretreated aspen pulp. Small scale reactions used in this method minimize the amount of enzyme required.

4.2 Introduction

Lignocellulosic biomass is a resource that is of great interest as a raw material to provide sugars for industrial applications, such as biofuels and bioproducts. The structural complexity and heterogeneity of lignocellulose in plant cell walls necessitates a diverse enzyme repertoire for complete hydrolysis of the constituent carbohydrates into fermentable monomeric sugars [14,110]. Complex cellulolytic enzyme cocktails are known to be secreted by saprophytic and phytoparasitic fungal species [97,359] and several commercial enzyme mixtures composed of cellulases, hemicellulases, pectinases, and other accessory proteins have been developed [360,361,362].

The hydrolytic efficiency of cellulase mixtures is limited by the crystallinity of cellulose and its association with hemicellulose and lignin. Biomass pretreatment significantly improves hydrolysis of lignocelluloses by altering the structural integrity of cellulose and removing hemicelluloses and lignin, but other limiting factors include non-productive adsorption of enzymes by lignin and other structural components, the presence of enzyme inhibitors, and the incomplete disruption of cellulose crystallinity [208]. These factors, together with the relatively slow turnover rates of cellulases, dictate the use of large amounts of enzymes for efficient hydrolysis, which increases the process cost. The recent emergence of Novozymes' cellic enzyme cocktail, Ctec 2 has cut the lignocellulolytic enzyme requirement by half, but further improvements are still needed for economic viability of lignocellulolytic bioprocesses [202].

Research has been directed towards development of cellulase enzymes and enzyme cocktails with improved hydrolytic performance. One approach is to screen microbial species for enzymes with novel properties [363,364,365]. This requires efficient and rapid activity assay techniques, preferably microplate-based to increase throughput, and assay conditions that mimic industrially-relevant hydrolysis using actual pre-treated substrates. Many screening methods use synthetic/model substrates, even though activity assays on model substrates hardly correlate well with the true hydrolytic capabilities of proteins on natural lignocellulosic feedstock [134]. In addition, biomass

consistencies reported so far for screening lignocellulolytic enzymes are typically low (≤ 5% dry weight per volume) [366,367] which contrast with the high biomass loading (≥ 15% dry weight per volume) required for economically viable industrial processes. High solids loading is imperative for high sugar and ethanol titers at reduced cost [368,369,245].

Many researchers prefer lower biomass consistencies due to the constraints associated with hydrolysis at high solids loading. It has been shown that cellulose conversion rate falls linearly with increasing consistency [370]. This is because there is a linear decrease in cellulase adsorption onto cellulose in biomass as substrate consistency increases. This is the result of (1) an increase in non-productive adsorption of enzyme onto lignin [371] and (2) increasing cellobiose concentrations, an efficient inhibitor of cellobiohydrolase adsorption to biomass [268]. Furthermore, biomass degradation products from pretreatment (e.g. furfural and hydroxymethyl furfural - HMF, solubilized lignin and phenolic derivatives of lignin), increase in concentration with solids loading, and are strong inhibitors of cell wall degrading enzymes [101,102] and/or fermentation. Another factor is that at high biomass loading, water availability falls, rheological effects and mass transfer limitations become highly unfavourable [370,372].

Commercialization of processes using lignocellulosic biomass will require efficient enzyme cocktails that are tolerant to high solids loading and its attendant problems. Direct screening of new biomass-degrading proteins on high consistency pretreated lignocellulosic biomass will allow identification of enzymes which demonstrate tolerance to solids loading and associated inhibitors. These can be directly used as crude enzyme cocktails, as additives for enhancing the performance of commercial cellulases, or as components of synthetic/minimal cellulolytic enzyme cocktails [176,263].

We have been engaged in mining fungal genomes for genes that encode biomass-active proteins, expressing them, and characterizing their catalytic properties. In this work, we describe a 96-well microplate-based technique for the screening of pure

and crude fungal protein preparations using natural pretreated lignocellulosic biomass at consistencies greater than 15% dry weight per volume.

4.3 Materials and Methods

Substrates and chemicals

Chemically and thermally pretreated lignocellulosic biomass samples used in this study were kindly donated by Prof. Jack Saddler of the University of British Columbia (Vancouver, BC). Thermo-mechanically pretreated Aspen pulp was a gift from FP Innovations (Montreal, QC). Avicel was purchased from Sigma-Aldrich (St. Louis, MO). Solka floc was purchased from (Brown, Berlin, NH). Amplex Red glucose/Glucose oxidase assay kit was purchased from Invitrogen (London, ON). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

Enzymes

Commercial cellulases Celluclast 1.5L (Cellulase from *Trichoderma reesei*) and Novozyme-188 (*β*-glucosidase from *Aspergillus niger*) were purchased from Sigma-Aldrich (St. Louis, MO). Recombinant enzymes were expressed in *A. niger* or *Pichia pastoris* host systems as described previously [355]. Where necessary, protein purification was done by anion exchange chromatography using 5ml HiTrap QXL column pre-packed with Q Sepharose XL anion exchanger from GE Healthcare Life Sciences (Baie d'Urfe, QC) as described previously [355]. The column was run on ÄKTA Chromatography System from GE Pharmacia (Baie d'Urfe, QC) as specified by the manufacturer.

Protein concentration

The benchmark cellulases, Celluclast 1.5L and Novozym 188 were diluted 100-fold in 50 mM sodium acetate buffer, pH 4.8, before use. Crude recombinant enzyme preparations were also appropriately diluted in the same buffer for concentration determination. To estimate protein concentrations, commercial and crude recombinant proteins were first precipitated in 3% TCA (2,2,2-trichloroacetic acid) as described elsewhere [373] followed by protein assay using the Pierce BCA protein assay kit in microplate format according to the manufacturer's instructions. Purified recombinant proteins were directly quantified by BCA protein assay without TCA precipitation. BSA was used as standard and all reactions were done in triplicate. The total protein concentrations obtained for Celluclast 1.5L and Novozym 188 were 142 mg/ml and 84 mg/ml, respectively.

Activity determination

In order to compare activity of commercial cellulases with values reported in the literature [374,361], cellulase activity in Celluclast 1.5L was determined via the microplate (60 µl) version of the dinitrosalicylic acid (DNS) method described by Xiao and co-workers using Whatman No.1 filter paper as substrate [374]. Glucose (0.02 – 0.2 mg) was used as standard. Cellulase activity was expressed in filter paper units (FPU). One FPU was defined as the amount of cellulase enzyme that released 2.0 mg of glucose equivalent from filter paper in 1 h at 50°C and pH 4.8 [375]. Cellulase activity in Celluclast was 52 FPU/ml. The β-glucosidase activity in Novozym 188 was determined as described previously [376] albeit with slight modification described in reference [355].

Recombinant enzymes were screened for activity in 96-well microplates via a modified BCA reducing sugar (for polysaccharide substrates [377]) or *p*-nitrophenol (for *p*-nitrophenol derivative substrates [376]) assay as described in reference [355]. All enzyme assays were carried out in triplicate. The appropriate monomeric sugar or *p*-

nitrophenol was used as standard for BCA reducing sugar or *p*-nitrophenol assay, respectively [378]. One unit of activity was defined as the amount of enzyme that released one µmole of product per minute under standard conditions.

Biochemical analysis of biomass

Structural carbohydrates and lignin in Aspen and Jackpine mechanical pulps were determined according to NREL's protocol NREL/TP-510-42618 of April 2008 [379] by our collaborators at FP Innovations, Pointe-Claire, QC. Composition analyses of all other pre-treated biomass samples were carried out at the Saddler laboratory at UBC.

Biomass preparation

Pretreated woody biomass was weighed (50 g) and suspended in 500 ml of 50 mM sodium acetate buffer, pH4.8. The mixture was homogenized in a 1 L steel cup of 115V seven-speed Waring Commercial Laboratory Blender (model E8140) at setting 4 for 5 min, followed by setting 7 for 2 min to reduce particle size or fiber length. The homogenate was transferred into a storage container, sodium azide added to a final concentration of 0.01% to prevent microbial growth and the preparation stored at 4˚C until further use.

Biomass consistency determination

To determine biomass dry weight, 1 ml of slurry was loaded into pre-weighed eppendorf tubes in triplicate using a wide-bore pipette tip and centrifuged at 10,000 × g for 10 min. The pellet was dried at 105˚C for 3 h, dry weight determined and consistency (% w/v) calculated. The consistency was adjusted with 50 mM sodium acetate buffer to 3% or less in order to facilitate pipetting.

Filter plate packing

A low protein-binding, multiscreen 96-well microfilter plate fitted with 1.2 µm pore size polyvinylidene fluoride membrane (Millipore, Billerica, MA) was adapted for hydrolysis reactions. Homogenized pretreated wood biomass slurry equivalent to 3 mg solids was loaded onto the microfilter plate using wide-bore pipette tips. The loaded plate was fitted on a 96-well collection plate and the cassette centrifuged at 1000 × g for 5 min to remove liquid from the biomass. The bottom outlets of the filter plate wells were then sealed using adhesive aluminum material and maintained on ice in a cold room (4 °C) with the top covered to prevent moisture loss. The volume of filtrate per well in the collection plate was measured. The difference in volume between the loaded slurry and filtrate was recorded as volume of residual biomass on the filter plate. The volume of enzyme mixture to be added to 3 mg biomass per well to reach 15% solid loading was then calculated.

Estimation of evaporative loss

To estimate evaporative loss from the reaction wells, four microfilter plates were packed with biomass as described above and 12 µl of 5 mM *p*-nitrophenol solution added per well. The plates were sealed, incubated at 50°C, one withdrawn every 24 h and stored at -20°C. There were six replicates per time point. The bottom seal was removed (top seal maintained in place) and liquid in the plate wells harvested by centrifugation at 3000 × g for 2 min at room temperature. The residual biomass was washed twice with 24 µl distilled water at 3000 × g for 2 min into the collection plate. To estimate residual concentration of *p*-nitrophenol, 60 µl of 1M Na_2CO_3 was added to each well, 80 ul of the mixture transferred to a 96-well reading plate, optical density measured at 410 nm and compared to a *p*-nitrophenol (0 – 1.125 mM) standard curve.

Assessing enzyme distribution

In order to estimate the minimum enzyme volume that will theoretically enable even distribution of the added biocatalyst throughout the packed biomass, three BioRad minispin columns (Mississauga, ON) were washed, resin discarded and loaded with pretreated biomass slurry to 3 mg dry weight. The loaded column was placed in a 50 ml Falcon tube and packed at 1000 × g for 5 minutes on a swinging bucket rotor. A coloured protein solution (0.5% haemoglobin) was quantitatively added to the centre of the biomass bed using a 500 µl syringe from Hamilton (Reno, NV) and colour distribution observed by eye until the protein was evenly distributed. The total volume of model protein added was recorded and termed 'critical enzyme volume'. The 'critical enzyme volume' was substantially less than the volume of enzyme or buffer required to reach target consistency.

Biomass hydrolysis: enhancement of lignocellulose hydrolysis

Recombinant proteins (crude or purified) were screened for the enhancement of cellulase-driven glucose release from thermomechanically pretreated (TMP) Aspen pulp at minimal protein loading. Commercial cellulase benchmark was prepared from Celluclast and Novozym 188 at a ratio of 9/1 in 50 mM sodium acetate buffer, pH 4.8. At a total protein loading of 4 mg/g substrate dry matter, the benchmark cellulase (3.6 mg/g dry matter, 90% total protein) was mixed with recombinant protein (0.4 mg/g dry matter, 10% total protein). Proteins were appropriately diluted in 50 mM sodium acetate buffer, pH 4.8 in 0.01 % sodium azide. To 8 µl residual biomass determined in the filter plate packing process, 12 µl of enzyme mixture was carefully dispensed at the centre of the biomass bed. Due to the low concentration of liquid in the bed, the dispensed enzyme mixture was quickly and uniformly absorbed into the biomass, resulting in a final consistency of 15%. Biomass samples loaded with 4 mg commercial

cellulase per gram dry matter (100 % protein loading) and those with 0.4 mg recombinant protein per gram dry matter were used as control. Two layers of parafilm were overlaid on the plate, a flat cell foam gasket placed on both sides (top and bottom) and each well sealed by securing the sandwich between two 304 stainless steel plates (1.6 mm thick, durometer hardness A40), using 2 inch medium duty Record G-clamps of 1.1/8 inch thread depth as illustrated in Figure 16. One clamp was positioned in each corner of the plate. The plates, each containing multiple reactions, were incubated at 50 °C, with one withdrawn at each time point and stored at -20 °C until product analysis. To collect samples for product analysis, the clamps and bottom aluminum seal were removed (leaving the parafilm seal on the top in place) and supernatant collected at 3000 × g for 2 min into 96-well polypropylene, pyramid bottom 1.0 ml microplates. Each well in the filter plate was washed twice by adding 294 µl double distilled water and collected into the collection plate at 3000 × g for 2 min, resulting in a 50-fold diluted hydrolysate. Hydrolysis reactions were carried out in triplicate.

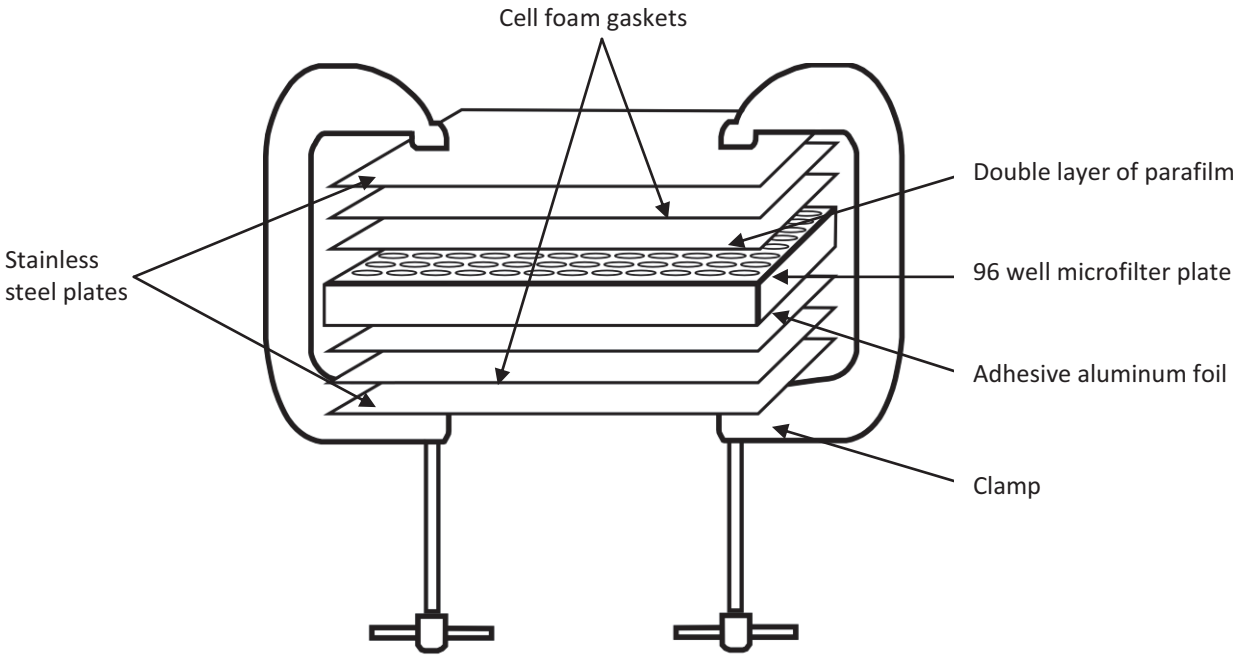


Figure 16. Schematic representation of the microplate assemblage

Biomass digestibility test

The protocol for Enzymatic Saccharification of Lignocellulosic Biomass from the National Renewable Energy Laboratory (NREL) (http://www.nrel.gov/biomass/pdfs/42629.pdf) [380] was adapted to the high-solids microplate strategy for assessment of the maximum digestible extent of various pretreated biomasses. Among the tested feedstocks were: TMP Aspen pulp, TMP Jack Pine pulp, Aspen kraft pulp, Black Spruce kraft pulp, organosolv pretreated Aspen, and organosolv pretreated Lodgepole Pine. Commercial cellulase (Celluclast 1.5 L and Novozym 188 at ratio of 9/1) was loaded onto 3 mg biomass at 128 mg cellulase per gram substrate dry matter at 15% biomass consistency. This protein loading corresponds to, at least 60 FPU/g cellulose for pretreated biomass with up to 80%

cellulose content. Plates were sealed, incubated and hydrolysate harvested as described above. All reactions were carried out in triplicate.

Product analysis

Hydrolysate was assayed for reducing sugar as described above. Glucose released was detected by glucose oxidase coupled assay [381,382,383] using the Amplex Red glucose/Glucose oxidase assay kit from Invitrogen as specified by the manufacturer. Fluorescence was measured using Wallac Victor 2 Multilabel Counter/multitask plate reader (Model 1420) from Perkin Elmer at an excitation wavelength of 544 nm and emission wavelength 590 nm. Glucose was used as standard. Test and benchmark sample readings were compared to standard curves to estimate reducing sugar and glucose concentrations. Reducing sugar and glucose yields were normalized against total carbohydrate and glucan contents of the biomass, respectively. In the calculations for reducing sugar conversion it was assumed that the reducing ends in the hydrolysate reacted with BCA reagent in the same way as glucose. The glucose yield, i.e. % gram glucose per gram cellulose, was calculated as shown in Equation 1 [380,384].

Equation 1. Calculation of % glucan conversion

Glucan conversion (%) = *100 %* × *[Glc] (mg/ml)* × *0.9* / { *[S] (mg/ml)* × *P (g/g)* }

Where *[Glc]* is the residual glucose concentration in hydrolysate expressed in mg/ml; *0.9* is the molecular mass correction factor in the conversion of cellulose to glucose equivalent; *[S]* is the dry biomass concentration at the beginning of hydrolysis expressed in mg/ml; P is the cellulose fraction in the dry biomass; and 100 % is the maximum percentage theoretical glucan conversion.

For instance; the enzymatic hydrolysis of 15 % (equivalent to 150 mg/ml) blackspruce kraft pulp with a glucan content of 76 % yields hydrolysate with glucose concentration of 17 mg/ml. The % glucan conversion from this reaction will be calculated as follows:

Glucan conversion $= 100\ \% \times 17\ mg/ml \times 0.9 / (150\ mg/ml \times 0.76)$

$= 13.4\ \%$

Enzyme dose response

To evaluate enzyme dose response, increasing amounts of commercial cellulase (2 – 128 mg per gram dry biomass) were loaded on 3 mg TMP Aspen pulp in microfilter plate. All downstream procedures for the determination of glucose or reducing sugar yields were performed as described above. Reactions were performed in triplicate.

4.4 Results

Concentration and activity of Commercial Cellulase

Total protein concentration in Celluclast 1.5L was slightly higher than that reported by Arantes and Saddler, 2011 (129.8 mg/ml [133,385]) and lower than that reported by Himmel's group (166 mg/ml [361]), possibly due to variations in batches of the commercial product. Cellulase activity in Celluclast was 52 FPU/ml as reported previously [133,385].

Composition of pre-treated biomass samples

The cellulose xylan and lignin content of biomass was as shown in Table 8 below. Other hemicelluloses, sulphur, extractives and ash were not determined. With respect to the pretreatment method, glucan content of biomass, in decreasing order, was as follows: kraft > organosolv > TMP. Similarly, the xylan content was in the order: TMP > Kraft > organosolv; and the lignin content was in the order: TMP > organosolv > Kraft.

Table 8. Carbohydrate composition of various pretreated biomass

Biomass Type	**Glucan (%)**	**Xylan (%)**	**Lignin (%)**
TMP Aspen pulp	48.6	20.4	22.2
TMP Jack pine pulp	42.8	18.3	24.2
Aspen kraft pulp S1	83.3	7.1	3.4
Aspen kraft pulp S2	69.7	11.2	5.8
Black spruce kraft pulp	75.1	9.9	7.8
Organosolv lodgepole pine S1	65.3	0.9	19.4
Organosolv lodgepole pine S2	59.0	4.2	21.0

A high-consistency microplate based screening method

In order to develop a high throughput-adaptable screening strategy for proteins at high substrate consistency, we customized a low protein-binding, multiscreen 96-well microfilter plate fitted with a polyvinylidene fluoride membrane as micro-scale reactor system for hydrolysis as shown in Figure 16. With this system, a number of fundamental questions were to be addressed: (1) would this microplate method be suitable for handling small amounts of biomass at consistencies higher than 15% without losing significant amounts of moisture at 50°C [i.e. is the seal of the microfilter plate effective in preventing moisture loss at this temperature], (2) is this method reproducible and (3) is this method sensitive to varying enzyme doses?

Evaporative loss

Evaporative loss from the microplate reaction wells at 50°C for up to 72 h was evaluated. A *p*-nitrophenol solution was used as an indicator to detect increase in concentration which would be indicative of moisture loss: *p*-nitrophenol absorbs maximally at 410 nm in alkaline solution. The mean of the test well to control/reference well (evaporative loss at time 0h) ratios of OD_{410} readings was 1.00 ± 0.09. The standard deviation of residual moisture in the test wells was 7.6 % after 24 h, 6.4 % after 48 h and 5.5 % after 72 h, relative to 9.8 % for the 0 h control (Figure 17). Overall, there was a 10 % loss of moisture in the first 24 h after which no apparent change in moisture content was observed (Figure 17).

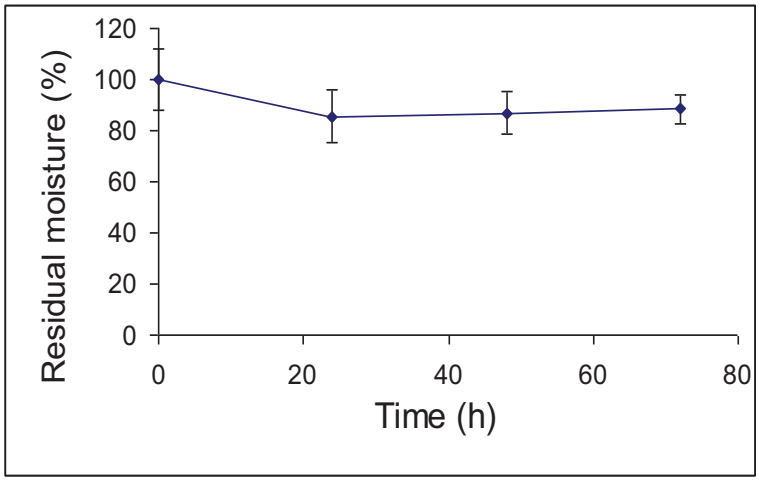


Figure 17. Evaporative loss from the high-consistency microplate reaction system incubated at 50°C. There were nine independent replicates per point and error bars represent standard error.

Method reproducibility

Reproducibility of the method was checked by running 15 replicates of TMP Aspen hydrolysis reactions at 50 °C and commercial cellulase (Celluclast supplemented with Novozym 188 at ratio of 9/1) loading of 4 mg/g dry biomass, results for which are shown in Table 9. For the various time points, the coefficient of variation (CV) for both glucose and reducing sugar yields were all between 4.6 % and 8.4 %. It was also observed that while the CV values increased with time for reducing end yields, there was a time dependent fall in CV for glucan conversion. Finally, as observed in the dose response analysis results, glucose and reducing sugar yields increased over the time course of reaction.

Table 9. High consistency microplate method reproducibility*

Replicate	Glucose yield (%)			Reducing end yield (%)		
number	12h	24h	48h	12h	24h	48h
1	7.7	12.3	14.6	12.3	16.9	20.9
2	9.2	11.2	13.3	12.6	16.9	22.8
3	9.6	10.2	13.5	11.5	15.7	20.1
4	8.0	10.4	13.0	10.9	17.6	21.2
5	8.3	10.7	15.2	12.3	15.6	21.4
6	7.5	12.0	14.9	13.3	17.3	19.4
7	8.1	9.9	15.5	12.4	16.4	18.9
8	8.0	9.8	13.3	12.9	14.3	20.7
9	8.7	10.9	14.2	11.9	15.0	20.4
10	8.3	10.7	14.0	12.7	16.3	19.4
11	7.2	11.5	14.9	12.2	15.6	18.2
12	8.4	11.3	14.7	12.4	17.3	18.1
13	7.6	11.0	15.5	12.2	15.4	22.2
14	8.9	12.8	15.6	12.3	14.4	20.9
15	9.2	11.5	14.3	12.4	15.0	21.9
Mean	**8.3**	**11.1**	**14.3**	**12.3**	**15.9**	**20.4**
Stdev	**0.7**	**0.9**	**0.9**	**0.6**	**1.1**	**1.4**
CV (%)	**8.4**	**7.7**	**6.0**	**4.6**	**6.7**	**6.9**

*Glucose and reducing sugar yields from 15 hydrolysis reaction replicates of 15 % TMP Aspen pulp by Celluclast. Reactions were run at 4 mg protein per gram dry biomass, 50 mM sodium acetate buffer, pH 4.8, 50 °C. Stdev denotes standard deviation and CV coefficient of variation

Enzyme dose response

In order to evaluate the sensitivity of biomass hydrolysis at high consistency to enzyme loading, enzyme dose response analysis was performed and results are as shown in Figure 18. In terms of glucose and reducing sugar yields, an enzyme dose-dependent response was observed between 2 mg and 32 mg cellulase/g dry biomass at 14 h, 24 h and 48 h time points (Figure 18). Beyond 32 mg protein/g dry matter there was no significant difference in yields at each time point, suggesting that the amount of carbohydrate left in the biomass (i.e. approx. 60 % and 40 % initial glucose and total carbohydrates respectively) was relatively inaccessible.

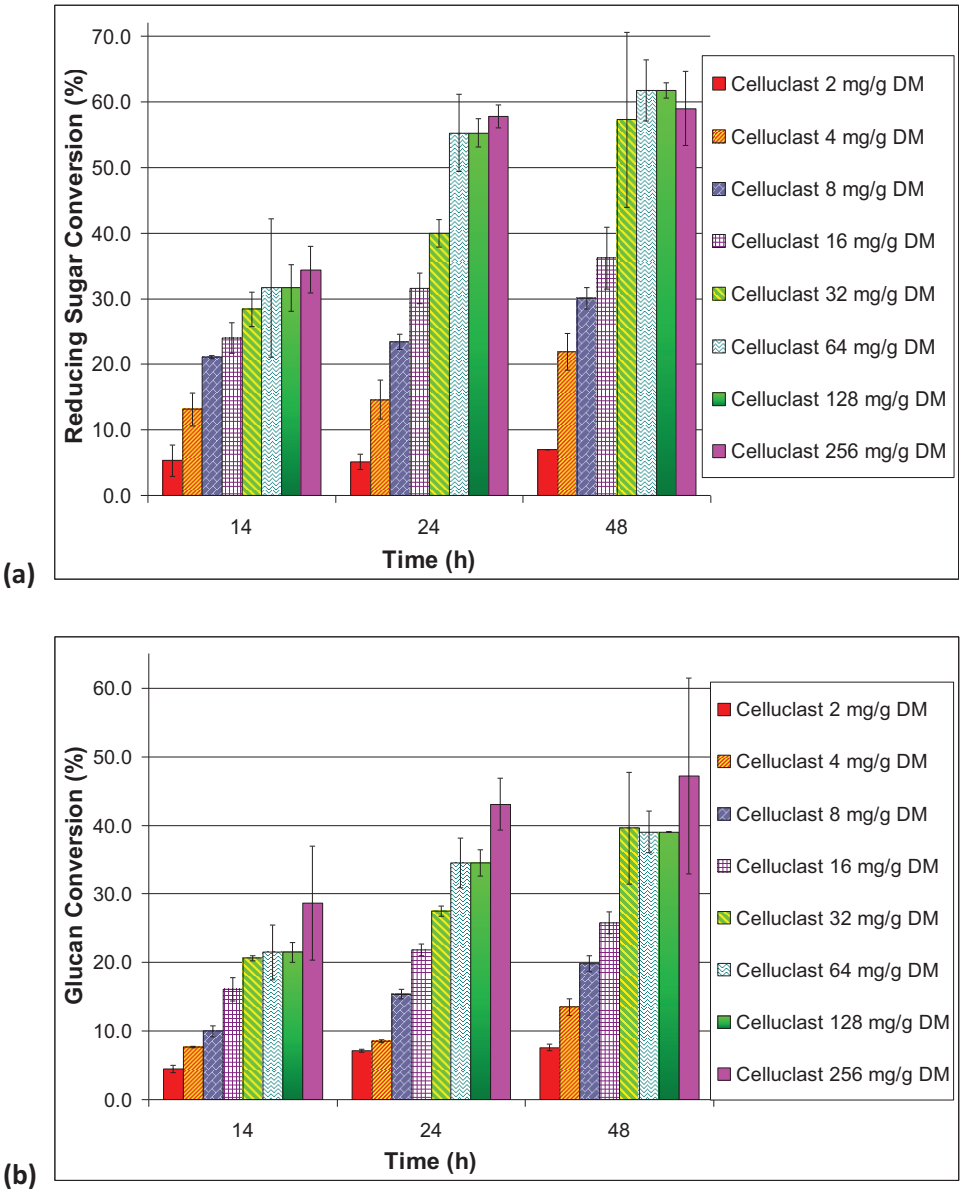


Figure 18. Enzyme dose response analysis using the high consistency microplate method. Increasing amounts of Celluclast (2 – 256 mg per gram dry biomass) were loaded on 3 mg TMP Aspen pulp in microfilter plate at 15 % consistency. Celluclast was supplement with Novozym 188 at ratio 9/1. Reducing sugar and glucose yields were normalized against total carbohydrate and glucan contents of the biomass, respectively. A: Reducing sugar yield, B: Glucose yield. Reactions were carried out in triplicate. Error bars represent standard deviation.

Hydrolysability assessment of pretreated biomass

Seven pretreated biomass samples were tested, including mechanically pretreated Aspen pulp, mechanically pretreated Jackpine pulp, two Aspen kraft pulp samples with pretreatment severity factors designated as 1 and 2 respectively, Black Spruce kraft pulp and two organosolv pretreated Lodgepole pine samples with pretreatment severity factors designated as 1 and 2 respectively. Biomass pre-treatment severity levels 1 and 2 differed by temperature, pressure, pre-treatment chemical concentration, residence time in reactor or a combination of these.

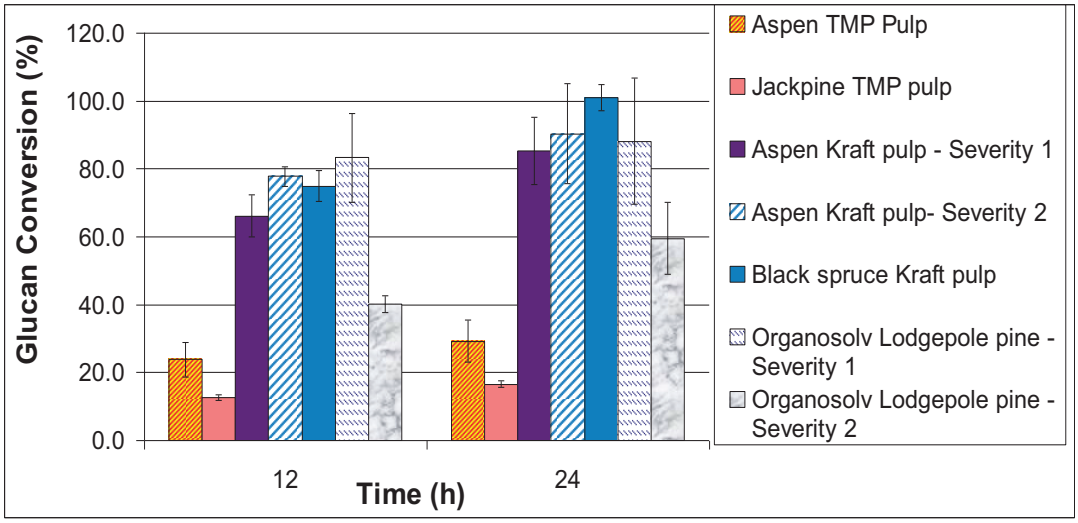


Figure 19. The digestibility of various pretreated biomasses by Celluclast at 15 % substrate consistency were tested to assess pre-treatment efficiency. Biomass pre-treatment severity levels 1 and 2 differ by temperature, pressure, pre-treatment chemical concentration or a combination of these. Reactions conditions: 3 mg biomass per well, Celluclast loading: 128 mg/g dry biomass (which corresponds to at least 60 FPU/g cellulose for pretreated biomass with up to 80% cellulose content), 50 mM sodium acetate buffer, pH 4.8, 50˚C. Celluclast was supplemented with Novozym 188 at ratio 9/1.

Hydrolysability tests on these substrates showed that thermomechanical pretreatment is less effective in improving cellulose accessibility to Celluclast than krafting or organosolv pretreatment (Figure 19). TMP pulps were far more recalcitrant, with Aspen TMP and Jackpine TMP releasing only 29 % and 18 % glucose respectively, after 24 h. In the same period, organoslv pretreated Lodgepole pine (pretreatment severity 2) yielded 60 % glucose, while Aspen kraft pulp (pretreatment severity 1 and 2) and organosolv pretreated Lodgepole pine gave up to 90 % glucose yields. Black spruce kraft pulp was the most digestible with glucose yield of 100 % after 24 h. This plate method was not suitable for hydrolysis of pure amorphous (phosphoric acid swollen cellulose - PASC) and microcrystalline cellulose (solka floc and avicel). Solka floc and avicel were found to pack very tightly, forming a relatively impermeable surface which did not allow liquid to percolate easily. The enzyme solution was therefore retained on top of the packed substrate bed without being distributed evenly in the substrate. In addition, PASC was found to clog the microfilter membrane, thereby obstructing the harvest of hydrolysate (data not shown).

Screening GH proteins for cellulase enhancement

The method was tested as a screening tool for enzymes that enhance cellulase hydrolysis of 15 % TMP Aspen pulp for up to 48 h at 50 °C, as measured by glucose released relative to a non-supplemented control. Two types of enzyme screening tests were performed: (1) binary enzyme mixtures in which 10 % of total cellulase was substituted by a single test enzyme preparation, (2) ternary mixtures in which 10 % of total cellulase was substituted by two test enzyme preparations (5 % each). Results of these two screenings strategies are shown in Figure 20 and Figure 21.

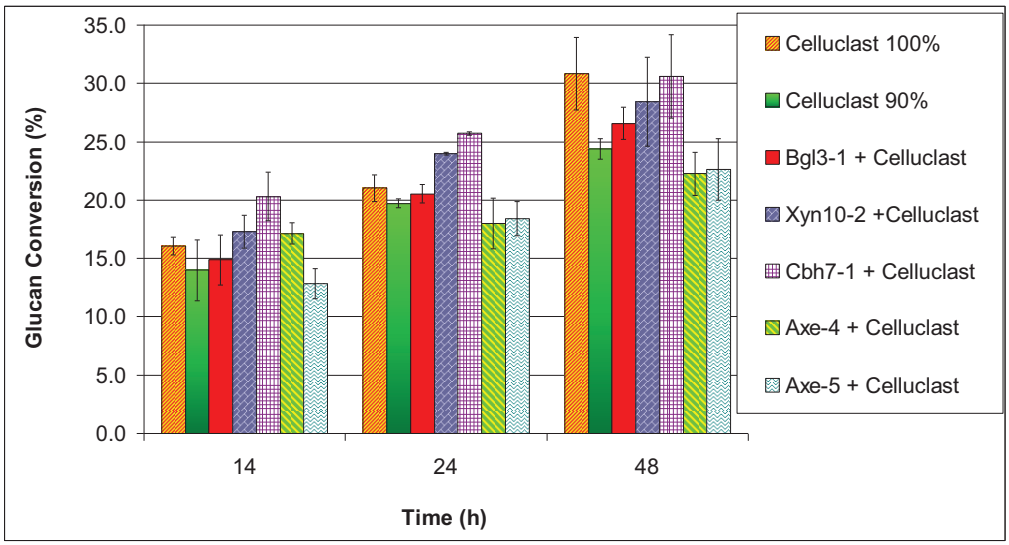


Figure 20. Binary mixtures - Screening fungal proteins for enhancement of cellulase hydrolysis of 15% TMP Aspen pulp. Bgl3-1 denotes crude *A. niger* GH3 *β*-glucosidase expressed in *A. niger*; Xyn10-2 denotes crude *T. terrestris* GH10 xylanase expressed in *A. niger*; Cbh7-1 denotes purified *A. niger* GH7 cellobiohydrolase expressed in *A. niger;* Axe-4 denotes crude *T. reesei* putative acetylxylan esterase expressed in *A. niger* and Axe-5 denotes crude *C. globosum* putative acetylxylan esterase expressed in *A. niger.*

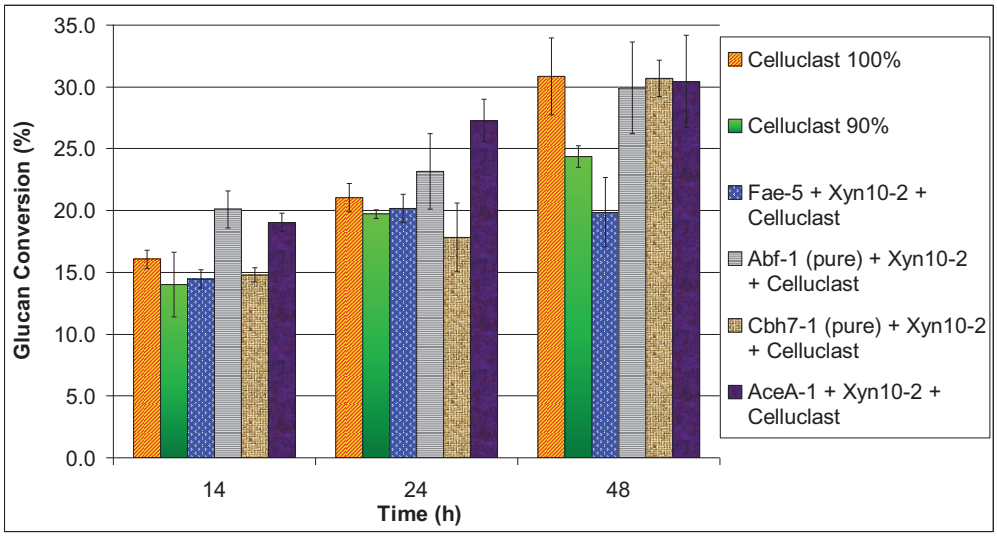


Figure 21. Ternary Mixtures - Screening fungal proteins for enhancement of cellulase hydrolysis of 15% TMPAspen pulp. Xyn10-2 denotes crude *T. terrestris* GH10 xylanase expressed in *A. niger*; Fae-5 denotes crude *C. globosum* putative feruloyl esterase expressed in *A. niger*; Abf54-1 denotes crude *A. niger* α-arabinofuranosidase expressed in *A. niger;* Cbh7-1 denotes purified *A. niger* GH7 cellobiohydrolase expressed in *A. niger* and Ace-1 denotes crude *A. niger* putative acetylesterase expressed in *A. niger.*

First, it was observed that at 24 h and 48 h time points, there was a significant difference between the amounts of glucose released in the 100 % cellulase benchmark compared with the 90 % cellulase control (Figure 20). Secondly, five proteins were tested for cellulase enhancement in binary combinations with Celluclast on TMP Aspen pulp. These include: crude GH3 *β*-glucosidase from *A. niger* (Bgl3-1), crude GH10 xylanase from *Thielavia terrestris* (Xyn10-2), purified GH7 cellobiohydrolase from *A. niger* (Cbh7-1), crude putative acetylxylan esterase from *Trichoderma reesei* (Axe-4) and crude putative acetylxylan esterase from *Chaetomium globosum* (Axe-5). Of the five protein preparations, Cbh7-1 and Xyn10-2 independently showed significant enhancement of glucose release from TMP Aspen by Celluclast for up to 24 h but not 48 h. Glucose yields from reactions containing Cbh7-1 were 20.3 % after 14 h and 25.7 %

after 24 h whereas those from the benchmark Celluclast were 16.1 % and 21 %, respectively. These data translate into 26 % and 18.3 % enhancement after 14 h and 24 h, respectively. After 48 h, the degree of enhancement by Cbh7-1 was insignificant. Glucose release from reactions containing each of the other test proteins was not significantly different from 90% benchmark control. After 48h, there was an apparent end point plateau where the glucose yields for the Cbh7-1 or Xyn10-2 supplemented samples was about 30 %, a value equal to that of the cellulase benchmark. Glucose yield from reactions containing Xyn10-2 was 24 % after 24 h and that from the benchmark Cellulase was 21%. This is equivalent to 14% enhancement. At all other time points, there was no significant difference between the yields from reactions containing only the benchmark and those containing Celluclast spiked with Xyn10-2.

Four protein preparations were screened for enhancement of a mixture of Xyn10-2 [found to enhance glucose yield in a binary mixture with Celluclast] at 5% total protein loading and Celluclast (90%). These included a putative feruloyl esterase from C. globosum (Fae-5), GH54 α-arabinofuranosidase from A. niger (Abf54-1), Cbh7-1 from A. niger and a putative acetylesterase (AceA-1) from A. niger. Results of the ternary mixture tested are represented in Figure 21. Pure α-arabinofuranosidase (Abf54-1) showed 25 % enhancement after 14 h but its effect was marginal after 24 h. AceA-1 conferred 18.8% and 30% enhancement in glucose yield after 14 h and 24 h respectively.

4.5 Discussion

At 15 % biomass consistency, the microplate system showed moisture loss of less than 10 % for up to 72 h. The fact that a fairly equal amount of moisture was lost at each time point suggests that this loss might be due to evaporation during sample handling at room temperature rather than during incubation at 50 °C i.e. some moisture could have been lost before plates were sealed and after unsealing for analysis of residual hydrolysate. The standard deviation of residual moisture obtained in this study

was between 5.5 % and 7.6 %. Studer *et al*. used a fluorescent assay strategy for leak testing in a custom-made 96-well steel microplate for combined pretreatment (at 120 – 180 °C) and enzyme hydrolysis system in screening cellulosic biomass, pretreatment conditions and enzyme formulations that enhance sugar release from biomass [367]. The standard deviation of the residual liquid volumes in the wells was 4.2 %. However, since their leak test utilised 250 µl of phosphate buffer (without biomass) at 180 °C for 55 min, it is not possible to compare their standard deviation with those obtained in this study.

Assessment of the method reproducibility revealed CV of between 4.6 % and 8.4 % for glucose and reducing sugar yields, respectively Table 9. Although a CV of ≤5 % is recommended for a reproducible dataset, values ≤ 8.4 % represent acceptable reproducibility in this case since fibrous biomass slurries are difficult to pipette with accuracy. Slight variations in quantities of biomass dispensed would increase CV values, possibly higher than 5 %. While the CV values increased with time for reducing end yields, there was a time dependent fall in CV for glucan conversion. In the plant cell wall microstructure, hemicellulose and lignin form a structural barrier around cellulose. Enzymes act on the relatively more abundant hemicellulose on the surface before gaining increased access to cellulose deep in the structure. In the early stages of the reaction, the overall reducing ends released do not depend much on cellulose accessibility and therefore vary less. On the contrary, the overall glucose yield will very much depend on cellulose accessibility which is relatively lower in the earlier phase of the reaction, with significant variations from one reaction system to another. Cellulose accessibility is progressively improved as hemicellulose is eroded to expose cellulose, thereby reversing the trend of the CV value between glucose and hemicellulose yields.

TMP aspen (15 % dry solids) shows a dose-dependent response in glucose yield at cellulase loadings in the range 2 – 32 mg/g DM (Figure 18). The reduced sensitivity to cellulase loadings in excess of 32 mg/g dry matter is likely due to saturation of enzyme binding surfaces on the accessible carbohydrates in the biomass. High relative

abundance of surface lignin and hemicellulose in biomass would reduce the enzyme saturation threshold. The TMP aspen pulp used in this study contained 51.4 % non-cellulosic material. Protein loadings higher than 32 mg/g dry biomass (equivalent to 60 % reducing sugar yield and 40 % glucan conversion after 48 h) would still leave behind significant amounts of non-cellulosic structural components in the biomass, which may limit accessibility of residfual cellulose to the cellulase enzymes [385]. Differential response was observed at very low protein loadings (i.e. 2 mg/g dry matter versus 4 mg/g dry matter – Figure 18a & b; as well as between the 90 % versus 100 % cellulase benchmark – Figure 19), suggesting that this method would be suitable for the identification of proteins which enhance cellulase-driven glucose release from lignocellulosic biomass at low enzyme loadings.

Celluclast 1.5L has been reported to possess a broad spectrum of activities, including cellulase (on filter paper, avicelase, CMCase and β-glucosidase), mannanase, pectinase and xylanase [133,175]. Glucose and reducing sugar yields in this study are thus demonstrative of the potential applicability of this method on protein cocktails with such repertoire of activities.

Seven pretreated wood biomass samples from aspen, jack pine, black spruce and lodgepole pine were tested in order to provide a basis to evaluate the effects of biomass type and pretreatment methods on biomass digestibility at high solids loading at micro-scale. This method provides a tool for assessing the efficiency of biomass pretreatment methods. TMP biomass (18-29 % glucose released in 24h) was far more recalcitrant to cellulase hydrolysis than organosolv and kraft pretreated biomass (60-100 % glucose released in 24h). Unlike the case of kraft or organosolv pretreated woody biomass which may not show significant differences in lignin composition between hardwood and softwood, thermomechanical pretreatment does not alter the original lignin content in wood [386]. Results in Figure 19 agree with previous findings that softwood (TMP Jackpine) is far more recalcitrant to cellulase degradation than hardwood (TMP Aspen) [387,388], which correlates with higher lignin content in

softwood than hardwood. Lignin acts as structural barrier which prevents cellulose accessibility, non-productively binds cellulases and may release soluble building blocks which inhibit or deactivate lignocellulolytic enzymes [101,102,370]. This method has shown differences in digestibility of biomass as a function of biomass type and pretreatment method. With respect to the pretreatment method, hydrolysability of biomass, in decreasing order, was as follows: kraft = organosolv > TMP. Softwood (Jackpine) TMP was more recalcitrant than hardwood (Aspen) TMP.

Of the five protein preparations tested in binary mixtures, Cbh7-1 and Xyn10-2 independently showed significant enhancement of glucose release from TMP Aspen by Celluclast for up to 24 h but not 48 h. Hydrolysis of lignocellulose is known to be impaired by hemicellulose and lignin in the substrate which limit accessibility of enzymes to the cellulosic chains [385]. Such 'blocking effect' imposed by xylan, can be reduced by supplementing Celluclast with xylanase. Sequential addition of xylanase followed by Celluclast in lignocellulose hydrolysis was found to reduce the blocking effect due to degradation of surrounding xylan to increase cellulose exposure to cellulase attack [133]. In addition, simultaneous addition of xylanase and Celluclast to lignocellulose has been reported to enhance glucose release to a greater magnitude than the sequential addition. This increase in performance is thought to be driven by the synergistic interaction between xylanase and cellulases [133]. The cellulase enhancement of glucose release by Xyn10-2 observed in this study might be due to both mechanisms: reduction in xylan 'blocking effect' and xylanase-cellulase synergistic interaction. A putative acetylesterase (Ace-1) from *A. niger*, which hydrolyses carboxylic ester bonds in biomass to release acetic acid, was found to enhance cellulase-driven hydrolysis of 15 % Aspen pulp in the presence of Xyn10-2.

The observed steady fall in degree of enhancement with time (26 %, 18 % and 0 % after 14 h, 24 h and 48 h, respectively) for Cbh7-1 could be due to: (1) progressive degradation of carbohydrates from biomass which increases the relative abundance of lignin onto which proteins bind non-productively [208], (2) product accumulation such

as cellobiose which leads to enzyme inhibition. Cellobiohydrolase adsorption onto cellulose is reduced when cellobiose occupies cellulose binding surfaces in the protein [268,389].

The cellobiohydrolases, Cel7A (Cbh-I) and Cel6A (Cbh-II), are known to constitute enormous portions (up to 60 % and 20 % respectively) of the total *T. reesei* secreted proteins found in Celluclast 1.5L [390,391,392,393]. Addition of a small amount of purified Cbh7-1 to Celluclast conferred such significantly high enhancement (26 % after 24 h) in the hydrolysis of TMP Aspen pulp. This observation raises questions: (1) does this Cbh7-1 show multiple activities other than the processive action cellobiohydrolases show on linear cellulose chains, (2) is the specific activity of Cbh7-1 significantly higher than those of Cel7A and Cel6A found in Celluclast, (3) does Cbh7-1 interact synergistically with other proteins in Celluclast in such ways that are more effective than those via which Cel7A and Cel6A can interact, and (4) does Cbh7-1 directly display synergism with Cel7A and Cel6A in Celluclast? While the scope of this study did not include synergistic interaction between these enzymes, it is worth noting that purified Cbh7-1 did not show activity on carboxymethyl cellulose (CMC-4M), birchwood xylan, oatspelt xylan, tamarind seed xyloglucan, polygalacturonate, pectin from citrus peel, lucost bean gum, pullulan and lichenan. Its specific activities on *p*-nitrophenyl-*β*-D-cellobioside and *p*-nitrophenyl-*β*-D-lactoside were very low (0.075 U/mg and 0.15 U/mg, respectively) which supports finding by other authors that cellobiohydrolases are noted for their slow activity [394,395]. The specific activity of the protein was not tested on microcrystalline cellulose such as avicel. It is difficult to compare the performance of Cbh7-1 against Cel7A and Cel6A since corresponding data for the latter pair was unavailable.

The thermal stability of the supplemented proteins at 50 °C was not tested. The ability of Celluclast alone to ‘catch-up’ after 48 h with the glucose yields of Celluclast supplemented with Cbh7-1 and Xyn10-2 might be due to thermal instability of Cbh7-1 and Xyn10-2 at 50 °C. Overall, this method was able to identify a GH7 cellobiohydrolase

from *A. niger* (Cbh7A-1), a GH10 xylanase from *T. terrestris* (Xyn10-2) and a putative acetylesterase (Ace-1) from *A. niger* as promising additives for enhancement of cellulase-driven hydrolysis of 15 % Aspen pulp.

In conclusion, we have developed a customized 96-well microfilter plate method for screening lignocellulolytic enzymes and pretreated biomasses at solids loading of 15 % w/v or higher. The customized microplate system was able to maintain residual moisture content of ≥ 90 % after 72 h, indicating limited effect from evaporative loss. The method proved satisfactorily reproducible and showed a dose-dependent response in sugar yields at protein loadings up to 32 mg/g dry biomass (≤60 mg/g glucan), was sensitive to differences in Celluclast loading of 0.4 mg, and proves useful in assessing the effectiveness of biomass pre-treatment methods on a broad spectrum of lignocellulosic biomass types. This method proves useful in screening fungal proteins for the enhancement of glucose release on high consistency biomass and in assessing the efficiency of biomass as a function of biomass type and pretreatment method.

Small scale reactions in this method minimize the amount of enzyme required and the 96-well plate format, approach is potentially high-throughput compatible. The use of high consistency biomass without control of interfering factors such as pretreatment-related inhibitors, low water concentration and rheological and mass transfer limitations would enable direct identification of potential synthetic cocktail candidates which would better tolerate high solids loading with associated effects.

Competing interests

The authors declare that they have no competing interests

Acknowledgements

This work was funded by Genome Canada, Génome Québec, Agriculture and Agri-Food Canada, the Bioconversion Network of the Natural Sciences and Engineering

Research Council of Canada and the US DOE. We thank Dr Jack Saddler for providing some of the substrates, Annie Bellemare for cloning the genes and producing the *A. niger* transformants of *A. niger* and *T. terrestris* proteins tested and David Taylor for production of the proteins.

Due to limited success in the identification of superior enzymes from the mesophilic filamentous fungi, *A. niger* via heterologous protein expression, characterization and direct screening on lignocellulosic biomass, an alternative approach was employed to develop biomass-liquefying enzymes from the secretome of thermophilic saprophytic fungi. *M. thermophila* was chosen for evaluation on a variety of pretreated wood biomass types as discussed in Manuscript IV.

Chapter 5 Manuscript IV

Myceliophthora thermophila as a source of proteins to enhance liquefaction of high consistency pretreated wood biomass

Enongene Ekwe[1, 2], Marcos Di Falco[1], Adrian Tsang[1, 3] and Justin Powlowski[1, 2]*
[1]Centre for Structural and Functional Genomics, [2]Department of Chemistry and Biochemistry, and [3]Department of Biology, Concordia University, 7141 Sherbrooke Street West, Montreal, QC H4B 1R6

*Corresponding author: Justin Powlowski Justin.Powlowski@concordia.ca

Manuscript

Abstract

Background

Temperatures above 50°C are advantageous for liquefaction of lignocellulosic feedstock at loadings in excess of the 15 % needed for commercially viable bioprocesses. However, commercial cellulases are labile at temperatures higher than 50°C and show reduced efficiency at high feedstock loadings. Fungal proteins that accelerate liquefaction of high loadings of feedstock at elevated temperatures could significantly impact the usability of lignocellulose feedstocks for bioprocesses.

Results

In this study the thermophilic saprotroph, *Myceliophthora thermophila* strain ATCC 42464, was grown on a variety of pretreated lignocellulosic biomass as sole carbon source, namely: organosolv pretreated lodgepole pine, organosolv pretreated poplar, acid catalysed steam exploded poplar, aspen kraft pulp, black spruce kraft pulp and aspen thermomechanical pulp. Unlike the steam and thermomechanically pretreated biomass types, the organosolv- and kraft- treated

biomass had higher cellulose, lower lignin and lower hemicellulose content which favoured the growth of *M. thermophila*. The secreted protein cocktails, found to contain up to 47 proteins, were capable of liquefaction of various pretreated biomass types at 15 % solids loading.

A highly-expressed GH7 cellobiohydrolase, *Mt*CBH7, the major protein detected from the secretome, was cloned, expressed in *Aspergillus niger*, biochemically characterized and tested for viscosity reduction of 15 % black spruce kraft pulp via custom-designed, roughened parallel plates. Although the enzyme showed low activity on various model substrates, it was capable of liquefaction of 15 % black spruce kraft pulp and boosted Celluclast hydrolysis of the same substrate at low protein loading.

Conclusions

This study provides insight on the group of proteins in the secretome of *M. thermophila* that concertedly liquefy 15 % black spruce kraft pulp. Although other *M. thermophila* proteins appeared to be involved, the results revealed that *Mt*CBH7 played a role in the liquefaction and enhancement of cellulase saccharification of 15 % BSKP, acting optimally at 55 °C. These findings make *Mt*CBH7 a promising enzyme candidate for application in biorefineries where enhanced liquefaction of lignocellulose is required, especially at temperatures above 50 °C.

Key words: Pretreated biomass, high-solids, liquefaction, thermotolerant, secretome, *Myceliophthora thermophila*, cellobiohydrolase, mass spectrometry, rheology

5.1 Introduction

Alternatives to petroleum are being sought to provide cleaner and more sustainable sources of energy and materials. Lignocellulosic biomass is widely considered a sustainable resource for renewable liquid transportation fuels,

materials and chemicals that ideally would be produced in a biorefinery setting [396,397]. Efficient, mild depolymerization of plant cell wall structural components is central to the realization of biorefineries. Among the several challenges facing commercially-viable bioconversion of lignocellulose to sugars are the high enzyme cost [205] and the need to develop thermostable biocatalysts with improved potential to accelerate solubilization of high-solids feedstock at elevated temperatures [246,247].

Commercial cellulases, sourced almost exclusively from mesophilic fungal species such as *Trichoderma reesei* and *Aspergillus niger*, are most effective at 50°C or below [241,242,246] and require improvement to make bioconversion of lignocellulose economically viable. For instance, a minimum ethanol concentration of 4% is required in fermentation broth for economy of scale, which in turn dictates that the concentration of starting lignocellulosic biomass be in excess of 15% dry solids [244,245]. Many studies have reported enzymatic conversion of lignocellulosic biomass at solids loadings of 15 – 40% [243,267,398,399]. Such high solid loadings benefit from treatment at temperatures above 50°C for: accelerated solubilization of solids and reduction of bulk viscosity (liquefaction) via favourable changes in the structure of feedstock material; improved mass transfer; potential increase in reaction rates of thermotolerant biocatalysts and shortening of reaction time [118,246]; and reduced process cost due to reduced water input in lignocellulose bioconversion [243].

The genus *Myceliophthora,* which phylogenetically includes six mesophilic and four thermophilic species, has emerged as a taxonomic group that attracts considerable research interest as a source of efficient thermophilic lignocellulolytic biocatalysts [258,400]. Its thermophilic species, *M. heterothallica, M. thermophila, M. hinnulea and M. fergussi*, are known to hydrolyse and grow efficiently on lignocellulosic biomass, secreting enzymes with temperature optima up to 70°C [241,256,257]. In addition, *M. thermophila*, reported to grow on cellulose at the same rate it grows on glucose, is known to secret a GH7 endoglucanase that liquefies hydrothermally pretreated wheat straw at 18% dry solids [259]. The genome of *M. thermophila* has been sequenced and is publicly available for the

research community to exploit for the further development of efficient industrial enzymes for lignocellulose degradation [118].

Enzymatic liquefaction of lignocellulosic biomass, thought to be caused by random desizing of constituent long chain polysaccharides, leads to a partial collapse of the structured biomass matrix, which in turn results in reduced porosity and water retention potential of the material [267]. These changes enhance the flowability of the feedstock, which can be measured by rheometry. Although several research groups as well as pilot and demonstration facilities are testing bioconversion technologies of lignocellulose at high-solid loadings [401], the evaluation of rheological properties of lignocellulosic biomass slurries remains a significant technical challenge. Rheometry of dilute acid pretreated corn stover slurries concluded that the parallel plate rheometer with roughened surfaces was most suited for measuring nearly all of the rheological properties of corn stover [402]. Other rheological studies have been reported on sawdust [403], barley straw [404] and wood pulp slurries [405,406]. These studies concluded that lignocellulosic biomass slurry can be described as heterogeneous, non-newtonian, viscoelastic fluid with particle sizes and shapes that vary widely.

In this study, *M. thermophile* was grown on various pre-treated wood substrates and its protein secretome were examined by mass spectrometry. Also, the ability of crude secretome preparation and purified enzymes to liquefy high-solids slurries was examined. Although a GH7 exoglucanase was a major factor in liquefaction of these samples, it was not the only one.

5.2 Results

Biomass composition

The cellulose, xylan and lignin contents of the eight biomass samples used in this study were as shown in Figure 22. In decreasing order of cellulose composition, the biomass categories were as follows: kraft pulps > organosolv > steam exploded > mechanical pulp biomass. Steam exploded biomass contained the highest percentage of lignin while the xylan content was highest in the mechanical pulp.

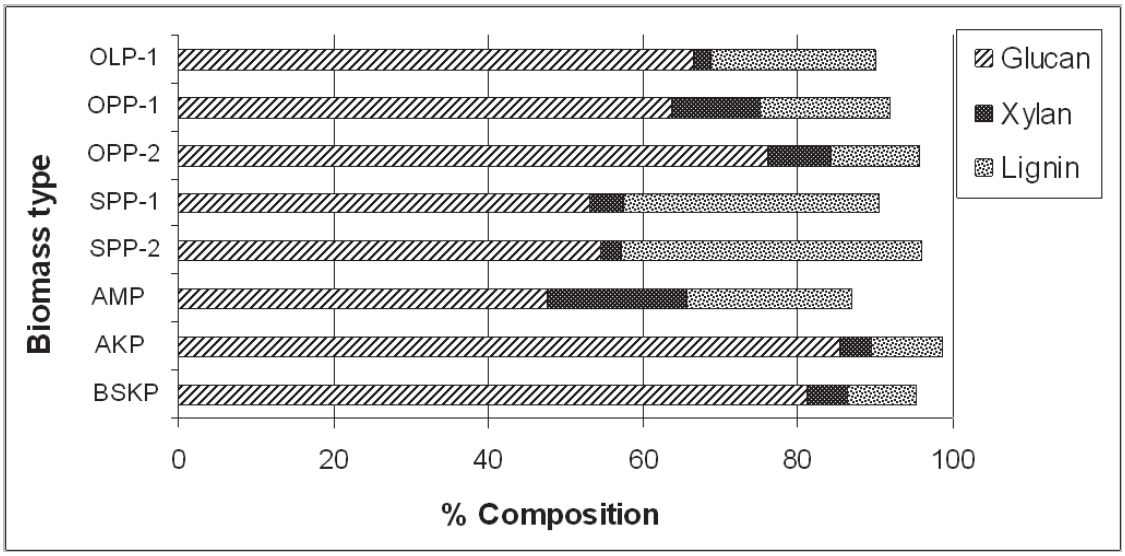


Figure 22. Cellulose, xylan and lignin content of biomass used in this study. OLP-1: Organosolv pretreated lodgepole pine at; OPP-1: organosolv poplar pretreated poplar at 160 °C; OPP-2: organosolv pretreated poplar at 175 °C; SPP-1: steam pretreated poplar at 200 °C. SPP-2; steam pretreated poplar at 220 °C; AMP: aspen mechanical pulp; AKP: aspen kraft pulp; BSKP: black spruce kraft pulp

Enzyme assays

Supernatants from cultures of *M. thermophila* strain ATCC 42464 grown on a variety of pretreated lignocellulosic biomass samples were assayed for cellulase, xylanase, mannanase and pectinase activities, and activity levels compared with the corresponding carbohydrate polymers in the culture carbon source. The cellulase activity in all samples was similar and did not correlate directly with the cellulose content of the culture carbon source (Figure 23a). Similarly, there was no apparent dependence of the mannanase and pectinase activities on the mannan and galactan contents of the biomass, respectively. It must be noted though, that the amounts of mannan and galactan were lower than 3.8% and 1.8% respectively (results not shown). However, the xylanase specific activity was directly correlated with the

xylan content in the carbon source (Figure 23b). No xylanase activity was detected in SPP-2.

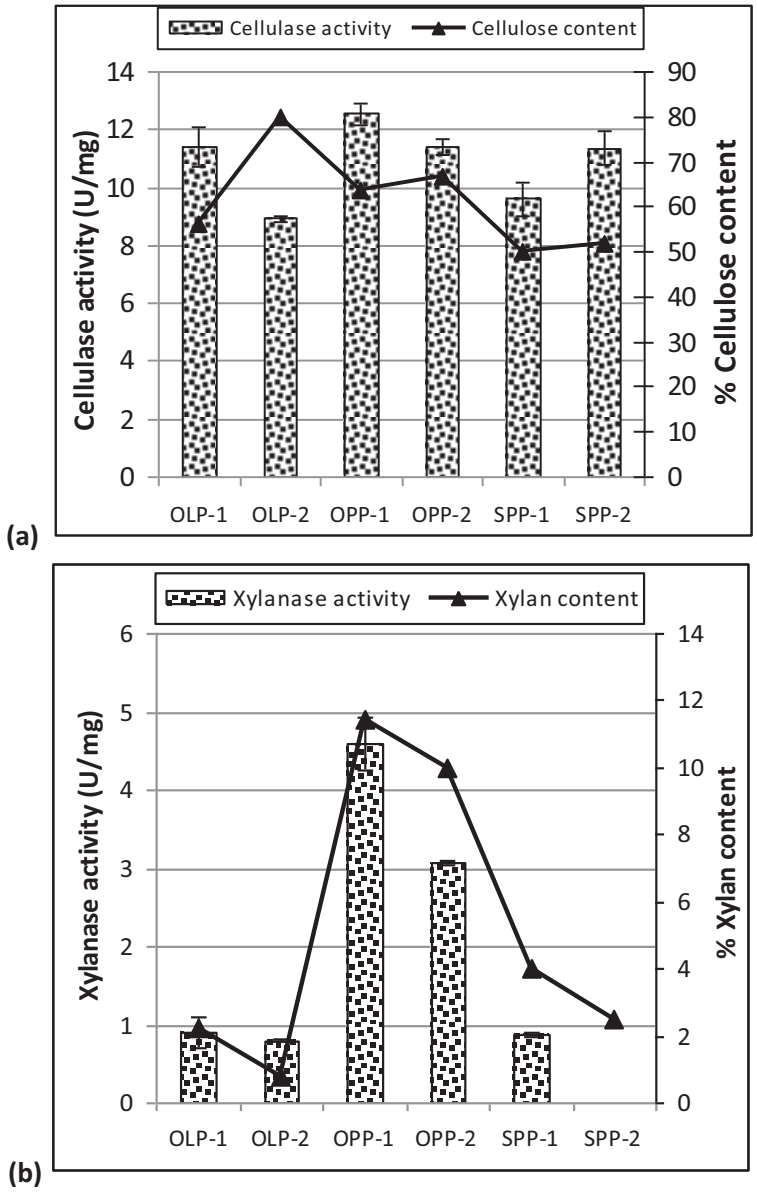


Figure 23. Relationship between (a) cellulase activity of *M. thermophila* culture supernatant and cellulose content and (b) xylanase activity of *M. thermophila* culture supernatant and xylan content of lignocellulose carbon source

The total protein concentrations of secretomes from cultures grown on the organosolv pretreated biomasses were two-fold or greater than those produced from steam exploded biomasses. Overall, the total amount of secreted protein, in decreasing order, was; organosolv pretreated hardwood > organosolv pretreated softwood >hardwood kraft pulp > softwood kraft pulp > steam exploded hardwood > hardwood mechanical pulp > softwood mechanical pulp (data not shown).

Biomass hydrolysis using *M. thermophila* supernatant

Supernatants from cultures of *M. thermophila* strain ATCC 42464 grown on variety of pretreated lignocellulosic biomass were tested for the liquefaction of 15% biomass at enzyme loading 5 mg/g DM. In this test, supernatant from cultures grown on a specific biomass type was used in the hydrolysis of the same type of biomass. Glucose release from the various biomass samples by *M. thermophila* supernatant was compared to that released by Celluclast under the same conditions (Figure 24). Biomass without enzyme was used as control and all background values were subtracted from the corresponding values of reactions with enzyme. The trend in glucose yield by *M. thermophila* supernatant was similar to that of Celluclast, although in all cases the yields from commercial enzyme were higher. With the exception of OPP-2 for which the amount of glucose released by *M. thermophila* was just 19% of the amount released by Celluclast, each of the other tested biomasses were hydrolyzed by *M. thermophila* supernatant releasing glucose equivalent to 60 – 83% of the amount released by the commercial cellulase. An inverse linear correlation was observed between the amount of glucose released by Celluclast or *M. thermophila* supernatant and the lignin content of biomass (Figure 24).

Enhancement of Celluclast glucose release by *M. thermophila* supernatants was also tested. As shown in Figure 25, the highest degree of enhancement of Celluclast activity was approximately 30% for OPP-2 and BSKP, followed by 13% and 19% for OPP-1 and AMP, respectively, over a 72 h period. The time course for the

OPP-2 and BSKP hydrolysis are shown in Figure 26. No significant improvement was observed for OLP1 and AKP, and minimal to no enhancement was observed for the two SPP substrates (results not shown).

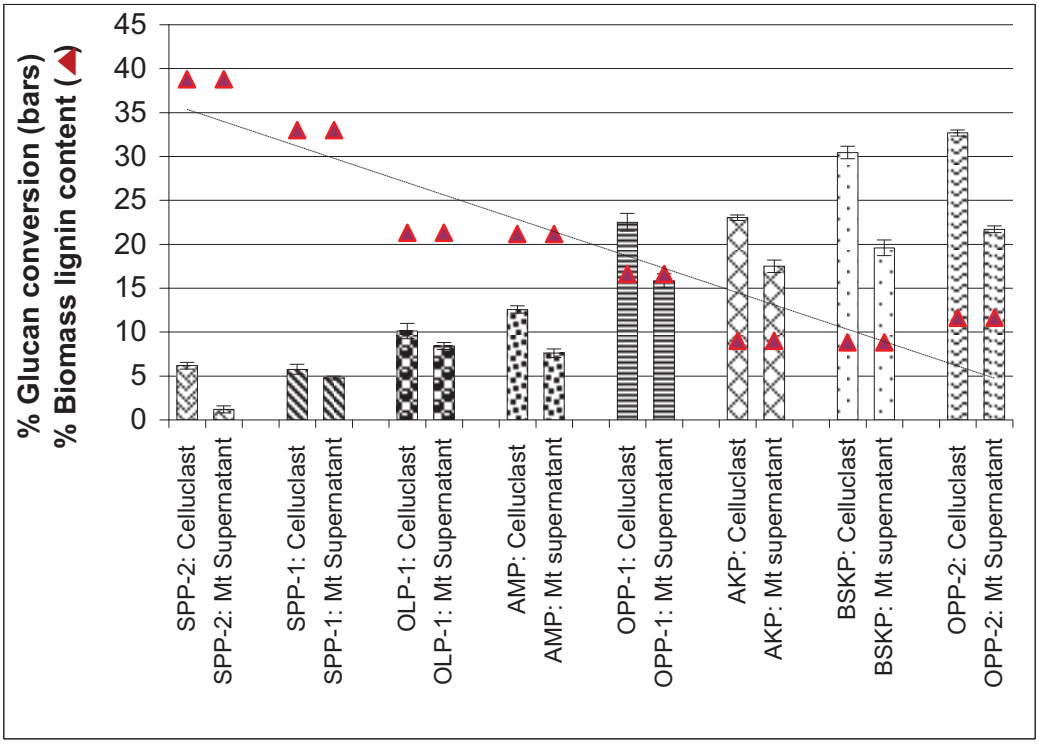


Figure 24. Comparison of glucose released by *M. thermophila* culture supernatant and that released by Celluclast on various substrates at 15% dry solids at 50 °C for 72h. Note the illustrated relationship between lignin content of biomass and enzymatic hydrolysability. Bars with identical pattern represent the same substrate. Dotted line represent linear trend in biomass lignin content.

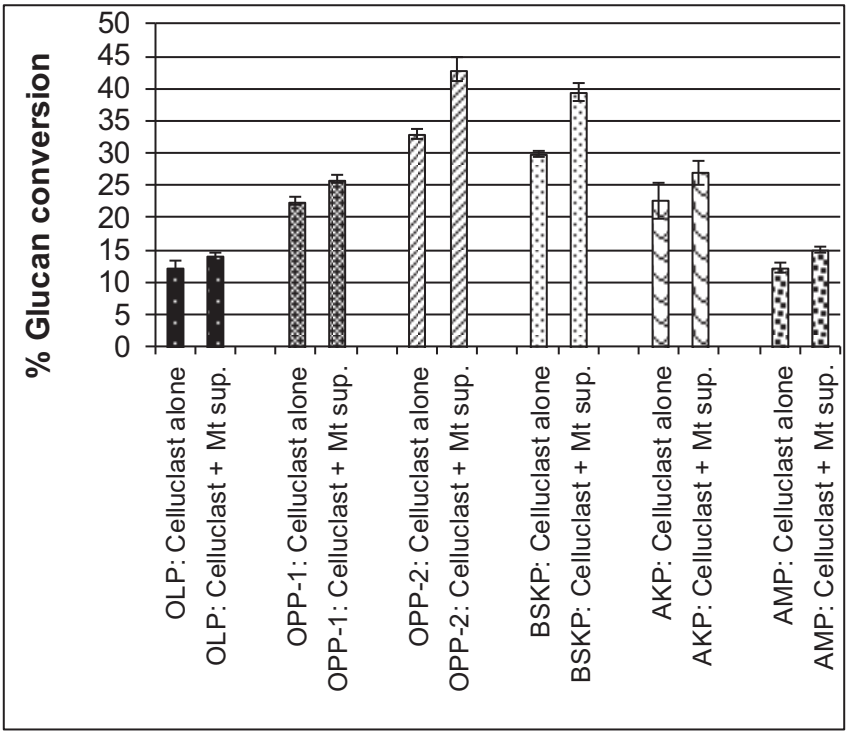


Figure 25. Enhancement of Celluclast in the release of glucose from from various pretreatred biomass substrates by *M. thermophila* secreted protein cocktail at 50 °C after 72 h. Substrate varieties are differentiated by pattern in the bar charts. Mt supernatant denotes *M. thermophila* supernatant. Supernatant used to supplement Celluclast in each case was obtained from culture containing the corresponding substrate as sole carbon source. Total protein loading: 5 mg/g DM. Celluclast and *M. thermophila* supernatant were mixed at a ratio of 9/1.

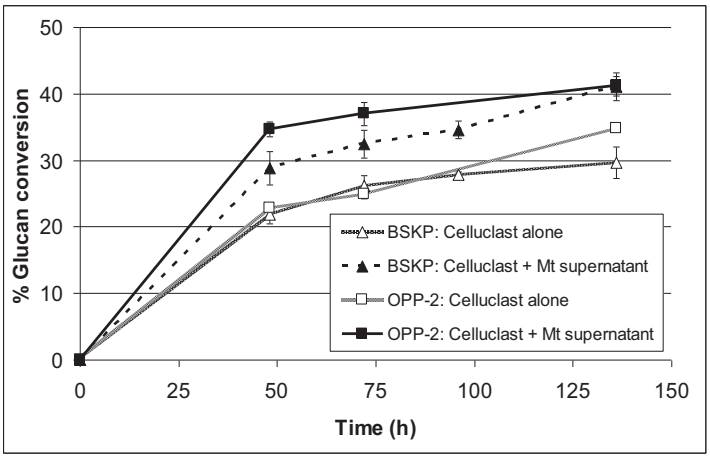


Figure 26. Enhancement of Celluclast in the release of glucose from 15 % BSKP and OPP-2 *by M. thermophila* secreted protein cocktail at 50 °C during a 136 h reaction. Total protein loading: 5 mg/g DM. Celluclast and *M. thermophila* supernatant were mixed at a ratio of 9/1.

SDS-PAGE and protein sequence analysis

M. thermophila culture supernatants grown on the various pretreated biomass types showed numerous protein bands on SDS-PAGE gels (Figure 27a). Some of these proteins are shown in Table 10. Some differences were observed in terms of proteins induced and their relative abundance from one substrate to another, shown as presence and absence of bands and band intensities. The prominent band at 72.2 kDa in supernatants from organoslv- and kraft- treated biomass (Figure 27a) represents the GH7 cellobiohydrolase, *Mt*CBH7 detected in the pooled sample, in which it accounted for >25% of the total by mass on SDS-PAGE gel. The mechanical pulps and steam pretreated substrates supported the production of very low amounts of protein due to low fungal growth as observed by eye.

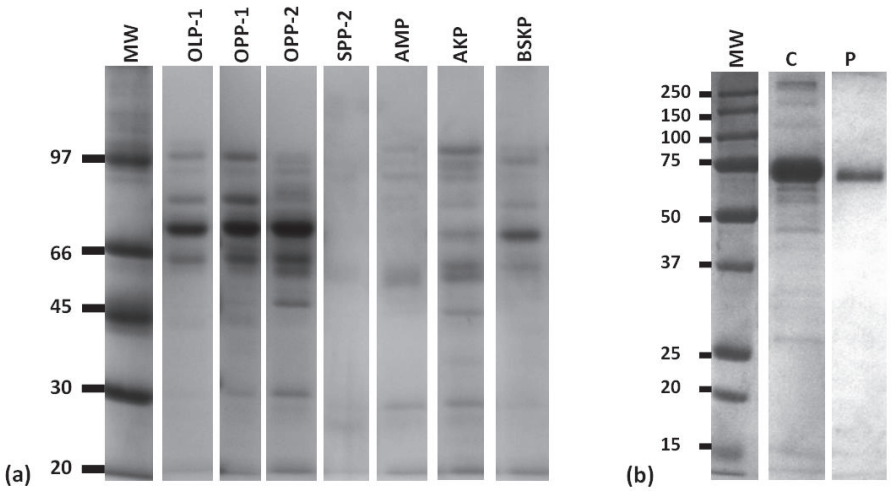


Figure 27. SDS-PAGE gel of: (a) *M. thermophila* culture supernatant grown on a variety of pretreated biomass types showing differential protein expression and (b) Recombinant *Mt*CBH7 expressed in *A. niger.* Sample volume loaded was 15 µl for each of the *M. thermophila* culture supernatants in (a) and 10 µl of crude and purified recombinant *Mt*CBH7. Lane 'C' represents crude *A. niger* culture supernatant, showing intense band of recombinant *Mt*CBH7; lane 'P' represents purified recombinant *Mt*CBH7 shown as a single band. MW denotes molecular weight marker.

Table 10. Protein composition of *M. thermophila* culture supernatant Fractions 2 and 3. The relative abundance was calculated as described in reference [420]

Fraction 2			Fraction 3		
Predicted protein activity	JGI protein ID	% Relative abundance	Predicted protein activity	JGI protein ID	% Relative abundance
Proteins with signal peptide					
Arabinoxylan arabinofuranosidase (AXH62A)	98003	2.9	Arabinoxylan arabinofuranohydrolase (AXH43G)	2303298	3.3
Cellobiohydrolase (CBH6A)	66729	22.7	β-glucoronidase (GUSB2)	114497	6.1
Cellobiohydrolase (CBH7A)	33936	8.1	Cellobiohydrolase (CBH7)	109566	37.4
Cuticle-degrading serine protease	2303011	1.9	Cellobiose dehydrogenase (CDH)	111388	3.5
Exo-β-galactanase (GAX43A)	50820	2.4	Endo-β-glucanase (EGL16)	90182	10.9
Polysaccharide monooxygenase (PMO)	80312	2.4	Endoglucanase (EGL6B)	2303045	14.3
Unknown	73270	1.7	Polysaccharide monooxygenase (PMO)	80312	6.5
Unknown	102138	14.8	Polysaccharide monooxygenase (PMO)	92668	7.9
Xylanase (XYN10D)	112050	13			
Xylanase (XYN10)	52904	1.8			
Xylanase (XYN11)	100068	4.2			
Xylanase (XYN11)	89603	1.5			
Proteins without signal peptide					
Exo-β-glucanase (EXG17B)	2315007	4.9	Endochitinase (CHI18A)	2308241	10.1
Superoxide dismutase (SODC)	2297816	17.7			

Proteins responsible for Celluclast enhancement and biomass liquefaction

In-solution digestion of the pooled supernatant and analysis by mass spectrometry revealed a repertoire of 42 secreted proteins, most of which are involved in plant cell wall deconstruction (result not shown). In order to identify protein targets that may be responsible for the Celluclast enhancement effect, pooled *M. thermophila* supernatant from cultures containing OLP1, OPP1, OPP2, SPP1, SPP2, AMP, AKP or BSKP as sole carbon source was fractionated by anion exchange chromatography. Each fraction was then tested for the enhancement of Celluclast hydrolysis of BSKP as shown in Figure 28. BSKP was selected because it was readily available and showed the highest increase in glucose yield (30%) when Celluclast was supplemented with *M. thermophila* supernatant at high-solids loading (Figure 25 and Figure 26). Fractions 2 and 3 showed improvement of Celluclast in the release of glucose from BSKP by 25 % and 27 %, respectively (Figure 28) after 48 h, which is equal to the effect of non-fractionated *M. thermophila* supernatant. However, after 96h the enhancing effect of non-fractionated *M. thermophila* supernatant was 46 %, the highest value observed, which surpassed the effects of fractions 2 and 3 by at least 20 % (Figure 28). No significant enhancement was observed for fractions 1, 4 and 5.

Analysis of fractions 2 and 3 by mass spectrometry revealed the presence of 14 and 9 proteins, respectively. Of these, 12 and 8, respectively, were predicted to have signal peptides for secretion (Table 10). The enhancing effects of both fractions were very similar. In this study, we focused on Fraction 3 which contained fewer proteins, with *Mt*CBH7 being the dominant constituent by mass as observed on 12 % SDS-PAGE gel (result not shown) and by spectral count in mass spectrometry. Among the other cellulose-active proteins present in Fraction 3 were EGL, PMO and CDH, which together amounted to 34 % of the total protein. With *Mt*CBH7 representing 37.4 % of the protein (determined by emPAI calculation), more than two and half times any of the other 8 constituents, this protein was cloned, characterized and its Celluclast-enhancing effect for glucose release from BSKP tested.

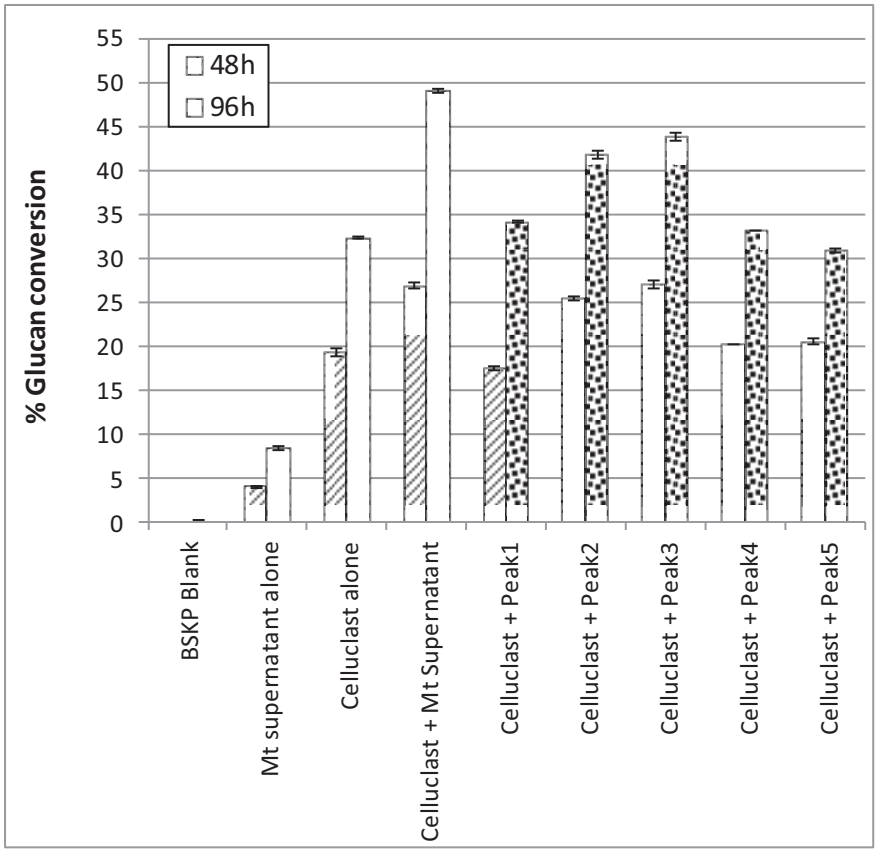


Figure 28. Evaluation of cellulase enhancement of glucose release from 15% BSKP by fractions of *M. thermophila* culture supernatant separated by anion exchange chromatography. Enzyme loading of 5 mg/g DM was used at 50 °C, pH 5.5. Each fraction, represented by peak# (where # is a number) was use to supplement the benchmark cellulase, Celluclast. Celluclast and *M. thermophila* supernatant or *M. thermophila* supernatant fractions were mixed at a ratio of 9/1.

Characterization of recombinant *Mt*CBH7

The ORF of *Mt*cbh7 encodes a 509 amino acid mature polypeptide, which is linked on the N-terminal to an extracellular secretory signal peptide of 17 amino acids (MYAKFATLAALVAGAAA), as predicted by SignalP v4.1. The mass and pI of the mature protein were 55,620 Da and pH 4.82, respectively. One N-glycosylation site

(Asp288) and 19 O-glycosylation sites were predicted using NetNGlyc v1.0 and NetOGlyc v4.0, respectively.

The SDS-PAGE mass of the native and recombinant *Mt*CBH7 were about the same. Native *Mt*CBH7 seen as the heaviest band in Figure 27a was 72.2 kDa in mass. Purified recombinant *Mt*CBH7 revealed a single protein band of 69.8 kDa on 12% SDS-PAGE gel (Figure 27b), which was confirmed by mass spectrometry to have the peptide fingerprint of native *Mt*CBH7. The protein stained positively on glycol protein staining (result not shown).

Recombinant *Mt*CBH7 was assayed on the substrates listed in Table 11 to investigate its catalytic specificity. Although the amino acid sequence of this protein matches a GH7 cellobiohydrolase, it showed very low cellobiohydrolase activity on 4-methylumbelliferyl cellobioside, 4-nitrophenyl-β-D-cellobioside, 4-nitrophenyl-β-D-lactoside and 4-nitrophenyl-β-D-glucoside. Similarly, its endoglucanase, xylanase, pectinase, pullulanase, laminarinase, avicelase and filter paper activities were very low. The protein did not show mannanase or galactomananase activity on locust bean gum.

The effects of pH and temperature on the activity of *Mt*CBH7 were evaluated using 4-methylumbelliferyl cellobioside. The pH and temperature optima were 5.5 and 55 °C, respectively (Figure 29a and Figure 30b). 80 % or more of the maximum activity was observed at pH 5 – 6, and temperatures from 50 – 60 °C. Above pH 6 or at a temperature of 60 °C, the activity fell sharply to below 30 % of the maximum. The protein reduced the dynamic viscosity of 0.5% CMC-4M/7M by 37 % in 10 min, compared to 87 % for the pooled *M. thermophila* supernatant (Figure 29c) at the same protein loading.

The effect of temperature on the hydrolysis of BSKP by purified *Mt*CBH7 is shown in Figure 30a. The highest amount of glucose was released at 55 °C. A similar pattern was observed for *M. thermophila* supernatant (Figure 30b). For both protein samples glucose released at 55 °C was at least 18% higher than at 50 °C or 60 °C.

Table 11. Activity of *M. thermophila* culture supernatant and purified *Mt*CBH7 on a variety of substrates at 50 °C.

'ND' denotes not detected; '–' denotes not tested

Substrate	Substrate concentration	Assay method	Activity (U/mg protein)	
			*Mt*CBH7	*Mt* Supernatant
4-Methylumbelliferyl cellobioside	1mM	4-Methylumbelliferone	0.10 ± 0.01	-
4-Nitrophenyl-β-D-cellobioside	1mM	*p*-Nitrophenol	0.09 ± 0.02	-
4-Nitrophenyl-β-D-lactoside	1mM	*p*-Nitrophenol	0.17 ± 0.01	-
4-Nitrophenyl-β-D-glucoside	1mM	*p*-Nitrophenol	0.32 ± 0.03	3.5 ± 0.2
Filter paper	3.2mg	Dinitrosalicylic acid	1.3×10^{-4}	2.7×10^{-3}
Avicel	0.2 %	BCA reducing sugar	0.04 ± 0.01	-
CMC 4M/7M	1%	BCA reducing sugar	0.42 ± 0.07	-
Birchwood xylan	1%	BCA reducing sugar	0.07 ± 0.02	-
Locust bean gum	0.4%	BCA reducing sugar	ND	-
Laminarin	1%	BCA reducing sugar	0.76 ± 0.07	-
Pullulan	1%	BCA reducing sugar	0.02 ± 0.00	-
Polygalacturonic acid	1%	BCA reducing sugar	0.15 ± 0.01	-

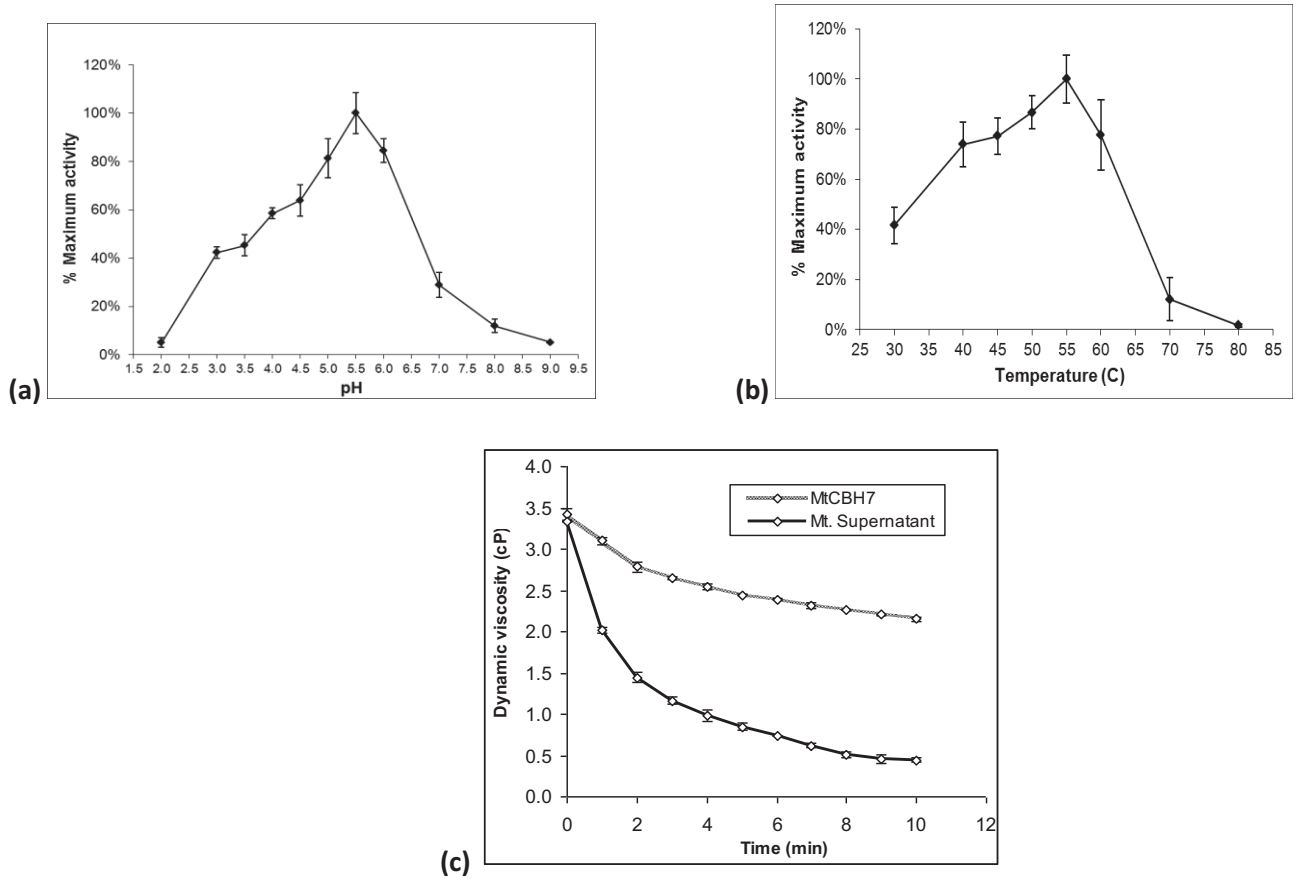


Figure 29. Effect of (a) pH and (b) temperature on the activity of MtCBH7 on 4-methylumbelliferyl cellobioside. Reduction of dynamic viscosity (cP) of 0.5% CMC 4M/7M by MtCBH7 is shown in (c). In this test, 50 µg enzymes was added to 10 ml of 0.5% CMC 4M/7M

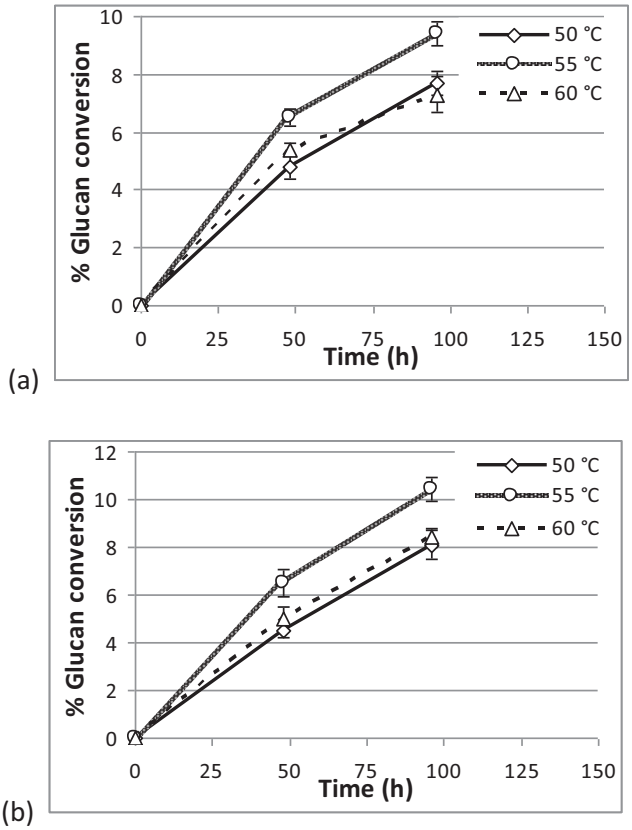


Figure 30. Effect of temperature on the hydrolysis of BSKP by (a) *Mt*CBH7 at total protein loading of 15 mg/g DM and (b) Pooled *M. thermophila* culture supernatant at total protein loading of 5 mg/g DM. *M. thermophila* supernatant represent pooled supernatant from *M. thermophila* cultures grown on various pretreated wood biomass as sole carbon source.

Cellulase enhancement by recombinant *Mt*CBH7

Results from trials with various Celluclast/recombinant *Mt*CBH7 combinatorial ratios are shown in Figure 31. At a protein loading of 5 mg/g DM, glucose yields by *Mt*CBH7 alone were more than 7 times lower than those from Celluclast alone at the 48 h and 84 h timepoints. Celluclast enhancement was only observed at Celluclast/*Mt*CBH7 combinatorial ratios of: 4:1, 7:3 and 3:2 after 48 h, which correspond to degrees of enhancement of 18 %, 26 % and 20 %, respectively.

However, after 84 h, enhancement was only observed at a Celluclast/*Mt*CBH7 ratio of 7:3, equivalent to 11 % increase in glucose yield. At the same total protein loading, the degree of Celluclast enhancement by purified *Mt*CBH7 (26 % gain in glucose release at 7:3 combinatorial ratio – Figure 31) was lower than by non-fractionated *M. thermophila* culture supernatant and supernatant Fractions 3, each of which showed 35 % gain in glucose release (Figure 26) after 48 h.

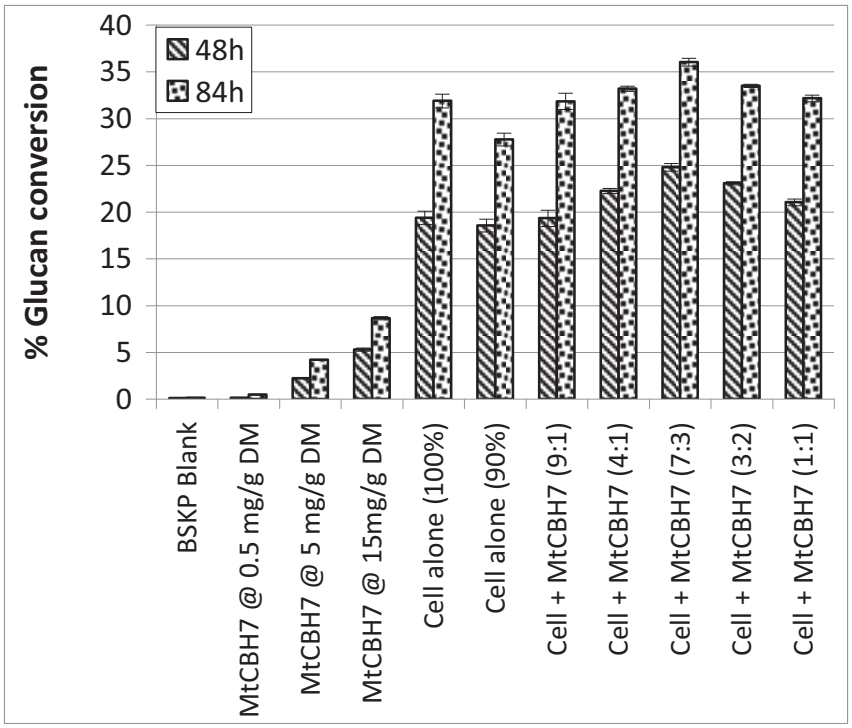


Figure 31. Evaluation of cellulase enhancement of glucose release from 15% BSKP by purified recombinant *Mt*CBH7. Unless otherwise stated; the enzyme loading was 5 mg/g DM. Hydrolysis was done at 50 °C and pH 5.5. Cell denotes Celluclast; 100 % denotes enzyme loading of 5 mg/g DM; 90 % denotes enzyme loading of 4.5 mg/g DM; 9:1 denotes the Celluclast/MtCBH7 combination ratio

Considering the number and concentration of sugars released, the pooled *M. thermophila* supernatant was more effective than purified *Mt*CBH7 in the hydrolysis of BSKP and the relatively more complex substrate AMP (Table 12). Particularly, arabinose, cellotriose and xylobiose were only detected in reactions containing *M. thermophila* supernatant. The carbohydrate conversion of BSKP was higher than of AMP for both protein samples.

Table 12. Sugars and oligosaccharides released by pooled *M. thermophila* culture supernatant and *Mt*CBH7 in the hydrolysis of 15 % BSKP and AMP at 55 °C for 84 h*

	% Conversion			
	BSKP		AMP	
Carbohydrate	*M. thermophila* supernatant	*Mt*CBH7	*M. thermophila* supernatant	*Mt*CBH7
Glucose	22.3 ± 1.76	9.8 ± 0.72	8.8 ± 0.91	3.1 ± 0.11
Xylose	1.7 ± 0.01	0.4 ± 0.03	3.3 ± 0.12	0.7 ± 0.03
Mannose	ND	ND	1.7 ± 0.09	ND
Arabinose	ND	ND	0.1 ± 0.01	ND
Galacturonic acid	0.1 ± 0.04	ND	0.3 ± 0.01	0.1 ± 0.01
Glucoronic acid	ND	ND	ND	ND
Cellobiose	1.7 ± 0.61	0.3 ± 0.02	1.1 ± 0.06	0.2 ± 0.01
Cellotriose	0.1 ± 0.01	ND	0.3 ± 0.02	ND
Xylobiose	0.1 ± 0.03	ND	ND	ND
Xylotriose	ND	ND	ND	ND

* Total protein loading was 5 mg/g DM for the pooled M. thermophila culture supernatant and 15 mg/g DM for purified MtCBH7

Fibre morphology

Following treatment with *Mt*CBH7, BSKP fibres were noticeably reduced in length after 48 h as seen in the low magnification images in Figure 32a. The thickness was reduced and the surface roughness increased (Figure 32b). Treatment of the substrate with pooled *M. thermophila* culture supernatant was more effective in than by *Mt*CBH7 as shown in Figure 32b.

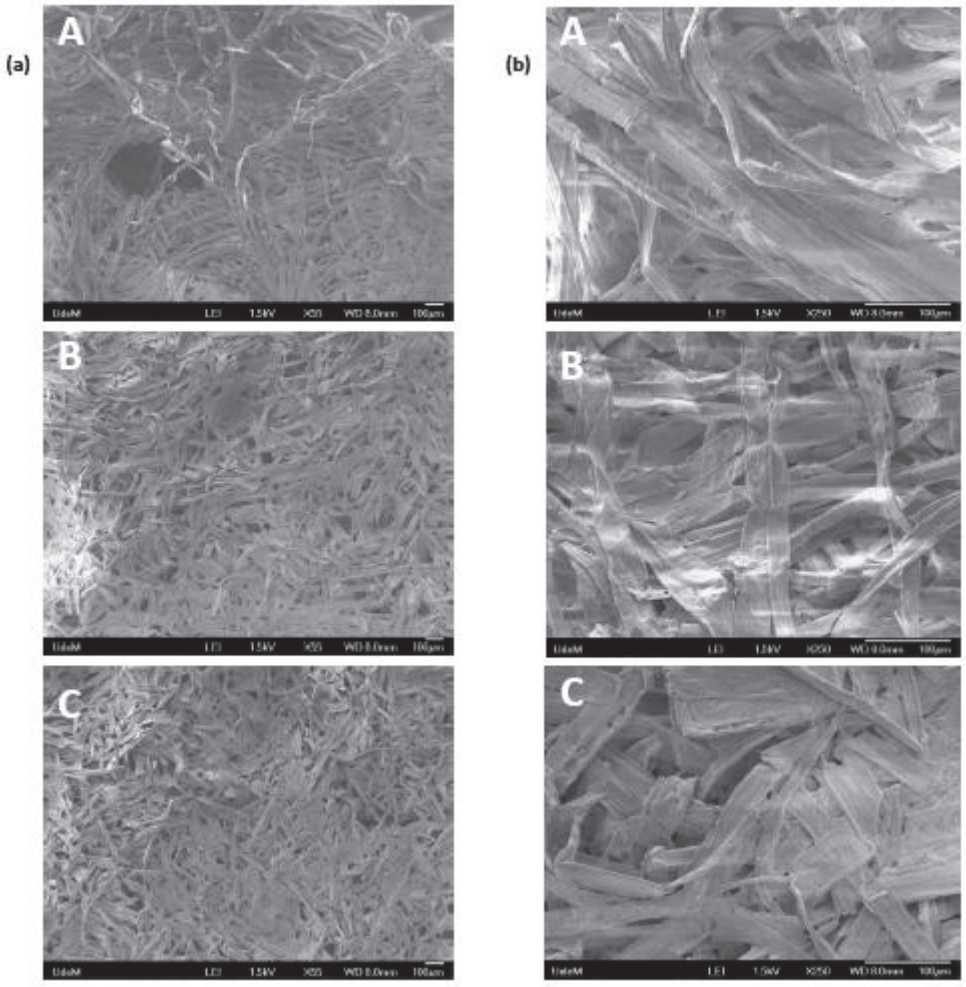


Figure 32. Fibre surface morphology of BSKP (A) before and (B) after treatment with *Mt*CBH7 for 48 h at total enzyme loading of 15 mg/g DM and (C) after treatment with pooled *M. thermophila* supernatant at total enzyme loading of 5 mg/g DM. Images are magnified (a) 55x and (b) 250x.

Rheology

In addition to solubilization of the biomass to release glucose, an increase in liquefaction of the 15 % biomass slurry was visually evident in vials containing the pretreated biomass to which *M. thermophila* culture supernatant was added, compared with those without enzyme, as illustrated in Figure 33 for OPP-2, BSKP and SPP-1. Similar results were obtained for OPP-1, OLP-2, AKP and AMP (not shown). No visible increase in liquifaction was observed for AMP and SPP-2. These rheological changes are not quantitative. Thus, the rheological effects of *M. thermophila* supernatant or purified *Mt*CBH7 on biomass were further examined using a parallel plate rotational rheometer on 15 % BSKP after enzyme treatment at 50 °C – 60 °C. BSKP was chosen because it was one of two substrates (with OPP2) on which *M. thermophila* supernatant produced the highest liquefaction effect as observed by eye (Figure 33).

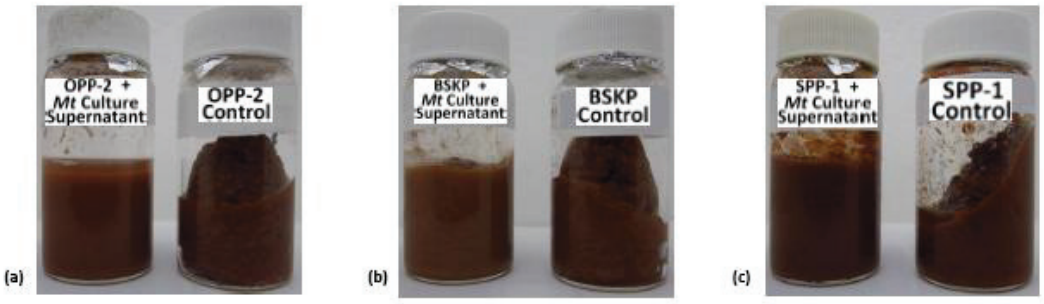


Figure 33. Examples of pretreated biomasses liquefied by *M. thermophila* supernatant after 72 h at 50 °C, pH 5.5 and 15 % dry solids. (a) OPP-2; (b) BSKP-2; (c) SPP-1; *M. thermophila* culture supernatant was loaded at 5 mg/g DM. No enzyme was added to control.

As shown in Figure 34a, b and c, the high complex viscosity (2,883 Pa.s) of the control BSKP sample without enzyme was unchanged after 96 h. *M. thermophila* culture supernatant alone reduced the complex viscosity by 1,603 Pa.s and 2,125 Pa.s after 48 h and 96 h, respectively (Figure 34a). Supplementation of Celluclast with *M. thermophila* supernatant (ratio 9:1) accelerated the reduction in complex viscosity by 2,796 Pa.s and 2,867 in 48 h and 96 h respectively. Interestingly, this extent of viscosity reduction was matched by Celluclast supplemented with *M. thermophila* supernatant Fraction 3. There was no significant improvement in complex viscosity reduction achieved by supplementing Celluclast with *M. thermophila* supernatant Fractions 1, 2, 4 or 5.

Supplementation of Celluclast with purified *Mt*CBH7 at various Celluclast/*Mt*CBH7 combinatorial ratios (Figure 34b) revealed that the most rapid reduction in the complex viscosity of BSKP by purified *Mt*CBH7 occurred at a ratio of 7:3. Combinatorial ratios of 9:1 and 1:1 were less effective. No significant difference in the extent of liquefaction at ratios of 4:1 and 3:2 was observed.

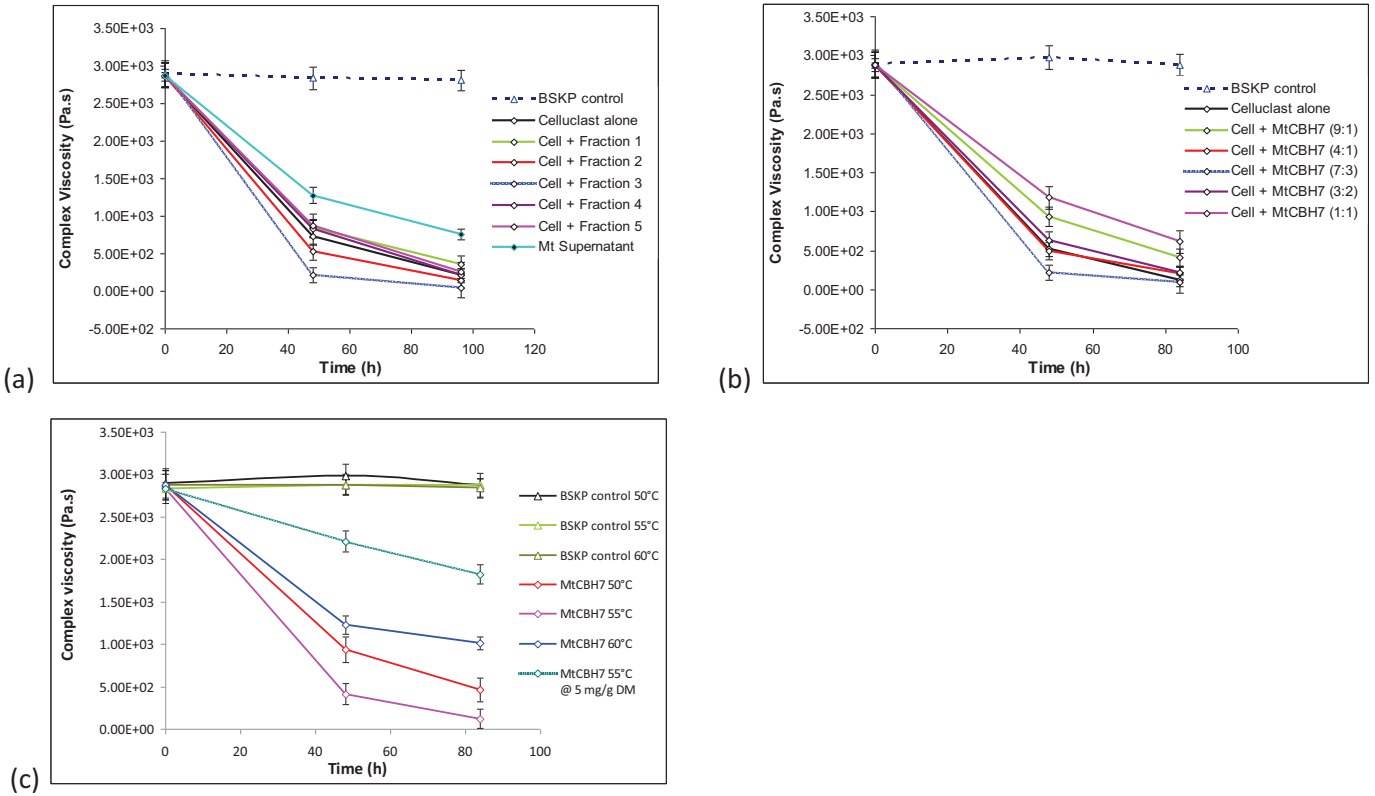


Figure 34. Liquefaction of BSKP by *M. thermophila* culture supernatant fractions and purified *M*tCBH7. (a) Effect of various *M. thermophila* culture supernatant fractions on Celluclast in the liquefaction of 15% BSKP at 50 °C, total protein loading of 5 mg/g DM and Celluclast/M. thermophila fraction ratio of 9:1. *M. thermophila* culture Supernatant was fractionated by anion exchange chromatography, denoted as Fractions 1-5. Cell denotes Celluclast; Mt supernatant denotes pooled *M. thermophila* supernatant from cultures grown on various pretreated wood biomass as sole carbon source. (b) Effect of purified *M*tCBH7 on Celluclast in the liquefaction of 15 % BSKP at 50 °C and 5 mg/g DM total protein loading. Celluclast/*M. thermophila* culture supernatant mixture ratios ranged between 9:1 and 1:1. (c) Direct liquefaction of BSKP by MtCBH7 at 50 – 60 °C and protein loading of 15 mg/g DM, except otherwise indicated.

The temperature dependence of complex viscosity reduction of BSKP by purified *Mt*CBH7 alone is shown in Figure 34c. Viscosity reduction was more effective at 55 °C than at 50 °C and 60 °C. Overall, Celluclast supplemented with pooled *M. thermophila* and *M. thermophila* supernatant Fraction 3 had higher liquefaction effect than Celluclast alone or Celluclast supplemented with purified *Mt*CBH7. The liquefaction potential of the former pair was almost three-fold that of *Mt*CBH7 alone. These results indicate that while *Mt*CBH7 plays an important role in liquefaction, other enzyme activities enhance this effect.

5.3 Discussion

In this study, *M. thermophila* strain ATCC 42464 was grown on eight different pretreated woody biomass types as sole carbon source to identify substrate-related differences in enzyme secretion. The secreted protein cocktail was tested for the liquefaction of high-consistency pretreated woody biomass and then fractionated and the biomass liquefaction effects of the fractions compared to that of non-fractionated protein cocktail in order to identify proteins responsible for viscosity reduction of high-consistency biomass. A GH7 cellobiohydrolase, *Mt*CBH7, identified as a potential candidate that plays a role in the liquefaction, was cloned, heterologously expressed, characterised and its biomass liquefaction effect tested by rheometry. This protein was also evaluated for its ability to enhance the release of glucose by Celluclast from high-consistency biomass.

M. thermophila strain ATCC 42464 was found to hydrolyse and grow efficiently on each of the following pretreated wood substrates: OLP-1, OLP-2, OPP-1, OPP-2, SPP-1, AKP, AMP and BSKP as sole carbon source. However, growth on SPP-2 and jackpine mechanical pulp – JMP (results not shown) – was found to be slower. The pH of SPP-2 was much lower (4.1) than the rest of the substrates (≥ 5.0), whereas *M. thermophila* prefers to grow at pH 5.5. Softwood mechanical pulps like JMP are known to be recalcitrance to enzymatic degradation due to high composition of lignin, which limits enzyme accessibility to the constituent carbohydrates.

Pooled culture supernatant from these substrates (excluding JMP) showed endoglucanase, xylanase, mannanase and pectinase activities that are relevant to the degradation of major constituent plant cell wall carbohydrates. This is consistent with previous reports that thermophilic *Myceliophthora sp.* grows efficiently on lignocellulosic substrates producing multiple cellulolytic, xylanolytic, mannanolytic and pectinolytic enzymes [117,257,407,408,409]. In this study, enzymes were identified in *M. thermophila* supernatant for the deconstruction of the major plant cell wall polysaccharides: endoglucanases, cellobiohydrolases, β-glucanases and PMOs for glucan; xylanases, xylosidases, xyloglucanase and arabinofuranosidases for xylan; β-mannanases for mannan; and pectate lyase, arabinogalactanase for galactan. The supernatant showed high cellulase, xylanase, mannanase and pectinase activities. As reported previously, the endoglucanase activities enabled a rapid reduction in viscosity of 0.5% CMC-4M/7M [259].

The amounts of cellulolytic, mannanolytic and pectinolytic activities in the culture supernatant did not show a direct correlation with the percentage of these components in the growth substrate. However, the xylanase activity correlated well with biomass xylan content, with high xylan content in substrates inducing high xylanase activity and low xylanase activity induced when xylan content was low (Figure 23). In previous studies, Badhan and co-workers concluded that the number (and in some cases, type) of detectable xylanase isoforms were different from one lignocellulosic substrate to another when *Myceliophthora sp.* was grown on various lignocellulosic substrates [408]. It is unclear whether differences in expression of specific xylanase genes influenced the observed correlation between xylanase activity and substrate xylan content in this study.

Higher total protein yield from cultures grown on the organosolv and kraft pretreated biomass relative to steam exploded biomass (results not shown) may be due to inhibitory compounds generated during steam explosion [16]. More severe pretreatment may significantly increase concentrations of such inhibitory compounds.

On the other hand, organosolv and Kraft samples were washed with distilled water following pretreatment to removed residual organic solvent and NaOH /Na2S, respectively which could interfere with microbial growth and enzyme catalytic function. Furthermore, it has been reported that kraft- and organosolv- pretreatment of wood biomass are more effective than steam explosion and thermo-mechanical comminution in the removal of the surrounding barrier of lignin and hemicelluloses to produce a cellulose-rich residue [16,410], with increased susceptibility to cellulase degradation [105].

Biochemical composition analysis of the eight biomass samples used in this study (Figure 22) confirmed that steam-exploded biomass contained the highest amount of lignin, while xylan was most abundant in mechanical pulp. The hydrolyzability of substrates by *M. thermophila* secreted proteins and Celluclast was strongly related to lignin content of the biomass (Figure 24) as reported previously [75,207].

As shown in Figure 25, *M. thermophila* culture supernatant enhanced Celluclast glucose release from 15 % OLP-1, OPP-1, OPP-2, BSKP, AKP and AMP by 17 – 30%. The enhancement of Celluclast hydrolysis of BSKP and OPP-2 by *M. thermophila* supernatant was sustained over 136 h, and had an optimum at 55 ºC (Figure 26). Protein mass spectrometry identified 47 secreted proteins (Supplementary Table 7), including three PMOs and a CDH. PMOs and CDHs are known to synergistically enhance cellulase hydrolysis of lignocellulose [136,163]. A GH7 endoglucanase previously reported to liquefy 18 % wheat straw [259], and a highly expressed CBH7 which, based on emPAI calculation, accounted for 14.5 % of the total protein (among the highest of all proteins in the cocktail) were also identified.

M. thermophila supernatant has previously been reported to boost Celluclast hydrolysis of lignocellulosic biomass [241]. Ion exchange chromatography to isolateprotein components responsible for Celluclast enhancement yielded 5 fractions, of which two were active (Figure 28). Each fraction contained different cellulose-degrading CBHs and PMOs as well as AXHs (Table 10), which hydrolyse arabinoxylans. Of

the twelve proteins identified in Fraction 2, nine are plant cell wall carbohydrate active, one is a protease and two are of unknown function, with the unknown proteins constituting more than 40 % of the total, as estimated by spectral count. CBHs, PMOs, CDH, XYNs, AXHs and other cell wall degrading auxiliary proteins such as those found in Fraction 2 have been reported to enhance Celluclase hydrolysis of lignocellulose by synergistic interaction with one another or with protein activities in Cellulase cocktail [134,135,136,137,175,176,177].

Studies that have previously reported liquefaction of high consistency lignocellulose by *M. thermophila* secreted enzyme cocktail used grasses (wheat and barley straw) as feedstock rather than woody biomass [241,243,411,259]. Here, the pooled *M. thermophila* supernatant was found to reduce the viscosity of all the woody substrates at 15 % consistency by more than 55 % after 48 h (Figure 6), with the exception of SPP-2 and AMP for which the viscosity was reduced by less than 10 % (rheology results not shown).

The ability of *M. thermophila* proteins to liquefy high solids loading of pretreated wood biomass was investigated further. Although both Fractions 2 and 3 from ion exchange chromatography enhanced of glucose release from BSKP, only Fraction 3 ignificantly enhanced Celluclast liquefaction of BSKP after 48 h (Figure 34). The GH7 endoglucanase reported previously to liquefy 18 % barley straw [259] was not found in either of these two fractions, but was identified in Fraction 5, which did not show any improvement of Celluclast liquefaction or glucose release from BSKP. None of the proteins identified in Fraction 3 has previously been reported to liquefy or enhance the liquefaction of high consistency lignocellulosic biomass. A major protein in Fraction 3 was *Mt*CBH7, which was examined further for its ability to liquefy 15 % BSKP.

Purified recombinant *Mt*CBH7 was found to release glucose from BSKP, with the optimum yield observed at 55 °C. Celluclast enhancement of glucose yield from BSKP was highest at Celluclast/*Mt*CBH7 combinatorial ratios of 7:3, where synergistic interaction between *Mt*CBH7 and Celluclastis optimal. Several studies have illustrated

the importance of combinatorial effects in optimizing enzyme cocktails [176,263]. The fall in degree of enhancement with time (Figure 31) may be influenced by the lack of mixing.

Analysis of sugars released by *M. thermophila* supernatant and *Mt*CBH7 from BSKP and AMP, reveal that glucose and xylose were the main products. The number of detected sugars and oligosaccharides released by *M. thermophila* supernatant was higher than those by *Mt*CBH7, as expected due to the multiple activities present in the former. The more diverse products from AMP compared to BSKP (Table 12) are likely due to the presence of higher amounts of hemicellulose in thermomechanically treated wood biomass.

Celluclast enhancement by *Mt*CBH7 of both the glucose release and liquefaction of BSKP peaked at a Celluclast/*Mt*CBH7 combinatorial ratio of 7:3 after 48 h, which suggests that hydrolysis was accompanied by liquefaction. This is consistent with previous reports that enzymatic liquefaction of lignocellulose potentiates hydrolysis [412]. The maximum BSKP hydrolysis and liquefaction effects of *Mt*CBH7 were observed at 55°C. The low hydrolysis and liquefaction effects observed at 60 °C indicates that the enzyme is limited in its ability to tolerate high temperature. As expected, reduction in viscosity by *Mt*CBH7 alone was lower than for *M. thermophila* supernatant, since the latter contains multiple supplementary enzymes suited for the deconstruction of various pretreated wood biomass types.

5.4 Conclusion

The future of lignocellulose bioconversion for bio-based commercial products is tied to high loadings of feedstock (≥15 %) and temperatures higher than 50 °C. Thermostable enzymes which liquefy and hydrolyse lignocellulose at high concentrations are necessary to achieve this goal. In this study the thermophilic saprotroph, *M. thermophila* strain ATCC 42464, grown on a variety of pretreated wood

biomass secreted complex mixtures of lignocellulolytic enzymes which concertedly liquefied 15 % pretreated wood biomass at 55 °C at low enzyme loading. The secreted proteins also enhanced Celluclast liquefaction and glucose release from a variety of pretreated wood biomass types at high consistency. Proteomic and chromatographic analysis of the secretome identified a highly expressed GH7 cellobiohydrolase (*Mt*CBH7) which was cloned and purified. *Mt*CBH7 was found to enhance Celluclast 1.5L in the release of glucose from 15 % organosolv pretreated poplar and BSKP. The recombinant protein was also found to reduce the viscosity of 15 % BSKP at 55 °C, although to a lower extent than the crude cocktail, indicating that it is only partially responsible for the observed liquefaction activity. Together, these results indicate that *Mt*CBH7 is a promising candidate for application in enhancing the hydrolysis of high loadings of pretreated wood biomass.

5.5 Materials and methods

Substrates, commercial enzymes and chemicals

Organosolv and steam exploded biomass samples used in this study were kindly donated by Prof. Jack Saddler of the University of British Columbia (Vancouver, BC). Mechanically- and kraft- treated wood samples were a gift from FP Innovations (Montreal, QC). Amplex Red glucose/Glucose oxidase assay kit was purchased from Invitrogen (London, ON). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Commercial cellulases, Celluclast 1.5L (Cellulase from *Trichoderma reesei*) and Novozyme-188 (β -glucosidase from *Aspergillus niger*) were purchased from Sigma-Aldrich (St. Louis, MO).

Biomass composition analysis

Eight pretreated woody biomass samples were used in this study: Organosolv pretreated lodgepole pine (OLP); organosolv pretreated poplar, pretreatment done at

160 °C (OPP-1); organosolv pretreated poplar, pretreatment done at 175 °C (OPP-2); acid catalysed steam pretreated poplar, pretreatment done at 200 °C (SPP-1); acid catalysed steam pretreated poplar, pretreatment done at 220 °C (SPP-2), aspen mechanical pulp (AMP); aspen kraft pulp (AKP) and black spruce kraft pulp (BSKP). Structural carbohydrates and lignin in BSKP were determined by FP Innovations, Pointe-Claire, QC, according to the NREL [379] procedure. Composition analysis of organoslov pretreated poplar was done at the Saddler laboratory at the University of British Columbia.

DNA Extraction and cloning of *Mt*cbh7

The sequence of the *Mt*cbh7 target gene was obtained from the *M. thermophila* ATCC42464 genome sequencing portal [413]. Total RNA was extracted using RNeasy Mini Kit (Qiagen) according to the protocol. Messenger RNA was isolated from total RNA by polyA selection using System III (Promega) and manufacturer's instructions. cDNA was prepared using oligo-dT primers and Superscript III (Invitrogen) and was used to generate PCR products. Primers were designed to be compatible the ligation-independent cloning (LIC) method, sequences (underlined) appended to the 5' ends of forward (TACTTCCAATCCAATCCATTTGACGATATGTACGCCAAGTTCGCGAC) and reverse (TTATCCACTTCCAATCCATTTGTTACAGGCACTGCGAGTACCAG) primers [414]. The gene was cloned in a LIC compatible expression vector, ANIp7, which was developed in-house. ANIp7 is maltose-inducible, integrative and carries the orotidine-5'-phosphate decarboxylase gene (pyrG) for selection [335]. The ligation product was transformed in DH5α competent cells, spread on LB plates containing 100 µg/ml ampicillin and incubated overnight at 37 °C. Positive clones were screened by standard PCR. Protoplasts of *A. niger* strain PY11 were transformed purified plasmids as described previously [336]. Transformants were selected on minimal medium without uracil and

uridine [337]. Supernatants from transformants were screened for recombinant protein production after growth in MMJ medium [306].

Expression and purification of *Mt*CBH7

The wild type *M. thermophila* strain ATCC 42464 was first cultured on yeast-starch agar [415] plates at 45 °C for 3 days. Spores were suspended in 0.5% NaCl / 0.02 % Tween 80 solution and stored at 4 °C for later use. For shake flask cultures, 2 % solution of blended BSKP or OPP in 10 times concentrated Trametes defined medium (TDM) was autoclaved and inoculated with the spores to a concentration of 1 million spores /ml under sterile conditions. 200 ml cultures were prepared in 2L shake flasks at 45 °C and 250 rpm under aerobic conditions. After 40 h, the culture supernatant was clarified at room temperature by centrifugation at 3,000 x g for 5 min and stored at -80 °C until further use.

Recombinant *Mt*CBH7 was expressed in *A. niger* strain PY11 host system as described previously [355]. The protein was purified by anion exchange chromatography using a 5 ml HiTrap QXL column pre-packed with Q Sepharose XL anion exchanger from GE Healthcare Life Sciences (Baie d'Urfe, QC) in 25 mM HEPES buffer pH 7.0, as described in reference [355]. The protein was then further purified by size exclusion chromatography using a Superdex 200 column in 25 mM HEPES buffer, pH7.0. Both columns were run on ÄKTA Chromatography System from GE Pharmacia (Baie d'Urfe, QC) as specified by the manufacturer at a flow rate of 2 ml/min.

Protein sequence analysis

The SignalP v4.1 web-based tool [416] was used to predict the signal peptide of recombinant *Mt*CBH7 and the mass and pI of the mature peptide were predicted by the ProtParam tool of ExPASY [417]. Glycosylation was predicted by NetNGlyc v1.0 [356]

and NetOGlyc v4.0 [357] tools, respectively of the Centre for Biological Sequence Analysis.

SDS-PAGE analysis

Proteins in crude supernatants and purification fractions were analysed via 12 % SDS-PAGE stained with Coomassie Brilliant Blue [418]. Gels were scanned on a Syngene G: BOX Chemi system and the image analyzed using Syngene Gene Tools to experimentally estimate the protein mass.

Proteomic analysis

For in-solution digestions, 5 ug of protein was incubated in 100 mM ammonium bicarbonate, 0.25 % AALS II (Morgantown, WV) and 5 mM dithiothreitol for 30 min. Iodoacetamide was added to a final concentration of 25 mM and further incubated for 30 min at 37 °C. Trypsin (150 ng) was added and incubated for 18 h at 37 °C. The digestion solution was acidified with trifluoroacetic acid (1 % final) and desalted using C18 ziptips™ (Millipore, Billerica, MA). Eluted peptides were dried in a SpeedVac and resuspended in 5% CAN:0.1 % formic acid (FA) solution. For protein identity determination on 12% SDS-PAGE gels, bands from lanes to which 2 µg of protein was loaded were cut, destained, reduced, cysteine-alkylated and in-gel digested with sequencing grade modified trypsin (Promega, Madison, WI) as previously described [419]. Peptides were extracted in 1 % FA and acetonitrile (ACN). The extracts were dried in a speedvac and resuspended in 5 % ACN:0.1 % FA solution. The digest (5 ul) was loaded onto a 15 cm x 75 µm i.d PicoFrit column (New Objective, Woburn, MA) packed with Jupiter 5 µm, 300 Å, C18 resin (Phenomemex, Torrance, CA) connected in-line with a Velos LTQ-Orbitrap mass spectrometer (Thermo-Fisher, San Jose, CA) for LC-MS/MS analysis. Peptides were separated on a linear gradient from 3 % ACN:0.1% FA to 90 % ACN:0.1 % FA in an Easy-LC II Nano-HPLC system (Thermo-Fisher). The capillary voltage

on the nanospray source was set at 1.9 to 2.1 kV. MS survey scans spanning the 350 to 2,000 *m*/*z* range were done at 60,000 resolution in the Orbitrap and the top ten multiply charged ions were selected for MS/MS CID fragmentation in the Velos-LTQ trap.

Raw mass spectrometric data were processed using Mascot Distiller ver. 2.4.1 (Matrix Science, London, UK). MS/MS fragmentation peaklist data was searched against the Sporotricum Thermophilum v2 database from the Joint Genome Institute containing 9110 protein sequence entries using Mascot version 2.4.1, followed by X! Tandem CYCLONE ver. 2010.12.01.1 on the subset of identified proteins. Scaffold ver.4.2.1 (Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. In order to ensure a false discovery rate of less than 1% at the peptide level, peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. The relative abundance of the various proteins in a given sample was calculated using the emPAI Calc open-source web application as described elsewhere [420].

Protein concentration estimation

Protein concentrations in the commercial cellulases, Celluclast 1.5L and Novozyme 188, as well as *M. thermophila* culture supernatants, were estimated by first precipitating protein using 2,2,2-trichloroacetic acid (TCA) followed by quantification using the Pierce BCA protein assay kit in microplate format, as described elsewhere [373,421]. Purified recombinant *Mt*CBH7 was directly quantified by BCA protein assay

without TCA precipitation. BSA was used as standard and all reactions were done in triplicate.

Enzyme assay and characterization of *Mt*CBH7

In order to compare activity of commercial cellulases with values reported in the literature [374,361], cellulase activity in Celluclast 1.5L was determined by the microplate version of the dinitrosalicylic acid (DNS) method as described previously [421,374] Glucose (0.02 – 0.2 mg) was used as standard. Cellulase activity was expressed in filter paper units (FPU), with one FPU defined as the amount of cellulase enzyme that released 2.0 mg of glucose equivalent from filter paper in 1 h at 50 °C and pH 4.8 [375]. Cellulase activity in Celluclast was 52 FPU/ml. β -Glucosidase activity in Novozyme 188 against *p*-nitrophenyl-β-D-glucopyranoside was determined as reported previously [376] albeit with slight modification described in reference [355]. *p*-Nitrophenol was used as standard.

The total endoglucanase, xylanase, mannanase and pectinase activities in crude *M. thermophila* supernatant were measured in 96-well microplates using the BCA reducing sugar assay [377] with modifications as described in reference [355]. Recombinant *Mt*CBH7 was similarly screened for activity via the BCA reducing sugar (for polysaccharide substrates) or *p*-nitrophenol [376] release (for *p*-nitrophenol-based substrates) assays as described previously [355]. The 4-methylumbelliferone (4-MU) based fluorimetric assay for estimating glycosidase (cellobiohydrolase) activity was also performed as described previously [422,423,424] with slight modification. 50 mM Britton-Robinson buffer, pH 5.0 (30 µl), was mixed with 10 µl of 1 mM solution of the fluorogenic substrate, 4-methylumbelliferyl cellobioside. Enzyme (10 µl) was added to the mixture and incubated at 40 °C for 30 min. At the end of the reaction, 20 µl of the reaction mixture was withdrawn and dispensed into 180 µl of glycine (1 %)/sodium carbonate (0.88 %) solution, pH 10.7, in a black flat-bottom 96-well fluorescence reading plate. Fluorescence was measured on a plate reader at excitation wavelength of 355

nm and emission wavelength of 460 nm. All substrates and activities tested are listed in Table 11. All enzyme assays were carried out in triplicate. The appropriate monomeric sugar, *p*-nitrophenol or 4-methylumbelliferone was used as standard for BCA reducing sugar, *p*-nitrophenol or 4-methylumbelliferone assays, respectively. One unit of activity was defined as the amount of enzyme that released one µmole of product per minute under standard conditions.

To determine the pH profile of *Mt*CBH7 activity, the standard assay reactions were performed at 40 °C in 50 mM Britton-Robinson buffer (pH 2.0 – 9.0). The effect of temperature on enzyme activity was determined using 50 mM Britton-Robinson buffer (pH 5.0) and incubating reaction mixtures at 30 – 80 °C. Reactions were initiated by adding an appropriate dilution of enzyme and incubated for 30 min. Reactions were terminated and analyzed for reducing sugar, *p*-nitrophenol or 4-methylumbelliferone released, as appropriate.

The ability of *M. thermophila* supernatant and recombinant *Mt*CBH7 to reduce viscosity of a 0.5 % solution of CMC 4M/7M (prepared by mixing CMC 4M and CMC 7M at a ratio of 1:1) was evaluated using an Ostwald viscometer incubated in a water bath at 50 °C. The substrate was diluted in 25 mM sodium acetate buffer, pH 5.5. The substrate solution (10 ml) was loaded in the viscometer and allowed to equilibrate at the incubation temperature. Enzyme (50 µg) diluted in the same buffer was added and the viscosity of the reaction mixture measured by drawing the fluid into the upper bulb via suction and recording the outflow time. Measurements were repeated at different time intervals for up to 15 min.

Biomass hydrolysis

Pretreated woody biomass was freeze-dried and blended in batches of 20 g using a Black and Decker home coffee bean grinder in the cold room (4 °C) to reduce particle size to below the 'mm' range. Biomass samples were then stored at -20 °C until later use.

Biomass hydrolysis was carried out at 15% dry solids in 20 ml scintillation vials in a total reaction volume of 5 ml. All enzyme solutions were in 50 mM sodium acetate buffer, pH 5.5 containing 0.04% sodium azide to prevent sugar loss due to microbial growth. Celluclast 1.5L and Novozym 188 were used as benchmark cellulase and supplementary β-glucosidase, respectively. In all cases, Celluclast 1.5L was supplemented with Novozym 188 as described previously [421] Reactions for the enhancement of glucose release by Celluclast 1.5L were performed by replacing 10 – 50% of total cellulase enzyme with recombinant *Mt*CBH7 or crude *M. thermophila* culture supernatants at a total enzyme loading of 5 mg/g dry matter (DM).

Culture supernatants were also tested for the liquefaction of 15% biomass. In these tests, supernatant from culture grown on a specific biomass type was used in the liquefaction of the same type of biomass at a total protein loading of 5 mg/g DM. In order to identify protein targets that may be responsible for biomass liquefaction, *M. thermophila* supernatants from cultures containing OLP1, OPP1, OPP2, SPP1, SPP2, AMP, AKP and BSKP as carbon source were pooled, buffer exchanged and chromatographed on a HiTrap QXL anion exchange column on ÄKTA purification system. Fractions obtained were tested for the enhancement of glucose release from BSKP by benchmark cellulase (Celluclast supplemented with Novozym 188). The enzymes loadings used to test biomass liquefaction by purified recombinant *Mt*CBH7 alone were 0.5, 5 and 15 mg/g DM. In all cases, enzyme solution was thoroughly mixed with biomass and the reaction vial sealed and incubated in a stationary mode at the appropriate temperature (50 – 60 °C). All reactions were performed in triplicate.

To determine yield of glucose and other sugars after hydrolysis, a small amount of slurry was transferred into 1 ml microfuge tube, centrifuged at 13,000 x g for 5 min, and the supernatant withdrawn for sugar analysis. Glucose was measured by glucose oxidase coupled assay [381,382,383] using the Amplex Red Glucose/Glucose oxidase assay kit from Invitrogen, as specified by the manufacturer. Glucose was used as standard and % glucan conversion was calculated as described in Chapter 4.

Identification and quantitation of the sugars released by *Mt*CBH7 alone was done by high pH anion-exchange chromatography coupled with pulse amperometric detection (Dionex ICS-5000 HPIC system from Thermo Scientific, Bannockburn, IL) [425]. Mono- and disaccharides were analysed on a CarboPac® PA20 column, and oligosaccharides on a CarboPac® PA100, column. The corresponding mono- or oligosaccharides were used as standard. Glucose released from the reaction containing benchmark cellulase supplemented with recombinant *Mt*CBH7 or *M. thermophila* culture supernatant was compared to that from reactions containing benchmark cellulase alone. The rest of the slurry was frozen at -20°C for subsequent rheometric analysis.

Rheometry

The liquefaction potential of *Mt*CBH7 and *M. thermophila* supernatant on 15% BSKP was assessed using the Anton Paar Physica MCR500 rotational rheometer with parallel plates (Anton Paar Canada, St. Laurent, Quebec). The plates were roughened with 100-grit sand paper as described previously [402]. A custom-made tight-fitting aluminium collar (Supplementary Figure 1) was mounted on the bottom plate to prevent biomass leakage from the gap. A 0.2mm clearance allowed free movement of the top plate within the collar. After initial stress amplitude and frequency sweeps for 15% BSKP, stress amplitude of 10 Pa and angular frequency of 100 rad/s were used for measurement of biomass complex viscosity. The gap between the plates was set at 1.8 mm, temperature at 25 °C and normal force between parallel plates set at 0 N. Biomass slurry from hydrolysis reactions at various temperatures (50 – 60 °C) and time points were loaded in the gap and measurements taken during a period of 120s. Each sample was measured three times, 3 min intervals separating each measurement.

Biomass fibre morphology

The effect of enzyme on the surface morphology of the biomass fibres was assessed by scanning electron microscopy (SEM) using JEOL JSM-7400F high resolution field emission SEM (JEOL Canada, Inc. St-Hubert, Quebec). Before and after hydrolysis, fibres were mounted on conducting tape, dried and observed for morphological changes. The voltage was set at 5 kV, current 20 µA and stage gap at 8 mm. At any chosen resolution, a minimum of six images were taken at random so as to evenly cover the surface area of the mounted sample. For each sample, changes in fibre morphology were only considered significant when observed in at least four images.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was funded by Genome Canada, Génome Québec, Agriculture and Agri-Food Canada, the Bioconversion Network of the Natural Sciences and Engineering Research Council of Canada, US DOE. We thank Nadeeza Ismael for cloning the *Mt*cbh7 gene, Annie Bellemare for producing the *A. niger* transformants and David Taylor for the protein production.

Other strategies for the development of efficient lignocellulose degrading proteins were examined in the literauture. For instance, NHCAPs have recently been recognized as a group of proteins which can significantly improve the efficiency of lignocellulose degradation when used to supplement cellulase cocktails. In Manuscript V, a comprehensive literature review of the research progress and potential industrial application of NHCAPs is discussed (see Supplementary Material M1 for a peer reviewed publication on these proteins).

Chapter 6 Manuscript V

Non-hydrolytic cellulose active proteins: Research progress and potential application in biorefineries

Enongene Ekwe[1,3], Ingo Morgenstern[1,2], Adrian Tsang[1,2], Reginald Storms[1,2] and Justin Powlowski[1,3]

[1]Centre for Structural and Functional Genomics, [2]Department of Biology, and [3]Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke Street West, Montreal, QC H4B 1R6

Industrial Biotechnology 2013, 9(3):123 – 131 (reproduced with permission)

6.1 Abstract

A group of proteins including expansins, swollenins and polysaccharide monooxygenases are non-hydrolytic cellulose active proteins (NHCAPs), which have been shown to significantly enhance cellulase-driven plant cell wall hydrolysis at reduced protein loadings. NHCAP-mediated enhancement of cellulase cocktails makes them important accessories in the molecular toolbox for producing fuels and chemicals from lignocellulose. This review presents an overview of the fundamental research advances on these groups of proteins, and considers their applicability to processes in the on-going pursuit of an economically viable biorefinery.

Key words: Non-hydrolytic cellulose active proteins (NHCAPs), expansin, swollenin, polysaccharide monooxygenase, synergy, lignocellulolysis.

6.2 Introduction

The search for sustainable, cleaner fuels and chemicals has intensified efforts to produce them from lignocellulose, a quest that has its beginnings more than five

decades ago [426]. Lignocellulose comprises cellulose (40-50%), hemicellulose (25-35%) and lignin (15-20%); the variation in relative abundance of constituent polymers is dependent on the type of biomass. Within lignocellulose, linear chains of β-1,4-linked D-glucose units are held tightly by intramolecular and intermolecular hydrogen bonding between hydroxyl groups and oxygen atoms present in the macromolecule to form compact, crystalline cellulose microfibrils [12].

Efficient enzymatic hydrolysis of lignocellulose to monomeric building blocks is essential in the commercialization of second generation liquid transportation biofuels and high value chemicals. In nature, fungal saprotrophs [121,241] and phytoparasites [264] secrete complex cocktails of lignocellulose degrading enzymes that function synergistically in plant cell wall deconstruction. These include the carbohydrate active enzymes (CAZymes; http://www.cazy.org) [126,280], and the fungal oxidative lignin enzymes[127]. Several commercial lignocellulolytic enzyme mixtures that include the core cellulases (endoglucanases, cellobiohydrolases, and β-glucosidases), hemicellulases and pectinases have been developed from a few fungal species [360,361,362].

The crystallinity of cellulose and its association with hemicellulose and lignin severely impair its susceptibility to enzymatic hydrolysis (saccharification) [208]. Other key factors contributing to inefficient hydrolysis include non-productive adsorption of enzymes by lignin and other structural components, and the presence of naturally occurring enzyme inhibitors. These factors increase processing costs [206,208]. Recent cost estimates for enzymes range between US$0.30 and US$0.50 per gallon of ethanol, making enzymes the second largest contributor (behind feedstock) to operating cost in the lignocellulose to ethanol bioprocess [427,428].

Research efforts to improve the efficiency of enzymatic saccharification of lignocellulosic biomass are broadly-based and include: development of energy crops that are less resistant to enzymatic hydrolysis; development of a broad spectrum of lignocellulosic biomass pretreatment methods that enhance enzymatic accessibility to

cellulose [203,204]; engineering enzymes with improved catalytic efficiency [223] (site-directed mutagenesis [224], directed evolution [225], construction of chimera or cellulosomes [226] and synergy engineering [209]; search of nature's diversity for more efficient enzymes and accessory proteins [429]; and designing synthetic enzyme cocktails with improved cellulose hydrolyzing capabilities [176,263].

Significant improvements in lignocellulose hydrolysis have been obtained by the inclusion of auxiliary hydrolytic enzymes (such as acetylxylan esterases, arabinofuranosidases, xylosidases, mannanases and pectinases) to core lignocellulolytic enzyme cocktails [176,209,263]. However, the resultant catalytic improvements are limited, possibly due to physical barriers imposed by insoluble or highly ordered polymer packaging in lignocellulosic biomass. High-level synergism may be possible in the presence of auxiliary proteins that can access insoluble biomass and disrupt highly ordered polymer packaging, thereby facilitating attack by hydrolytic enzymes.

A group of non-hydrolytic cellulose active proteins (NHCAPs) including expansins, swollenins and polysaccharide monooxygenases (previously-known as glycoside hydrolase family 61 [GH61] proteins) have been shown to act synergistically with cellulases in lignocellulose hydrolysis [161,430,431,432]. Unlike the core cellulases for which synergistic mechanisms have been extensively studied [433], NHCAPs work concertedly with cell wall hydrolases via mechanisms which are currently poorly understood. Due to their potential ability to enhance lignocellulolysis, NHCAPs are attracting interest as prospective catalytic components in biorefineries. This article presents an overview of the fundamental research advances on NHCAPs and their potential impact on the commercialization of lignocellulose-based fuels and chemicals.

6.3 Classification and structure of non-hydrolytic cellulose active proteins

Classification of expansins

Expansins are thought to occur in almost all terrestrial plants ranging from bryophytes to angiosperms [434]. Two expansins and their encoding genes were originally isolated from cucumber cell walls [435,436] Since then, other plant expansins and expansin-like proteins, from organisms such as bacteria, fungi and nematodes, have been identified and characterized.

Expansins and expansin-like proteins are classified into families using the systematic nomenclature [437,438] shown in Table 13. The four main plant expansin families are designated EXPA (α-expansins), EXPB (β-expansins), EXLA (Expansin-like A) and EXLB (Expansin-like B). In this classification, expansin-like family X (EXLX) is the fifth and newest expansin family and represents the sub-family of expansin comprising non-plant proteins with distant homology to EXPAs and EXPBs [437]. These include the swollenins discussed below.

Table 13. Nomenclature and symbols of expansin superfamily (adapted from Kende *et al*, 2004 [437])

Original Name	Former Symbol	New Symbol	New name
α-Expansin	*EXP* or *EXPα1*	*EXPA*	*EXPANSIN A*
β-Expansin	*EXPB* or *EXPβ1*	*EXPB*	*EXPANSIN B*
Expansin-like	*EXPL* or *EXPβ2*	*EXLA*	*EXPANSIN – LIKE A*
Expansin-related	*EXPR* or *EXPβ3*	*EXLB*	*EXPANSIN – LIKE B*
Various	*EXP* or none	*EXLX*	*EXPANSIN – LIKE X*

Structural properties of expansins

Expansins contain between 250-275 amino acids, with limited sequence conservation between expansins of different families (20-40%) [437,438,439]. A secretory signal peptide directs export to the cell exterior. With the exception of loosenins, described below, expansins comprise two distinct domains: an N-terminal catalytic domain (domain I) and a C-terminal tryptophan-rich putative polysaccharide binding domain (domain II) interconnected by a short linker [440] (reviewed by Kende *et al.* [437]). The crystal structures of two expansins have been solved [441,442,443] Domain I, which is organized into a DPBB (*double psi beta barrel*) fold, is structurally related to the catalytic domain of GH45 proteins (mainly fungal *β*-1,4-D-endoglucanases [http://www.cazy.org/]), whereas the C-terminal domain II is related to group II pollen allergen of grasses [441,442,444]. Domain II of an expansin-like protein (*Bs*EXLX1) lacks any significant sequence similarity with members of the 62 currently known CBM families in the CAZy database, and was proposed as the founding member of CBM family 63 [441].

Classification and structural properties of swollenins

Swollenins are classified as members of the EXLX family. The first swollenin characterized was isolated from the filamentous fungus *Trichoderma reesei* [173]. Like expansins, swollenins have a catalytic domain I and cellulose binding domain interconnected by a linker region [173,174]. However, a single domain EXLX ("loosenin") from the white-rot basidiomycete *Bjerkandera adusta*, *Ba*LOOS1, lacks the carbohydrate binding C-terminal domain II present in plant expansins, with both cell-disrupting and polysaccharide binding activities bundled in domain I [440].

Structure of polysaccharide monooxygenases

Polysaccharide monooxygenases (PMO) constitute a group of small (22 – 45 kDa) proteins previously classified as family 61 glycosyl hydrolases (GH61). It was recently shown that PMOs are Cu-dependent metalloenzymes which oxidatively depolymerize crystalline cellulose in the presence of small molecule redox active agents such as gallate and ascorbate, or the enzyme, cellobiose dehydrogenase [432,445,446,447]. These proteins are structurally similar to chitin-binding proteins, such as CBP21 from *Serratia marcescens* and CelS2 from *Streptomyces coelicolor* [157,158,159], belonging to the carbohydrate-binding module family 33 (CBM33), and which can oxidatively cleave polysaccharides [432,448,449].

Despite the structural and functional similarity of PMOs with CBM33 proteins, there is little evidence that these enzyme families share phylogenetic ancestry. PMOs appear to be restricted to fungi [450], while CBM33 proteins are found mainly in bacteria and viruses [451,167]. Within the fungal kingdom, PMOs are widespread and occur in organisms with different ecological lifestyles and nutritional modes. Most cellulolytic fungal species harbour multiple PMO genes. For instance, the *T. reesei* genome contains six PMO genes [452] whereas that of *Thielavia terrestris* harbours eighteen [453]; the highest number so far (over forty) [454] has been detected in *Chaetomium globosum*. White-rot fungi tend to contain more copies of PMO genes than brown-rot fungi [284].

A recent analysis of 143 PMO genes indicated that 26% of the encoded PMOs are associated with CBM1 (a domain that is involved mainly in cellulose binding [455], and two are associated with the chitin-binding CBM18 (reviewed by Lo Leggio *et al*.) [156] The Cel2S protein from *S. coelicolor* contains both a CBM33 domain and a CBM2 domain, and is active in cellulose cleavage [156,449].

As of February 2013, there were 249 GH61 (PMO) entries in the CAZy database [http://www.cazy.org/] [280]. Of these, five structures have been solved, namely: polysaccharide monooxygenases PMO –2 and PMO–3 from *Neurospora crassa*, GH61A from *Thermoascus aurantiacus,* GH61E from *T. terrestris* and GH61B from *T. reesei*

[156]. None of these five PMO structures is associated with a CBM1 and none shows any evidence of a binding pocket for soluble polysaccharides i.e. large surface clefts, crevices or holes which are commonly found in or proximal to the active sites of conventional glycoside hydrolases. Instead, PMOs (and CBM33) proteins present planar surfaces that appear to be suitable for binding to crystalline polysaccharide surfaces (reviewed by Lo Leggio *et al*.[156]). In the known PMO structures, a divalent metal ion binding site that preferentially binds Cu^{2+} is formed by an essential methylated N-terminal histidine residue, together with a second essential histidine residue and a tyrosine side chain [432,447,156]. For more information, see the review on the structure of PMOs by Lo Leggio *et al* [156].

Classification of polysaccharide monooxygenases

At least two types of PMOs have been functionally distinguished thus far. Type-I PMOs incorporate one oxygen atom from O2 into C-H bonds adjacent to glycosidic linkages generating products oxidized on C1 (reducing end), while Type-II PMOs yield products oxidized at C4 (non-reducing end) [156,162] see Figure 35.

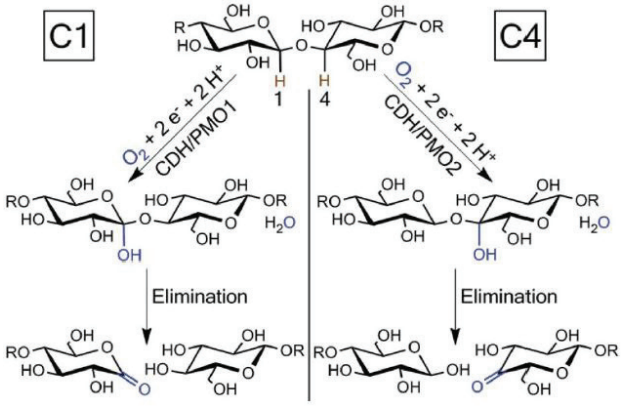


Figure 35. Proposed reaction pathway for CDH-assisted oxidative cleavage of cellulose by GH61 (PMO) – adapted from Beeson *et al.*, 2012 [162]. Note that the oxidation of carbon C1 or C4 is spontaneously followed by a non-reversible elimination reaction, which produce aldonic acid and 4-ketoaldose, respectively

6.4 Activities of non-hydrolytic cellulose active proteins

Activity of expansins

Expansins loosen plant cell wall compactness and disrupt cellulose crystallinity without themselves having cellulolytic [169,170] or xyloglucan endotransglycosylase activities [168]. Several studies have shown that expansins (*Bs*EXLX1 and maize *β*-expansin) bind to cell wall polysaccharides [441,456,457]. Expansins are thought to loosen cell wall matrix polysaccharides by a mechanism involving the disruption of non-covalent bonds between cellulose microfibrils or cellulose and other matrix polymers. This allows cellulose microfibrils to slip or "creep" past each other within cell walls subjected to mechanical stress. For instance, decrystallization of cellulose by expansin involves binding of the protein onto H-bonded cellulose coated with other polysaccharides. As the protein drifts along the chain, disruption of intermolecular and intramolecular H-bonding occurs within the interacting matrix carbohydrates to facilitate slippage [171,172].

A recent report on the structure-function relationship of recombinant *Bs*EXLX1 revealed that although both domains work together, domain I is the main catalytic domain responsible for 'creep activity' (cell wall loosening), while domain II is responsible for binding to cellulose and whole cell wall matrix polysaccharides, indicative of a CBM [441]. Using site-directed mutagenesis, the same study pinpointed various residues involved in polysaccharide binding and wall loosening activities.

Regardless of their limited sequence identity, both EXPA and EXPB type expansins show similar plant cell wall loosening effects, are activated by reducing agents, and are active under mildly acidic (below pH 6) conditions but not at neutral pH [435,458]. Evidence of the non-hydrolytic nature of expansins is that their cell wall loosening action is not time-dependent and their effects on the mechanical properties of the cell wall are fully reversible upon protein denaturation [168]. Some studies have shown that expansins are co-expressed with endoglucanase and xyloglucan endotransglycosylase in response to auxin, suggesting that these proteins may act

synergistically in cell expansion [459,460,461]. Differences in substrate specificity have been observed between EXPAs and EXPBs, with EXPA showing preference for dicot cell walls whereas the EXPB prefer those of grasses: grass cell walls are composed of unusual hemicelluloses (e.g. mixed-link glucan and glucuronoarabinoxylan) in addition to the usual matrix polymers present in dicot cell walls (i.e. xyloglucan, pectin, etc.) [462]. Although the cell wall loosening activity of many members of the EXPA and EXPB families has been demonstrated, there is a dearth of experimental evidence for this activity in EXLA and EXLB family proteins.

Activity of swollenins

Swollenins are expansin-like proteins that cause loosening, partial disruption and swelling of plant cell walls [173]. As is the case for expansin, the catalytic domain of swollenin harbours cell-wall polysaccharide loosening activity that partially disrupts, loosens and reduces cellulose intensity (decrystallization) without release of reducing sugars [173,174]. The original work on *T. reesei* swollenin characterization is complemented by binding studies of swollenins from other fungal species. *Aspergillus fumigatus* swollenin, *Af*Swo1, binds to microcrystalline cellulose (avicel) and chitin, shows negligible activity toward carboxymethyl cellulose (CMC) and avicel, and does not bind to xylan [430]. Swollenin 2 (Swo2) from *Trichoderma pseudokoningii* S38 shows slight hydrolytic activity on xylan and yeast cell wall glycan, but not on CMC, cotton fibre, filter paper or cellulose powder CF11 [463].

Activity of polysaccharide monooxygenases

In the presence of molecular oxygen and reducing agents, PMOs are involved in the oxidative disruption of crystalline cellulose [160] and chitin [432], tying their function to the structural homology they share with CBM33 [156,157,158,159]. Their mechanism of action and the nature of carbohydrate oligomers released (i.e. oxidized cello- or chito- oligomers) are markedly different from those of canonical hydrolytic

cellulases (i.e. non-oxidized oligomers) [160,161]. Although C1, C4 or, to a lesser degree, C6 [164] in the glucose ring may be oxidized by PMOs, aldonic cellodextrins resulting from C1 oxidation are the most abundant products [162,163]. This oxidative cleavage of cellulose into a variety of native and oxidized cellodextrins with varying degrees of polymerization (DP) [160] renders lignocellulosic biomass more susceptible to attack by classical endo- and exo-cellulases [432,447,160].

Synergy between PMO and the secreted fungal flavocytochrome, cellobiose dehydrogenase (CDH), appears to promote cellulose degradation by canonical hydrolases [160,165,166,167] Among other suggested roles [464,465,466], CDHs are thought to bind cellulose, oxidize cellodextrins, maltodextrins and lactose to their corresponding lactones [464], and reduce Cu(II) to Cu(I) in the PMO catalytic cycle [156,445]. Most fungal genomes contain PMO and CDH encoding genes [453,467] which are expressed and up-regulated together [160]; a few others (e.g. the brown-rot fungus, *Postia placenta)* have multiple PMO encoding genes but none encoding CDH [468].

6.5 Cellulose hydrolysis with NHCAP additives

Expansins as additives to lignocellulolytic enzyme cocktails

The expansins most commonly used in research as additives to enzyme mixtures for lignocellulose hydrolysis are the EXLX family proteins. Unlike the plant expansin families, which have not been successfully expressed in non-plant host systems [430] most EXLX members are readily expressed heterologously in a diverse number of host systems including bacteria, fungi and plants [430,173,442]. Consequently, the EXLX proteins are promising candidates for improvement of industrial-scale lignocellulose hydrolysis.

Several studies of cellulose cocktail enhancement have used the *β*-expansin-like protein from *Bacillus subtilis (Bs*EXLX1). This protein showed a degree of synergism of 5.7 with low cellulase loadings (<0.6FPU/g cellulose) on filter paper, but no improvement was observed at higher cellulase loading [430]. Glucose yields were,

however, too low for *Bs*EXLX1 to be considered applicable in industrial lignocellulolysis. In a second study, *Bs*EXLX1 and crude cellulase mixtures from *Aspergillus* and *Penicillium* species increased glucose, xylose, arabinose and total reducing sugar yields from alkali pretreated wheat straw by 45% - 112%, compared to the cellulase preparation Celluclast 1.5L alone [469]. By replacing 32% of cellulase with recombinant *Bs*EXLX1, the total cellulase loading was reduced by 21%, with a net gain in glucose, xylose, arabinose and total reducing sugar yield of between 10% and 13% each. Another expansin-like protein from *Hahella chejuensis* (*Hc*EXLX2) was reported to confer a 4.6-fold enhancement in reducing sugar released from filter paper by Celluclast 1.5L [470], but no tests were performed on natural lignocellulose.

Cellulase loading often used for assessing digestibility of pretreated lignocellulosic biomass is 5FPU/g glucan or higher [206,471], which exceeds loadings at which enhancement is observed for *Bs*EXLX1 and *Hc*EXLX2 [442,430,470]. Preferential binding of recombinant *Bs*EXLX1 to lignin, rather than cellulose, has led to the suggestion that bacterial expansins could be used as lignin blocker in the enzymatic hydrolysis of lignocellulose. Other substrate binding studies on *Bs*EXLX1 conclude that the success of expansin dosing in biomass conversion is dependent on the substrate pretreatment method [472].

The first report using a plant expansin to reduce doses of *T. reesei* cellulase in the hydrolysis of microcrystalline cellulose was in 1998 [430]. A small amount (0.1mg of enzyme/g cellulose) of *β*-expansin-D from corn pollen was reported to enhance hydrolysis of acid-pretreated yellow poplar sawdust by a *T. reesei* cellulase mixture [473]. The improved rate of saccharification was observed in the glucan conversion range of 55–80%. When tested on a synthetic binary cellulase mixture containing *Acidothermus cellulolyticus* endoglucanase and *T. reesei* CBHI, (8.35 mg total protein /g biomass), *β*-expansin-D did not enhance saccharification, which suggests that for the cell wall loosening corn *β*-expansin-D has to interact with one or more component(s) in the *T. reesei* cellulase complex mixture missing in the synthetic binary mixture [473].

Wei *et al.* studied synergism between an α-expansin from cucumber and a mix of purified fungal cellulases, hemicellulases and pectinases [474]. Cucumber α-expansin shows concentration-dependent enhancement of hydrolysis of homogeneous model substrates (pectin and carboxymethylcellulose) [474], but it is difficult to extrapolate these results to complex lignocellulosic biomass. Moreover, all the previous studies employing expansins as additives in lignocellulolytic enzyme cocktails used biomass concentrations lower than 2%, in contrast to the high loadings required in industrial bioprocesses [243].

Swollenins as additives to lignocellulolytic enzyme cocktails

Swollenins have also been reported to synergistically enhance cellulase activity. One study reported increased release of reducing sugar from endoglucanase hydrolysis of cotton fibre pretreated with *Ba*LOOS1, the increase of which was proportional to the amount of *Ba*LOOS1 used [440]. When tested with a cocktail of commercial cellulases and xylanases on lignocellulosic substrate (*Agave tequilana* bagasse) *Ba*LOOS1 rendered the biomass 7.5 times more susceptible to enzymatic hydrolysis. Nonetheless, *Ba*LOOS1 is inactive at temperatures above 40˚C [440], limiting its use in lignocellulolytic bioprocess designed for higher temperatures.

The effects of a number of swollenins on hydrolysis of model cellulose substrates have been reported. Pretreatment of filter paper and avicel by *T. reesei* swollenin, produced in *Kluyveromyces lactis*, was found to enhance cellulase hydrolysis of both substrates [475]. However, *T. reesei* swollenin is produced only at a low level by *T. reesei* (≤25 μg/l), limiting its usefulness for industrial applications [173,174]. *Trichoderma asperellum* swollenin enhances hydrolysis of avicel by endoglucanase from *Fervidobacterium nodosum* [476]. Recombinant *Aspergillus oryzae* swollenin promotes cellulase hydrolysis of filter paper by up to 80% in a concentration-dependent fashion [174]. A concentration-dependent increase (61-74%) in cellulase-mediated

saccharification of crystalline cellulose was observed when 20µg AfSwo1/g cellulose was added to a commercial cellulase mixture from *T. reesei* and *Aspergillus niger* [430].

Some swollenins appear to exhibit low but significant hydrolytic activity toward polysaccharides, making it difficult to determine whether it is the wall loosening or the hydrolytic activity that interact synergistically with cellulases in lignocellulosic biomass hydrolysis [431].

Polysaccharide monooxygenases in lignocellulose deconstruction

The current generation of commercial cellulase cocktails are known to contain PMOs. Cellic CTec2, manufactured by Novozymes A/S, contains extra PMOs that have significantly contributed to the improved performance compared to its predecessors, notably Celluclast 1.5L [163]. Coexpression of *T. terrestris* or *T. aurantiacus* PMOs with cellulases in a commercial hypercellulolytic *T. reesei* strain leads to a 2-fold reduction in the total protein loading and cost of hydrolysis of acid-catalyzed steam-exploded corn stover [432]. Using the *T. aurantiacus* PMO as model protein, Harris and co-workers concluded that stimulatory effects of PMOs on cellulases are not due to interaction with any single specific enzymatic activity, but rather with the overall activity of the cellulase complex or substrate [432].

The cellulase enhancement activity of PMOs seems to be influenced by the presence of other cell wall-derived material such as lignin or hemicellulose present in the substrate [432]. For instance, lignin boosted the cellulase-enhancing performance of a recombinant *Myceliophthora thermophila* PMO in the degradation of hydrothermally pretreated wheat straw, soda-pretreated spruce/pine pulps and acid-catalyzed steam pretreated spruce. Up to 40% enhancement was achieved on the cellulase hydrolysis of pretreated spruce [477]. A mixture containing 10.5mg of *M. thermophila* PMO and 16.5mg mixture of purified cellulases (a combination of endoglucanase, cellobiohydroase and β-glucosidase) was found to have the same hydrolyzing activity on pretreated lignocellulosic biomass as 33 mg of purified cellulolytic enzymes, the

equivalence of 20% reduction in enzyme loading [477]. Canella *et al.*, used pretreated wheat straw as representative lignocellulosic biomass and filter paper as pure cellulosic substrate to demonstrate that *T. reesei* PMO additives in Cellic CTec2 does not require any reducing co-factor for activity in the presence of lignin [163], implying that lignin can function as electron donor for PMO activity. This also suggests that PMOs may exploit adventitious copper and reductants present in natural lignocellulosic biomass for their activity, making such metal-induced lignocellulolytic enhancement feasible at low cost. The role of PMOs in the degradation of cellulose by fungal enzymes is represented in Figure 6 [162].

One challenge associated with the use of PMO additives in cellulase mixtures for the deconstruction of lignocellulose is the production of non-fermentable gluconic and cellobionic acids (see Figure 36). Cannella *et al.* found a positive correlation between the gluconic acid yield and cellulose conversion when Cellic CTec2 was used to hydrolyze pretreated wheat straw at high (30%) solids loading. Unlike CTec2 which contains PMO additives, Celluclast 1.5L (supplemented with Novozyme 188) does not produce any detectable gluconic acid from the same substrate [163]. Gluconic acid was found to be a more potent *β*-glucosidase inhibitor than glucose, while Cellic CTec2 was found to hydrolyse cellobionic acid at a much slower rate than cellobiose. In the same study, gluconic acid production was found to be lower at higher temperatures without compromising the overall cellulose conversion [163] This suggests that temperature could be used, to a certain extent, to control aldonic acid levels and its inhibitory effect on *β*-glucosidase.

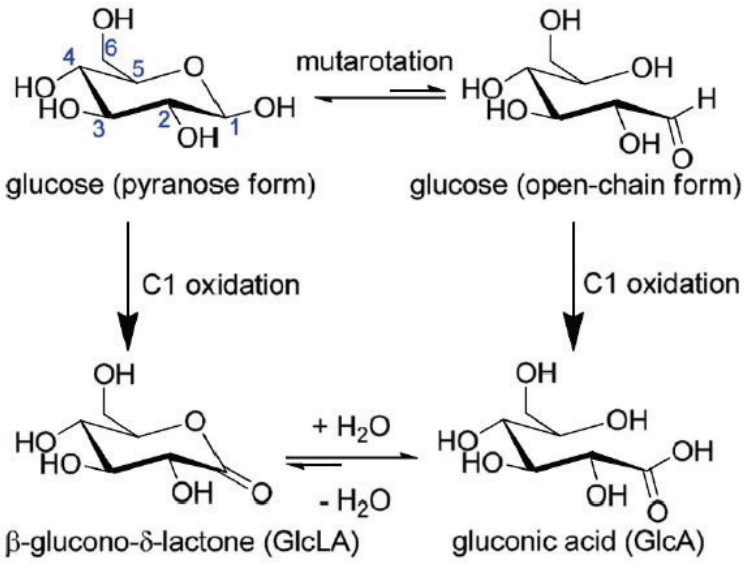


Figure 36. Gluconic acid formed from C1 oxidation of glucose – adapted from Cannella *et al*., 2012 [163]

6.6 Other non-hydrolytic cell wall active proteins

Other non-hydrolytic cell-wall active proteins have been reported to cause expansion or degradation of plant cell wall polysaccharides. Two non-hydrolytic proteins that regulate cell wall expansion in cowpea hypocotyl were reported in 1995 [478]. However, details of the description and characterization of these proteins are not available. Another non-hydrolytic protein termed 'acid wall protein' was found to stimulate activity of exo- and endoglucanases in maize coleoptile walls [479]. It is not clear, however, whether the action of the 'acidic wall protein' renders cell wall polymers more accessible to hydrolysis by exo- and endoglucanases. A cellulose-induced protein Cip1, was identified in *T. reesei* [480]. The fact that the sequence of Cip1 not fit into any of the known GH families, that it is co-expressed with canonical cellulases [480], and

that it contains a cellulose-binding module, suggests that it may play a non-hydrolytic role in biomass degradation.

6.7 Concluding Remarks

This review highlights the growing research interest on NHCAPs, particularly, toward their potential application in commercial scale enzymatic biomass conversion. Fundamental research progress on NHCAPs has been significant in recent years, but much remains to be learned. For instance, due to the large superfamily size of expansins, the expression, regulation and functional roles of many expansin genes have yet to be explored, which might unlock vital information toward the use of these proteins in biomass degradation. While the mechanistic reactivity of a number of PMOs has been clearly established as copper- and reductant-dependent polysaccharide (cellulose and chitin) monooxygenases, sequence diversity in the family is large, and it remains to be seen how similar the as-yet uncharacterized PMOs are. Furthermore, many of their catalytic properties (e.g. pH and temperature optima; specific activity) have not yet been extensively studied. Finally, it is still unclear on which specific polymers expansins and swollenins act to loosen the cell wall and what molecular mechanisms are involved.

Very few studies have focused on the direct use of NHCAPs as booster additives in enzyme cocktails for biomass degradation. With the exception of studies leading to the development of the Cellic CTec enzyme family, the small number of studies that have tested NHCAPs for lignocellulosic biomass hydrolysis have used relatively homogeneous model substrates (such as avicel, phosphoric acid swollen cellulose and filter paper) which do not match the complexity of natural lignocellulosic biomass. In addition, these studies usually employed solids loadings of 5% or less (which is far below loadings in excess of 15% required industrially) [243] or have used thermolabile NHCAPs (e.g. *Ba*LOOS1 tested by Quiroz-Castañeda *et al.*[440]). To overcome the challenges related to enzyme inefficiency in lignocellulose degradation, research on NHCAPs could be focused on thermostable NHCAPs with tolerance to high solids loading. High

temperatures facilitate biomass solubilization while high solids loadings ensure high product yields at low operation cost. Thermophilic fungal saprotrophs represent an interesting, rich gene reservoir from which such proteins could be sourced. Several research groups are now focusing on such thermophilic fungal species for the development of improved lignocellulolytic enzyme cocktails.

Acknowledgements

Research on lignocellulose-active proteins at Concordia University is supported by funds from Genome Canada, Génome Québec and the Bioconversion Network of the Natural Sciences and Engineering Research Council of Canada.

Chapter 7 Conclusion and outlook

Resullts obtained in this project indicate that *N. benthamiana* is as effective a production host for *A. niger* glycoside hydrolases as *A. niger* itself. The recombinant glycoside hydrolases produced in the four host systems showed similar properties with the exception of xylanases produced in *E. coli* and *P. pastoris* as well as polygalaturonases in *P. pastoris*. These proteins displayed biochemical properties that are substantially different from those produced in the parent host. Since both the native and recombinant xylanases are not glycosylated, it is unlikely that the differences in biochemical properties are due to protein glycosylation. The success rates of producing the heterologous GHs in *P. pastoris* and *A. niger* were similar, whereas the rate in *E. coli* was poor. Developing *N. benthamiana* into a more efficient expression system will provide an interesting opportunity for commercial scale *in planta* production of cell wall

degrading proteins of fungal origin, which can be used to supplement cellulase mixtures or to design synthetic biomass degrading enzyme mixtures.

Direct screening methods for plant cell wall degrading enzymes (such as those expressed in these four host systems) on natural lignocellulosic substrates are desired to facilitate extrapolation of the protein catalytic efficiency to industrial bioconversion processes. A medium- to high-throughput adaptable screening method was successfully developed in this project and it appeared suitable for screening fungal cell wall degrading enzymes at high substrate loading. Through this method, a GH7 cellobiohydrolase from *A. niger*, a GH10 xylanase from *T. terrestris* and a putative acetylxylan esterase from *A. niger* were identified as promising additives for the enhancement of glucose release from 15% AMP by cellulases. This method was also found suitable for evaluating the hydrolysability of pretreated biomass as a function of biomass type and pretreatment efficiency. The method was found to be reproducible, required minimal amount of enzyme, showed a dose-dependent response in sugar yields at Celluclast loadings up to 32 mg/g and maintained residual moisture content of ≥ 90% after 72h. This method presents a tool for screening the large number of biochemically characterized proteins reported in the literature on natural lignocellulosic biomass to evaluate their catalytic potential in industrial bioconversion processes. This method is however not suitable for pure, non-fibrous cellulosic substrates like avicel, PASC and solka flock due to clogging of the microfilter membranes and uneven distribution of enzyme in the substrate.

The alternative approach employed in this study to develop efficient lignocellulytic enzymes from thermophilic filamentous fungi, involved growing *M. thermophila* on various pretreated wood biomass types as sole carbon source and analyzing the secreted proteins. Growth of the strain was better in organosolv- and kraft- treated biomass substrates with high cellulose, low lignin and low hemicellulose contents. Unlike the cellulase activities of the secreted protein cocktails which were not affected by the glucan content of the carbon source, the xylanase activities correlated

strongly with xylan content. This indicates that high xylan content in biomass induced high xylanase activity. The secreted supernatants were capable of rapid reduction in the viscosity of a variety of 15% pretreated wood biomass. In terms of glucose released, *M. thermophila* secreted protein cocktails were able to release up to 75% the amount of glucose released by the baseline commercial cellulase, Celluclast 1.5L. It was also able to boost the hydrolytic capability of Celluclast 1.5L on various pretreated wood biomass types.

The results further provide insight on the group of proteins in the secretome of *M. thermophila* that concertedly liquefy high consistency biomass. Forty-seven secreted proteins were detected in the secretome of the strain. Among these was a highly-expressed GH7 cellobiohydrolase, *Mt*CBH7, which accounted for >25% of the total secreted protein. Analysis of partially fractionated *M. thermophila* secreted protein cocktail revealed that the fraction containing *Mt*CBH7 confered the highest degree of enhancement on Celluclast in the release of glucose from 15% BSKP.

Purified recombinant *Mt*CBH7A showed very little activity on the major plant cell wall polysaccharides, but was capable of reducing the viscosity of 15% BSKP, acting optimally at 55 °C. At low protein loading, *Mt*CBH7 boosted the Celluclast liquefaction of pretreated wood biomass at high solids loading and increased Celluclast glucose yield from the same substrate. Although other proteins in the *M. thermophila* secretome appear to be involved in the liquefaction of pretreated biomass, *Mt*CBH7 plays a significant role. *Mt*CBH7 therefore appears to be a promising enzyme candidate for application in biorefineries where enhanced liquefaction of lignocellulose is required.

Optimization of *M. thermophila* growth on various biomass types could be an interesting strategy to increase the production of lignocellulolytic proteins from this strain. For instance, testing various mixtures of pretreated biomass types as sole carbon source would identify those that better support growth and protein secretion by the strain. Fine-tuning the temperature, pH and nutrient requirements could also improve protein yield. It is also necessary to test *Mt*CBH7 on other pre-treated lignocellulosic

biomass types in order to compare the liquefaction effects to that observed on BSKP. It will be interesting to see the reproducibility of the hydrolytic and liquefaction effects of *Mt*CBH7 as well as *M. thermophila* secreted protein cocktails from various biomass types at larger scale. Bioreactor systems of 100 – 500 mL capacity will provide valuable data with regards to the whether the liquefaction effects of *Mt*CBH7 and *M. thermophila* supernatants can be extrapolated to larger scale. This will also provide enough hydrolysate for assessing the fermentability of the sugars released.

A similar approach can be used to evaluate other thermophilic lignocellulose-degrading fungal species such as *T. terrestris, Myceliophthora heterothallica, Myceliophthora fergusii, Thermoascus aurantiacus,* etc. for the identification of novel thermotolerant biomass liquefying enzymes. Also, studies that are focused on developing *M. thermophila* into a hypersecretory strain for the production of plant cell wall degrading protein cocktails will improve on protein yield for direct application in commercial biomass conversion processes.

The supplementation of cellulase mixtures with NHCAPs is yet uncommon. Although experimental work was not done on these proteins in this work, the potential impact of NHCAPs on biomass conversion is huge, as in the case of PMOs on the Cellic CTec cellulase enzyme family. Studies evaluating and improving on the effects of NHCAPs on the deconstruction of complex natural lignocellulose are necessary. For instance, attention should be paid on the cloning, expression, characterization and use of NHCAPs as supplements of cellulase mixtures. Focus on thermostable NHCAPs which can boost Cellulase hydrolysis of lignocellulose at high solids loading will be suitable for bioprocesses at temperatures >50 ºC for improved biomass solubilization and product yields and low operation cost. Thermophilic fungal saprotrophs represent an interesting, rich gene reservoir from which such proteins could be sourced for the development of improved lignocellulolytic enzyme cocktails.

Appendices

Supplementary Table 1. Summary of reported properties of genes and enzymes used in this study

Gene name	Enzyme Activity	JGI protein ID	Expression host	Specific activity	pH opt	Temp opt (°C)
abf54B	arabinofuranosidase	200605	native & *A. niger*		3.8	56
abn43A	arabinanase	203143	native & *A. niger*		4.6	51
agl31A	alpha-glucosidase	214233	*A. nidulans*			
agu67A	alpha-glucuronidase	56619	native		3.5	40
axh62A	arabinoxylan-arabinofuranhydrolase	55136	*A. niger*			
bgl3A	beta-glucosidase	56782	*S. cerevisiae* & *P. pastoris*			
cbh7A	cellobiohydrolase	53159	*A. niger*			
cbh7B	cellobiohydrolase	51773	*A. niger*			
egl12A	endoglucanase	211053	*E. coli*			
egl5B	endoglucanase	209376	*E. coli*			
inu32A	inulinase	52928	*E. coli*			
lac35A	beta-galactosidase	51764	*S. cerevisiae*			
mel27A	alpha-galactosidase	37736	*A. niger*			
mel27B	alpha-galactosidase	207264	*A. niger*			
mnd2A	beta-mannosidase	138876	native	30U	4	70
pga28A	endo-polygalacturonase	214598	*A. niger*	990U	4.2	
pga28E	endo-polygalacturonase	46255	*A. niger*	30U	3.8	60
pga28II	endo-polygalacturonase	182156	*A. niger*	2747U	4.2	
pgx28C	exo-polygalacturonase	172944	*A. niger*	202U	3.7	
rhg28B	rhamnogalacturonase	211163	*A. niger*	0.4U	4.1	
suc32A	beta-fructofuranosidase	198063	native	3290U	5.5	50
xeg12A	xyloglucanase	52011	*A. niger*	83U	5	50
xyl3D	xylosidase	205670	*A. nidulans*	60U		
xyn10A	*xylanase	57436	native		5.5	60
xyn11B	xylanase	52071	native	2265U	5	55

Data source: https://mycoCLAP.fungalgenomics.ca. *The biochemical property reported here comes from the xylanase A of *Aspergillus kawachii* (Genbank accession no. BAA03575) which displays >97 identity in amino acid sequence with xyn10A of *A. niger*.

Supplementary Table 2. *Sequences* of oligonucleotide primers used to amplify genes for expression in *Escherichia coli*

Gene name	Protein ID	Forward primer*	Reverse primer*
abf54B	200605	5′-tacttccaatccaatgcc GCTGACCCCGGAGCATGCTC	5′-ttatccacttccaatgtta TACAACCGGCCATCCGCTAGAGAAG
abn43A	203143	5′-tacttccaatccaatgcc GCTGACCCCGGAGCATGCTC	5′-ttatccacttccaatgtta TACAACCGGCCATCCGCTAGAGAAG
agl31A	214233	5′-tacttccaatccaatgcc AGCCAGTCACTCTTATCCACCACTG	5′-ttatccacttccaatgtta CCATTCCAATACCCAGTTTTCCGCC
agu67A	56619	5′-tacttccaatccaatgcc GAGGATGGGTACGATGGCTG	5′-ttatccacttccaatgtta ATCACCCACAGATAGTGTCCAC
axh62A	55136	5′-tacttccaatccaatgcc AAATGCGCTCTTCCGTCGACATATAG	5′-ttatccacttccaatgtta CTGCTTCAAGGTAAGAACTCCTGG
cbh7A	53159	5′-tacttccaatccaatgcc CAGCAAGCCGGAACGCTCAC	5′-ttatccacttccaatgtta TGCGGAAGCGCTGAAGGTCG
cbh7B	51773	5′-tacttccaatccaatgcc CAGCAGGTTGGCACCTACAC	5′-ttatccacttccaatgtta CAAACACTGCGAGTAGTACGC
egl12A	211053	5′-tacttccaatccaatgcc CAGACGATGTGCTCTCAATATGAC	5′-ttatccacttccaatgtta GTTGACACTGGCGGTCCAG
egl5B	209376	5′-tacttccaatccaatgcc GTGCCTCATGGCTCCGGAC	5′-ttatccacttccaatgtta GAGATACGTCTCCAGGATATCCAG
inu32A	52928	5′-tacttccaatccaatgcc AATGATTACCGTCCTTCATACCACTTC	5′-ttatccacttccaatgtta TTCAAGTGAAACACTCCGCACGTC
mel27A	37736	5′-tacttccaatccaatgcc TCTATCGAGCAGCCCAGCCTC	5′-ttatccacttccaatgtta GGCCAACTGAACGCCACTTGG
mel27B	207264	5′-tacttccaatccaatgcc GTTCGACCCGATGGCGTGGG	5′-ttatccacttccaatgtta CGAATTCAACATCAACACAAACACCT CCC
pga28A	214598	5′-tacttccaatccaatgcc GCTCCTGCGCCTTCCCG	5′-ttatccacttccaatgtta CTGATCGCAAGAAGCACCAGAG
pga28II	182156	5′-tacttccaatccaatgcc TCTCCTATCGAAGCTCGAGACAGC	5′-ttatccacttccaatgtta ACAAGAGGCCACCGAAGGGAAG
pgx28C	172944	5′-tacttccaatccaatgcc GTTCCACACTCCAGCAGAGC	5′-ttatccacttccaatgtta ATTAGATAATGGGCTGTAGGTCCC
suc32A	198063	5′-tacttccaatccaatgcc GCCTCGCCTTCCATGCAG	5′-ttatccacttccaatgtta CCGAACCCAAGTACTCAACG
xeg12A	52011	5′-tacttccaatccaatgcc GCCTCTATCAGCCGCCGTAG	5′-ttatccacttccaatgtta CTCGATGGAAACGGAGTACTCG
xyn10A	57436	5′-tacttccaatccaatgcc CCCATTGAACCCCGTCAGGC	5′-ttatccacttccaatgtta GAGAGCATTTGCGATAGCAGTGTATG
xyn11B	52071	5′-tacttccaatccaatgcc GTTCCCCACGACTCTGTGG	5′-ttatccacttccaatgtta CTGAACAGTGATGGACGAAGATC

*The sequences corresponding to the ligation-independent annealing sites are shown in lowercase whereas the gene-specific sequences are presented in uppercase.

Supplementary Table 3. Sequences of oligonucleotide primers used to amplify genes for expression in *Pichia pastoris*

Gene name	Protein ID	Forward primer*	Reverse primer*
abf54B	200605	5'-ccgctcgagaaaaga GGCCCCTGTGACATCTACGAAGCC	5'-ctagctagcac CGAAGCAAACGCCGTCTCAATCTCAA AGC
abn43A	203143	5'-ccgctcgagaaaaga GCTGACCCCGGAGCATGCTC	5'-ctagtctagagc TACAACCGGCCATCCGCTAGAGAAG
agl31A	214233	5'-ccgctcgagaaaaga AGCCAGTCACTCTTATCCACCACTG	5'-ctagtctagagc CCATTCCAATACCCAGTTTTCCGCC
agu67A	56619	5'-ccgctcgagaaaaga GAGGATGGGTACGATGGCTGG	5'-ctagtctagagc ATCACCCACAGATAGTGTCCACTTC
cbh7A	53159	5'-ccgctcgagaaaaga GTGCCTCATGGCTCCGGAC	5'-ctagtctagagc GAGATACGTCTCCAGGATATCCAG
egl12A	211053	5'-ccgctcgagaaaaga CAGACGATGTGCTCTCAATATGAC	5'-ctagtctagagc GTTGACACTGGCGGTCCAG
egl5B	209376	5'-ccgctcgagaaaaga GTGCCTCATGGCTCCGGAC	5'-ctagtctagagc GAGATACGTCTCCAGGATATCCAG
inu32A	52928	5'-ctagtcgacaaaaga AATGATTACCGTCCTTCATACCACT TCAC	5'-ctagtctagagc TTCAAGTGAAACACTCCGCACGTCC
mel27B	207264	5'-ccgctcgagaaaaga GTTCGACCCGATGGCGTGG	5'-ctagtctagagc TCCCACCACCAAAACCGCGAC
mnd2A	138876	5'-ctagtcgacaaaaga ATGATCCTGCAACCGGTGTCTAATG ATC	5'-ctagtctagagc CTCGCGCATCGACTTCTTCCTCG
pga28A	214598	5'-ctagtcgacaaaaga GCTCCTGCGCCTTCCCG	5'-ctagtctaga GCCTGATCGCAAGAAGCACCAGAG
pga28II	182156	5'-ctagtcgacaaaaga TCTCCTATCGAAGCTCGAGACAGC	5'-ctagtctagagc ACAAGAGGCCACCGAAGGGAAG
pgx28C	172944	5'-ccgctcgagaaaaga GTTCCACACTCCAGCAGAGC	5'-ctagtctagagc ATTAGATAATGGGCTGTAGGTCCC
rhg28B	211163	5'-ctagtcgacaaaaga CAGCTCTCCGGTTCCGTGGG	5'-ctagtctagagc CACGTAGCAGACACCCTGCTCG
suc32A	198063	5'-ccgctcgagaaaaga GCCTCGCCTTCCATGCAGACG	5'-ctagtctagagc CCGAACCCAAGTACTCAACGCAAAG
xeg12A	52011	5'-ccgctcgagaaaaga GCCTCTATCAGCCGCCGTAGC	5'-ctagtctagagc CTCGATGGAAACGGAGTACTCGGAC
xyl3D	205670	5'-ccgctcgagaaaaga CAGGCCAACACCAGCTACGTCG	5'-ctagtctagagc CTCCTTCCCCGGCCACTTCAG
xyn10A	57436	5'-ccgctcgagaaaaga CCCATTGAACCCCGTCAGGC	5'-ctagtctagagc GAGAGCATTTGCGATAGCAGTGTATG
xyn11B	52071	5'-ccgctcgagaaaaga GTTCCCCACGACTCTGTGG	5'-ctagtctagagc CTGAACAGTGATGGACGAAGATC

*The sequences corresponding to the ligation-independent annealing sites are shown in lowercase whereas the gene-specific sequences are presented in uppercase.

Supplementary Table 4. Sequences of oligonucleotide primers used to amplify genes for expression in *Aspergillus niger*

Gene name	Protein ID	Forward primer*	Reverse primer*
abf54B	200605	5'-ccaactttgtacaaaaaagcaggct ATGTTCTCCCGCCGAAACC	5'-ccaactttgtacaagaaagctgggt TTACGAAGCAAACGCCGTCTC
abn43A	203143	5'-ggggacaagtttgtacaaaaaagcaggct ATGTATCAACTCCTATCAGTTGC	5'-ggggaccactttgtacaagaaagctgggt TTATACAACCGGCCATCC
axh62A	55136	5'-ccaactttgtacaaaaaagcaggct ATGAAATTCCTCAAAGCCAAGGGTAG	5'-ccaactttgtacaagaaagctgggt TCACTGCTTCAAGGTAAGAACTCC
cbh7A	53159	5'-ggggacaagtttgtacaaaaaagcaggct ATGCATCAACGTGCCCTTCTC	5'-ggggaccactttgtacaagaaagctgggt TTATGCGGAAGCGCTGAAGG
egl12A	211053	5'-ccaactttgtacaaaaaagcaggct ATGAAGCTCCCCGTGTCACTTG	5'-ccaactttgtacaagaaagctgggt CTAGTTGACACTGGCGGTCCAG
egl5B	209376	5'-ccaactttgtacaaaaaagcaggct ATGAAGTTTCAGAGCACTTTGCTTCTTG	5'-ccaactttgtacaagaaagctgggt TCAGAGATACGTCTCCAGGATATCCAG
inu32A	52928	5'-ccaactttgtacaaaaaagcaggct ATGTTGAATCCGAAGGTTGCCTAC	5'-ccaactttgtacaagaaagctgggt TCATTCAAGTGAAACACTCCGCAC
mel27B	207264	5'-ggggacaagtttgtacaaaaaagcaggct ATGCGGTGGCTTCTCACCTC	5'-ggggaccactttgtacaagaaagctgggt CTAACATTGCCCTCCCACCAC
mnd2A	138876	5'-ccaactttgtacaaaaaagcaggct ATGCGCCACAGCATCGG	5'-ccaactttgtacaagaaagctgggt TCATCCGCGAACCTTCTGG
pga28A	214598	5'-ccaactttgtacaaaaaagcaggct ATGCCTTCTGCCAAGCCTTTG	5'-ccaactttgtacaagaaagctgggt TTACTGATCGCAAGAAGCACCAG
pga28II	182156	5'-ggggacaagtttgtacaaaaaagcaggct ATGCACTCGTTTGCTTCTCTTCTC	5'-ggggaccactttgtacaagaaagctgggt CTAACAAGAGGCCACCGAAGG
pgx28C	172944	5'-ccaactttgtacaaaaaagcaggct ATGTCTGTCTTCAAGGCATC	5'-ccaactttgtacaagaaagctgggt TTAATTAGATAATGGGCTGTAGGTCC
suc32A	198063	5'-ggggacaagtttgtacaaaaaagcaggct ATGAAGCTTCAAACGGCTTCC	5'-ggggaccactttgtacaagaaagctgggt TCACCGAACCCAAGTACTCAAC
xeg12A	52011	5'-ccaactttgtacaaaaaagcaggct ATGAAGGTTCTCGCTCTTTCC	5'-ccaactttgtacaagaaagctgggt TTACTCGATGGAAACGGAGTAC
xyl3D	205670	5'-ccaactttgtacaaaaaagcaggct ATGGCGCACTCAATGTCTCG	5'-ccaactttgtacaagaaagctgggt CTACTCCTTCCCCGGCC
xyn10A	57436	5'-ccaactttgtacaaaaaagcaggct ATGGTTCAGATCAAGGTAGCTG	5'-ccaactttgtacaagaaagctgggt CTAGAGAGCATTTGCGATAGCAG
xyn11B	52071	5'-ccaactttgtacaaaaaagcaggct ATGCTCACCAAGAACCTTCTCC	5'-ccaactttgtacaagaaagctgggt TTACTGAACAGTGATGGACGAAGATC

*The attB1 and attB2 recombination sequences are shown in lowercase whereas the gene-specific sequences are presented in uppercase.

Supplementary Table 5. Sequences of oligonucleotide primers used to amplify genes for expression in *Nicotiana benthamiana*

Gene name	Protein ID	Forward primer*	Reverse primer*
abf54B	200605	5′-ggggacaagtttgtacaaaaaagcaggcttg GGCCCCTGTGACATCTACGAAGC	5′-ggggaccactttgtacaagaaagctgggtc CGAAGCAAACGCCGTCTCAATC
abn43A	203143	5′-ggggacaagtttgtacaaaaaagcaggcttg TCGGGTGTTTGTACCACCCATG	5′-ggggaccactttgtacaagaaagctgggtc GAAGTCGATCGTGTTCCACCC
agu67A	56619	5′-ggggacaagtttgtacaaaaaagcaggcttg GAGGATGGGTACGATGGCTG	5′-ggggaccactttgtacaagaaagctgggtc ATCACCCACAGATAGTGTCCAC
axh62A	55136	5′-ggggacaagtttgtacaaaaaagcaggcttg AAATGCGCTCTTCCGTCGACA	5′-ggggaccactttgtacaagaaagctgggtc CTGCTTCAAGGTAAGAACTCC
bgl3A	56782	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGAGGTTCACTTTGATCGAG	5′-ggggaccactttgtacaagaaagctgggtc GTGAACAGTAGGCAGAGACGCCCG
cbh7A	53159	5′-ggggacaagtttgtacaaaaaagcaggcttg CAGCAAGCCGGAACGCTCAC	5′-ggggaccactttgtacaagaaagctgggtc TGCGGAAGCGCTGAAGGTCG
cbh7B	51773	5′-ggggacaagtttgtacaaaaaagcaggcttg CAGCAGGTTGGCACCTACACCA	5′-ggggaccactttgtacaagaaagctgggtc CAAACACTGCGAGTAGTACGC
egl12A	211053	5′-ggggacaagtttgtacaaaaaagcaggcttg CAGACGATGTGCTCTCAATATG	5′-ggggaccactttgtacaagaaagctgggtc GTTGACACTGGCGGTCCAGTT
egl5B	209376	5′-ggggacaagtttgtacaaaaaagcaggcttg GCGTCTGTGTTTGAATGGTTCG	5′-ggggaccactttgtacaagaaagctgggtc GAGATACGTCTCCAGGATATCC
inu32A	52928	5′-ggggacaagtttgtacaaaaaagcaggcttg AATGATTACCGTCCTTCATA	5′-ggggaccactttgtacaagaaagctgggtc TTCAAGTGAAACACTCCGCAC
lac35A	51764	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGAAGCTTTCCTCCGCTTGT	5′-ggggaccactttgtacaagaaagctgggtc GTATGCACCCTTCCGCTTCTTG
mel27A	37736	5′-ggggacaagtttgtacaaaaaagcaggcttg TCTATCGAGCAGCCCAGCCTCCT	5′-ggggaccactttgtacaagaaagctgggtc GGCCAACTGAACGCCACTTGGGA
mel27B	207264	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGCGGTGGCTTCTCACCTC	5′-ggggaccactttgtacaagaaagctgggtc ACATTGCCCTCCCACCACCA
mnd2A	138876	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGATCCTGCAACCGGTGTCT	5′-ggggaccactttgtacaagaaagctgggtc CTCGCGCATCGACTTCTTCCTC
pga28A	214598	5′-ggggacaagtttgtacaaaaaagcaggcttg GCTCCTGCGCCTTCCCGCGTCTCCG	5′-ggggaccactttgtacaagaaagctgggtc CTGATCGCAAGAAGCACCAGA
pga28E	46255	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGGTGACTTCTAGCTCGGTGA	5′-ggggaccactttgtacaagaaagctgggtc ACAGCTGATATCGTCCGGCACA
pga28II	182156	5′-ggggacaagtttgtacaaaaaagcaggcttg TCTCCTATCGAAGCTCGAGAC	5′-ggggaccactttgtacaagaaagctgggtc ACAAGAGGCCACCGAAGGGAAG
pgx28C	172944	5′-ggggacaagtttgtacaaaaaagcaggcttg GTTCCACACTCCAGCAGAGCAT	5′-ggggaccactttgtacaagaaagctgggtc ATTAGATAATGGGCTGTAGG
rhg28B	211163	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGCTTCTCGACAAGCTCTCT	5′-ggggaccactttgtacaagaaagctgggtc CACGTAGCAGACACCCTGCTCGCC
suc32A	198063	5′-ggggacaagtttgtacaaaaaagcaggcttg GCCTCGCCTTCCATGCAGAC	5′-ggggaccactttgtacaagaaagctgggtc CCGAACCCAAGTACTCAACGC
xeg12A	52011	5′-ggggacaagtttgtacaaaaaagcaggcttg GCCTCTATCAGCCGCCGTAGCGAC	5′-ggggaccactttgtacaagaaagctgggtc CTCGATGGAAACGGAGTACTCG
xyl3D	205670	5′-ggggacaagtttgtacaaaaaagcaggcttg ATGGCGCACTCAATGTCTCGT	5′-ggggaccactttgtacaagaaagctgggtc CTCCTTCCCCGGCCACTTCAGCA
xyn10A	57436	5′-ggggacaagtttgtacaaaaaagcaggcttg CCCATTGAACCCCGTCAGGC	5′-ggggaccactttgtacaagaaagctgggtc GAGAGCATTTGCGATAGCAGTG
xyn11B	52071	5′-ggggacaagtttgtacaaaaaagcaggcttg GTTCCCCACGACTCTGTGGC	5′-ggggaccactttgtacaagaaagctgggtc CTGAACAGTGATGGACGAAGATC

*The attB1 and attB2 recombination sequences are shown in lowercase whereas the gene-specific sequences are presented in uppercase.

Supplementary Table 6. Substrates used for biochemical assays

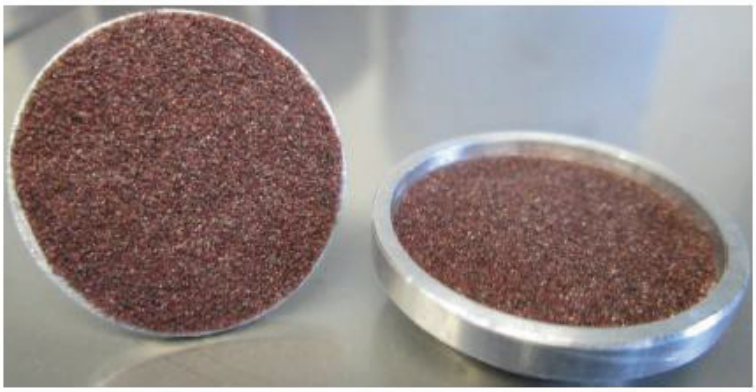
Enzyme Activity	Substrate	Supplier	Substrate Concentration Used
alpha-galactosidase			
alpha-glucosidase			
alpha-glucuronidase			
arabinoxylan-arabinofuranhydrolase	4-nitrophenyl-L-arabinofuranoside	Sigma-Aldrich	5mM
arabinanase	arabinan	Megazyme	1%
arabinofuranosidase	4-nitrophenyl-L-arabinofuranoside	Sigma-Aldrich	5mM
beta-fructofuranosidase			
beta-galactosidase			
beta-glucosidase			
beta-mannosidase	4-nitrophenyl-beta-D mannoside	Sigma-Aldrich	5mM
cellobiohydrolase	4-nitrophenyl-cellobioside	Sigma-Aldrich	5mM
endoglucanase	CM-cellulose 4M	Megazyme	1%
endo-polygalacturonase	polygalacturonic acid	Sigma-Aldrich	1%
exo-polygalacturonase	polygalacturonic acid	Sigma-Aldrich	1%
inulinase	chicory inulin	Sigma-Aldrich	1%
rhamnogalacturonase			
xylanase	Birchwood xylan	Sigma-Aldrich	1%
xyloglucanase	xyloglucan (amyloid)	Megazyme	1%
xylosidase	4-nitrophenyl-beta-D-xyloside	Glycosynth	5mM

Supplementary Table 7. List of *M. thermophila* secreted proteins identified by mass spectrometry in pooled culture supernatants*

Gene name	Predicted enzyme activity	JGI protein ID
Gan53A	Arabinogalactanase	43163
Axh43G	Arabinoxylan arabinofuranohydrolase	2301869
Axh62A	Arabinoxylan arabinofuranosidase	98003
Cdh	Cellobiose dehydrogenase	81925
Cat3	Catalase	80916
Cbh6A	Cellobiohydrolase	66729
Cbh7	Cellobiohydrolase	109566
Cdh	Cellobiose dehydrogenase	111388
Cdp	Cuticle-degrading protease	2303011
Chdh	Choline dehydrogenase	2305709
Chi18A	Endochitinase	50608
Egl5A	Endoglucanase	86753
Egl6B	Endoglucanase	2303045
Egl7	Endoglucanase	111372
Egl7	Endoglucanase	116157
Egl45D	Endoglucanase	76901
Exg55	Exo-1,3-β-glucanase	2305407
Exg55	Exo-1,3-β-glucanase	2294895
FaeB	Feruloyl esterase	33936
Gal5F	Galactanase	2301720
GalM1	Aldose 1-epimerase	103702
Gan53A	Arabinogalactanase	43163
Gub16	Hypothetical GH16 β-glucanase	90182
GusB2	β-Glucuronidase	114497
-	Hypothetical GH16 glycosidase	2315557
Man5A	β-Mannanase	84297
Ply1B	Pectate lyase	52463
Pmo	Polysaccharide monooxygenase	80312
Pmo	Polysaccharide monooxygenase	92668
Pmo	Polysaccharide monooxygenase	112089
Sdh	L-Sorbose 1-dehydrogenase	2299749
Tyr	Uncharacterized tyrosinase	60685
-	Uncharacterized protein	2299856
-	Uncharacterized protein	2298860
-	Unknown protein	58635
-	Unknown protein	2297957
-	Unknown protein	2310464
-	Unknown protein	2311687
-	Unknown protein	112399
-	Unknown protein	73270
-	Unknown protein	102138
Xeg74A	Xyloglucanase	116384
Xyn10	Xylanase	116553
Xyn10	Xylanase	112050
Xyn10F	Xylanase	52904
Xyn11	Xylanase	89603
Xyn11B	Xylanase	100068

*The cultures were prepared with various pretreated wood biomass types as sole carbon source

(a)



(b)

Supplementary Figure 1. A customized aluminium collar for the bottom plate of a rotational parallel plate rheometer for measuring rheology of pretreated wood biomass. (a) The top and bottom plates roughened with 100 grid sand paper. (b) Plates mounted on Anton Paar Physica MCR500 rotational rheometer

References

1 Türe S, Uzun D, Ture IE, 1997. The potential use of sweet sorghum as a non-polluting source of energy. *Energy* 22:17–9

2 de la Rue du Can S, Price L, 2008. Sectoral trends in global energy use and greenhouse gas emissions. *Energy Policy* 36:1386–1403

3 Sheehan J, Cambreco V, Duffield J, Garboski M, Shapouri H, 1998. An overview of biodiesel and petroleum diesel life cycles. *A report by US Department of Agriculture and Energy*. p.1–35

4 Reijnders L, Huijbrets MAJ, 2009. Biofuels for road transport: A seed to wheel perspective; *Green Energy and Technology,* Springer-Verlag, London Ltd

5 International Energy Agency, 2008. World Energy Outlook

6 Bernstein M, Bromley P, Hagen J, Hassell S, Lempert R, Munoz J, Robalino D, 1999. RAND, Developing Countries & Global Climate Change: Electric Power Options for Growth, Report prepared for the Pew Center on Global Climate Change

7 Greene L, Schafer A, 2003. Reducing Greenhouse Gas Emissions from US Transportation, Report prepared for the Pew Center on Global Climate Change

8 Demirbas A, 2007. Progress and trends in biofuels. *Prog Enger Combust Sci* 33(1):1-18.

9 Pacala S, Socolov R, 2004. Stabilization wedges: solving the climate problem for the next 50 years with current technologies. *Science* 305:968-971

10 Farrel AE, Plevin RJ, Turner BT, Jones AD, O'hare M, Kammen DM, 2006. Ethanol can contribute to energy and environmental goals. *Science* 311:506-508

11 Service R, 2007. Cellulosic ethanol: biofuel researchers prepare to reap a new harvest. *Science* 135:1488-1491

12 Gray KA, Zhao L, Emptage M, 2006. Bioethanol. *Curr Opin Chem Biol* 10(2): 141–146

13 Wyman CE, 1996. Handbook on Bioethanol: Production and utilization. Wasington: Taylor Francis

14 Jorgensen H, Kristensen JB, Felby C, 2007. Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities. *Biofuels Bioprod Bioref* 1:119–134

15 Perez J, Dorado JM, Rubia TD, Martinez J, 2002. Biodegradation and biological treatment of cellulose, hemicellulose and lignin: An overview *Int Microbiol* 5:53–63

16 Kumar P, Barrett DM, Delwiche MJ, Stroeve P, 2009. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Ind Eng Chem Res* 48(8): 3713–3729

17 Sarkar P, Bosneaga E, Auer M, 2009. Plant cell walls throughout evolution: towards a molecular understanding of their design principles. *J Exp Bot* 60:3615-3635

18 Limayem A, Ricke SC, 2012. Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Progress Energy Combust Science* 38:449-467

19 Menon V, Rao M, 2012. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Progress Energy Combust Sc* 38:522-550

20 Limayem A, Ricke SC, 2012. Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Progress Energy Combust Science* 38:449-467

21 Mohanram S, Amat D, Choudhary J, Arora A, Nain L, 2013. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. *Sust Chem Processes*, 1:15

22 Sánchez OJ, Cardona CA 2008. Trends in biological production of fuel ethanol from different feedstocks. *Bioresour Technol*, 99:5270–5295.

23 Galbe M, Wallberg O, Zacchi G, 2011. 6.47- Techno-Economic aspects of ethanol production from lignocellulosic agricultural crops residue, Publisher: Clarence Yue

24 Solomon BD, 2010. Biofuels and sustainability. *Ann N Y Acad Sci* 1185:119-134.

25 Rosegrant, M. W. 2008. Biofuels and grain prices: impacts and policy responses. Washington, DC: International Food Policy Research Institute

26 Hein L, Leemans R, 2012. The impact of first-generation biofuels on the depletion of the global phosphorus reserve. *Ambio* 41(4): 341–349.

27 Crutzen PJ, Mosier AR, Smith KA, Winiwarter W, 2007. N2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. *Atmosph Chem Physical Discuss* 7:11191–11205

28 Wicke B, Dornburg V, Junginger M, Faaij A, 2008. Different palm oil production systems for energy purposes and their greenhouse gas implications. *Biomass Bioenergy*. ;32:1322–1337

29 Patzek TW, 2004. Thermodynamics of the corn-ethanol biofuel cycle. *Critical Rev Plant Science*. 23:519–567

30 Tilman D, Socolow R, Foley JA, Hill J, Larson E, Lynd L, Pacala S, Reilly J, Searchinger T, Somerville C, Williams R, 2009. Beneficial biofuels—The food, energy, and environment trilemma. *Science*. ;325:270–27

31 Cushion E, Whiteman A, Dieterle G, 2010. Bioenergy development; issues and impacts for poverty and natural resource management. Washington, DC: World Bank.

32 Haberl H, Beringer T, Bhattacharya SC, Erb K-H, Hoogwijk M, 2010. The global technical potential of bio-energy in 2050 considering sustainability constraints. *Current Opin Environ Sustainability*, 2:394–403

33 Vries SC, Ven GWJ, Ittersum MK, Giller KE, 2010. Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. *Biomass Bioenergy* 34:588–601

34 Leemans R, Amstel AR, Battjes C, Kreileman GJJ, Toet AMC, 1996. The land cover and carbon cycle consequences of large-scale utilizations of biomass as an energy source. *Global Environ Change* 6:335–357

35 Haigler CH, Ivanova-Datcheva M, Hogan PS, Salnikov VV, Hwang S, Martin K, Delmer DP, 2001. Carbon partitioning to cellulose synthesis. *Plant Mol Biol* 47: 29–51

36 Wang M, Wu M, Huo H, 2007. Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environ Res Lett* 2:1–13.

37 DeBolt S, Campbell JE, Smith R Jr, Montross M, Stork J, 2009. Life cycle assessment of native plants and marginal land for bioenergy agriculture in Kentucky as a model for southeastern USA *Global Change Biology- Bioenergy*, 4: 380-389

38 Lebo SE Jr, Gargulak JD, McNally TJ, 2001. *Lignin. Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Inc

39 Klemm D, Philipp B, Heinze T, Heinze U, Wagenknecht W, 1998. Comprehensive cellulose chemistry; *Wiley-VCH Verlag GmbH*: Berlin, Germany

40 Zhang Y-HP, Lynd LR, 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol Bioeng* 88:797

41 Zhang Y-HP, Lynd LR, 2005. Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. *Biomacromol* 6:1510-1515

42 Arantes V, Saddler JN, 2010. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol Biofuels*, 3-4:1-11

43 Beguin P, Aubert JP, 1994. The biological degradation of cellulose. *FEMS Microbiol Rev* 13:25–58

44 Wikipedia [http://en.wikipedia.org/wiki/Bacterial_cellulose]

45 Kuhad RC, Singh A, Eriksson KE, 1997. Microorganisms and enzymes involved in degradation of plant fiber cell walls. *Adv Biochem Eng/Biotechnol* 57:45–125

46 Scheller HV, Ulvskov P, 2010. Hemicelluloses. *Annu Rev Plant Biol* 61:263–289

47 Xu N, Zhang W, Ren S, Liu F, Zhao C, Liao H, Xu Z, Huang J, Li Q, Tu Y, Yu B, Wang Y, Jiang J, Qin J, Peng L, 2012. Hemicelluloses negatively affect lignocellulose crystallinity for high biomass digestibility under NaOH and H2SO4 pretreatments in Miscanthus. *Biotechnol Biofuels* 5:58

48 John MJ, Anandjiwala RD, 2008. Recent developments in chemical modification and characterization of natural fiber-reinforced composites. *Polymer Composites*, 29(2): 187-207

49 Biofuel outlook blog: Biomass, biofuels and biochemicals; [http://biofueloutlook.blogspot.ca/2010/09/enzymatic-deconstruction-of-xylan-for.html]

50 Pu Y, Kosa M, Kalluri UC, Tuskan GA, Ragauskas AJ, 2011. Challenges of the utilization of wood polymers: how can they be overcome? *Appl Microbiol Biotechnol* 91:1525–1536

51 Balakshin MY, Capanema EA, Chang HM, 2007. MWL fraction with a high concentration of lignin–carbohydrate linkages: isolation and 2D NMR spectroscopic analysis. *Holzforschung* 61:1–7

52 BalakshinMY, Capanema EA, Chen CL, Gracz HS, 2003. Elucidation of the structures of residual and dissolved pine kraft lignins using an HMQC NMR technique. *J Agric Food Chem* 51:6116–6127

53 Pandey A, 2009. Handbook of plant-based biofuels. Florida: *CRC press*

54 Perez J, Dorado JM, Rubia TD, Martinez J, 2002. Biodegradation and biological treatment of cellulose, hemicellulose and lignin: An overview. *Int Microbiol* 5:53–63

55 Hatfield R, Vermerris W, 2001. Lignin formation in plants. the dilemma of linkage specificity, *Plant Physiol* 126:1351-1357

56 Goring DAI, 1963. Thermal softening of lignin, hemicelllulose and cellulose. *Pulp & Paper Magazine Canada*. 64(12):517-527

57 Zhang Y, 2010. Hydrothermal liquefaction to convert biomass into crude oil. *Biofuels from Agricultural Wastes and Byproducts.* Ed: Blaschek HP, Ezeji TC and Scheffran J, Blackwell Publishing. ISBN: 978-0-813-80252-7

58 Kamm B, Kamm M, 2004. Principles of biorefineries. *Appl Microbiol Biotechnol* 64(2):137–146

59 US DOE; Pacific Northwest National Laboratory, 2007. Top value-added chemicals from biomass; Volume II—Results of screening for potential candidates from biorefinery lignin; Edited by Holladay JE, Bozell JJ, White JF and Johnson D.

60 McKendry P, 2002. Energy production from biomass (part 1): overview of biomass. *Biores Technol* 83:37–46

61 Bridgwater AV, 2003. Renewable fuels and chemicals by thermal processing of biomass. *Chem Eng J* 91:87–102

62 Maschio G, Lucchesi A, Stoppato G, 1994. Production of syngas from biomass. *Biores Technol* 48:119–26

63 Koppejan J, van Loo S (editors), 2012.The handbook of biomass combustion and co-firing; *Routledge, Taylor and Francis Group*

64 NREL, Renewable energy sources charged with energy from the sun and originated from the earth-moon interaction – Vol.1 – Direct combustion of biomass; Ralp P. Overend; *Encyclopedia of life support systems*

65 Zinoviev S, Mueller-Langer F, Das P, Bertero N, Fornasiero P, Kaltschmitt M, Centi G, Miertus S, 2010. Next-generation biofuels: survey of emerging technologies and sustainability issues. *Chem Sus Chem* 3(10):1106–1133

66 Demirbas A, Arin G, 2002. An overview of biomass pyrolysis. Energy Sources 5:471–478

67 Demirbas A, 2007. Progress and recent trends in biofuels. *Progr Energy Combust Sc* 33:1–18

68 Wang D, Czernik S, Montane D, Mann M, Chornet E, 1997. Biomass to hydrogen via fast pyrolysis and catalytic steam reforming of the pyrolysis oil or its fractions. *Ind Eng Chem Res* 36:1507–18

69 Demirbas A, 2009. Biorefineries: current activities and future development. *Energ Convers Manage* 50:2782–2801

70 Zhang Y, 2010. Hydrothermal liquefaction to convert biomass into crude oil (Chapter 10). *Biofuels from agricultural wastes and byproducts* ; Edited by Hans P. Blaschek, Thaddeus C. Ezeji and Jürgen Scheffran J; Blackwell Publishing. ISBN: 978-0-813-80252-7

71 NREL/ Pacific Northwest National Laboratory, 2013. Whole algae hydrothermal liquefaction technology pathway; Technical Report: NREL/TP-5100-58051, PNNL-22314. http://www.nrel.gov/docs/fy13osti/58051.pdf

72 Kumar R, Singh S, Singh OV, 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J Ind Microbiol Biotechnol* 35 (5): 377–91

73 Van Eylen D, van Dongen F, Kabel M, de Bont J, 2011. Corn fiber, cobs and stover: Enzyme-aided saccharification and co-fermentation after dilute acid pretreatment. *Bioresour Technol*. doi:10.1016/j.biortech.2011.02.049

74 Sakuragi H, Kuroda K, Ueda M, 2011. Molecular breeding of advanced microorganisms for biofuel production. *J Biomed Biotechnol* doi:10.1155/2011/416931

75 Várnai A, Siika-aho M, and Viikari L, 2010. Restriction of the enzymatic hydrolysis of steam-pretreated spruce by lignin and hemicellulose. Enzyme. *Microb Technol* 46: 185-193

76 Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M et al., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour Technol* 96:673–686

77 Chen WH, Pen BL, Yu CT, Hwang WS, 2011. Pretreatment efficiency and structural characterization of rice straw by an integrated process of dilute-acid and steam explosion for bioethanol production. *Biores Technol* 102:2916–2924

78 Sun Y, Cheng JJ, 2005. Dilute acid pretreatment of rye straw and Bermuda grass for ethanol production. *Biores Technol* 96:1599–1606

79 Zhang J, Ma X, Yu J, Zhang X, Tan T, 2011. The effects of four different pretreatments on enzymatic hydrolysis of sweet sorghum bagasse. *Biores Technol* 102(6):4585-9

80 Kaar WE, Gutierrez CV, Kinoshita CM, 1998. Steam explosion of sugarcane bagasse as a pre-treatment for conversion to ethanol. *Biomass Bioenergy* 14:277–287

81 Chen H, Liu L, 2007. Unpolluted fractionation of wheat straw by steam explosion and ethanol extraction. *Bioresour Technol* 98:666–676

82 Bals B, Rogers C, Jin M, Balan V, Dale B, 2010. Evaluation of ammonia fibre expansion (AFEX) pretreatment for enzymatic hydrolysis of switchgrass harvested in different seasons and locations. *Biotechnol Biofuels* 3:1

83 Alizadeh H, Teymouri F, Gilbert TI, Dale BE, 2005. Pretreatment of switchgrass by ammonia fiber explosion (AFEX). *Appl Biochem Biotechnol* 6:1133–1141

84 Murnen HK, Balan V, Chundawat SPS, Bals B, da Costa Sousa L, Dale BE, 2007. Optimization of ammonia fiber expansion (AFEX) pretreatment and enzymatic hydrolysis of Miscanthus × giganteus to fermentable sugars. *Biotechnol Prog* 23:846-850

85 Mirahmadi K, Kabir MM, Jeihanipour A, Karimi K, Taherzadeh MJ, 2010. Alkaline pretreatment of spruce and birch to improve bioethanol and biogas production. *BioResources* 6:928–938

86 Li H-Q, Li C-L, Sang T, Xu J, 2013. Pretreatment of *Miscanthus lutarioriparious* by liquid hot water for efficient ethanol production. *Biotechnol Biofuels* 6:76

87 Suryawati L, Wilkins MR, Bellmer DD, Huhnke RL, Maness NO, Banat IM, 2009. Effect of hydrothermolysis process conditions on pretreated switchgrass composition and ethanol yield by SSF with *Kluyveromyces marxianus* IMB4. *Process Biochem* 6:540–545

88 Pérez J, Ballesteros I, Ballesteros M, Sáez F, Negro M, Manzanares P, 2008. Optimizing liquid hot water pretreatment conditions to enhance sugar recovery from wheat straw for fuel-ethanol production. *Fuel* 6:3640–3647

89 Kim S, Holtzapple MT, 2005. Lime pretreatment and enzymatic hydrolysis of corn stover. *Bioresour Technol* 6:1994–2006

90 Xu J, Cheng JJ, Sharma-Shivappa RR, Burns JC, 2010. Sodium hydroxide pretreatment of switchgrass for ethanol production. *Energy Fuels* 6:2113–2119

91 Brosse N, Sannigrahi P, Ragauskas A, 2009. Pretreatment of *Miscanthus x giganteus* Using the Ethanol Organosolv Process for Ethanol Production. *Ind Eng Chem Res* 6:8328–8334

92 Zhu J, Gleisner R, Scott C, Luo X, Tian S, 2011. High titer ethanol production from simultaneous enzymatic saccharification and fermentation of aspen at high solids: a comparison between SPORL and dilute acid pretreatments. *Bioresour Technol* 6:8921–8929

93 Lan TQ, Gleisner R, Zhu JY, Dien BS, Hector RE, 2013. High titer ethanol production from SPORL-pretreated lodgepole pine by simultaneous enzymatic saccharification and combined fermentation. *Bioresour Technol* 6:291–297

94 Hideno A, Inoue H, Tsukahara K *et al*., 2009. Wet disk milling pretreatment without sulfuric acid for enzymatic hydrolysis of rice straw. *Biores Technol* 100(10):2706–2711

95 Jin Y, Jameel H, Chang H-M, Phillips R, 2010. Green liquor pretreatment of mixed hardwood for ethanol production in a repurposed kraft pulp mill. *J Wood Chem Technol* 30(1):86-104

96 Feng L, Chen ZL, 2008. Research progress on dissolution and functional modification of cellulose in ionic liquids. *J Mol Liq* 142:1–5

97 Sánchez C, 2009. Lignocellulosic residues: biodegradation and bioconversion by fungi. *Biotech Adv* 27:185 – 194

98 Kurakake M, Ide N, Komaki T, 2007. Biological pretreatment with two bacterial strains for enzymatic hydrolysis of office paper. *Curr Microbiol* 54:424-428

99 Salvachúa D, Prieto A, López-Abelairas M, Lu-Chau T, Martínez AT, Martínez MJ, 2011. Fungal pretreatment: an alternative in second-generation ethanol from wheat straw. *Bioresour Technol* 102:7500 - 7506

100 Liu H, Zhu JY, 2010. Eliminating inhibition of enzymatic hydrolysis by lignosulfonate in unwashed sulfite-pretreated aspen using metal salts. *Biores Technol* 101: 9120–9127

101 Ani Tejirian, Feng Xu, 2010. Inhibition of Enzymatic Cellulolysis by Phenolic Compounds, *Enzyme Microbial Technol.* doi:10.1016/j.enzmictec.2010.11.004

102 Ximenes E, Kim Y, Mosier N, Dien B, Ladisch M, 2011. Deactivation of cellulases by phenols. *Enzyme Microb Technol* 48: 54–60

103 Du B, Sharma LN, Becker C, Chen SF, Mowery RA, van Walsum GP, Chambliss CK, 2010. Effect of varying feedstock-pretreatment chemistry combinations on the formation and accumulation of potentially inhibitory degradation products in biomass hydrolysates. *Biotechnol Bioeng* 107(3):430-40

104 Pienkos PT, Zhang M, 2009. Role of pretreatment and conditioning processes on toxicity of lignocellulosic biomass hydrolysates. *Cellulose* 16:743–762

105 Zheng Y, Pan Z, Zhang R, 2009. Overview of biomass pretreatment for cellulosic ethanol production. *Int J Agric Biol Eng* 2(3):51-68

106 Phillips RB, Jameel H, Chang HM, 2013. Integration of pulp and paper technology with bioethanol production*. Biotechnol Biofuels* 6:13

107 Wyman CE, Dale BE, Elander RT, Holtzapple M, Ladisch MR, Lee YY, 2005. Coordinated development of leading biomass pretreatment technologies. *Bioresour Technol* 96:1959–1966.

108 Yang B, Wyman CE, 2008. Pretreatment: the key to unlocking low-cost cellulosic ethanl. *Biofuels Bioprod Bior* 2:26–40

109 Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ, 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresour Technol* 101:4851–4861

110 Sticklen MB, 2008. Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. *Nat Rev Genet* 9:433–443.

111 Dodd D, Cann IK, 2009. Enzymatic deconstruction of xylan for biofuel production. *Global Change Biol Bioenergy* 1(1): 2–17

112 Martinez AT, Speranza M, Ruiz-Duenas FJ, Ferreira P, Camarero S, *et al*., 2005. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *Int Microbiol* 8:195–204

113 Akin DE, Rigsby LL, Sethuraman A, 1995. Alterations in the structure, chemistry and biodegradation of grass lignocellulose treated with white rot fungi *Ceriporiopsis subvermispora* and *Cyathus stercoreus*. *Appl Environ Microbiol* 61:1591–1598

114 Gold MH, Alic M, 1993. Molecular biology of the lignin degrading basidiomycetes *Phanerochaete chrysosporium*. *Microbiol Rev* 57:605–622

115 Stukenbrock EH, McDonald BA, 2009. Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions, *Current Review MPMI* 22(4):371–380

116 de Jonge R, Bolton MD, Thomm BPHJ, 2011. How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. *Curr Opinion Plant Biol* 14(4):400–406

117 Martinez D, Berka RM, Henrissat B, *et al*., 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnol* 26(5): 553–60

118 Berka RM, Grigoriev IV, Otillar R, Salamov A, Grimwood J, Reid I *et al,* 2011. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris. Nature Biotechnol* 29:922–927

119 van den Brink J, van Muiswinkel GCJ, Theelen B, Hinz SWA, de Vries RP, 2013. Efficient plant biomass degradation by thermophilic fungus *Myceliophthora heterothallica. Applied Env Microbiol* 79(4):1316–1324

120 Adav SS, Li AA, Manavalan A, Punt P, Sze SK, 2010. Quantitative iTRAQ secretome analysis of *Aspergillus niger* reveals novel hydrolytic enzymes. *J Proteome Res* 9: 3932–3940

121 Tsang A, Butler G, Powlowski J, Panisko EA, Baker SE, 2009. Analytical and computational approaches to define the *Aspergillus niger* secretome. *Fungal Gen Biol* 46(1): S153-S160

122 Lum G, Min XJ, 2011. FunSecKB: the Fungal Secretome KnowledgeBase. *Database* (Oxford). doi: 10.1093/database/bar001

123 US DOE Joint Genome Institute [http://genome.jgi.doe.gov]

124 Rubin EM, 2008. Genomics of cellulosic biofuels. *Nature* 454: 841–845

125 Carbohydrate-active enzymes database; CAZy [http://www.cazy.org]

126 Henrissat B, Bairoch A, 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316:695-696

127 Levasseur A, Piumi F, Coutinho PM, Rancurel C, Asther M, Delattre M, Henrissat B, Pontarotti P, Asther M, Record E, 2008. FOLy: an integrated database for the classification and functional annotation of fungal oxidoreductases potentially involved in the degradation of lignin and related aromatic compounds. *Fungal Genet Biol* 45(5): 638-45

128 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*. *Fung Genet Biol* 46:427-435

129 Henrissat B, Bairoch A, 1993. New families in the classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem J* 293:781-788

130 Henrissat B, 1991. A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem J* 280:309-316

131 Koshland DE, 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev Camb Philos Soc* 28: 416-436

132 Sinnott ML, 1990. Catalytic mechanisms of enzymic glycosyl transfer. *Chem Rev* 90:1171-1202

133 Hu J, Arantes V, Saddler JN, 2011. The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnol Biofuels* 4:36

134 Berlin A, Gilkes N, Kilburn D, Bura R, Markov A, Skomarovshy A, Okunev O, Gusakov A, Maximenko V, Gregg D, Sinitsyn A, Saddler J, 2005. Evaluation of novel fungal cellulase preparations for ability to hydrolyse softwood substrates – evidence for the role of accessory enzymes. *Enzyme Microb Technol* 37(2):175-184

135 Zhang M, Su R, Qi W, He Z, 2009.Enhanced enzymatic hydrolysis of lignocellulose by optimizing enzyme complexes. *Appl Biochem Biotechnol*. 160(5):1407-1414

136 Ekwe E, Morgenstern I, Tsang A, Storms R, Powlowski J, 2013. Non-hydrolytic cellulose active proteins: research progress and potential application in biorefineries. *Ind biotechnol* 9(3): 123-131

137 Pothiraj C, Kanmani P, Balaji P, 2006. Bioconversion of lignocellulose materials. *Mycobiol*, 34(4):159-165

138 Vlasenko E, Schulein M, Cherry J, Xu F, 2010. Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. *Bioresour Technol* 101(7):2405–2411

139 de Vries RP, Nadal M, van den Brink J, Vivas-Duarte DA, Stålbrand H, 2011. Fungal degradation of plant oligo- and polysaccharides. In: Grunwald P (ed) Carbohydrate modifying enzymes and microorganisms. Pan Stanford publishing Pte. Ltd., Singapore

140 Henrissat B, Driguez H, Viet C, Schulein M, 1985. Synergism of cellulases from *trichoderma reesei* in the degradation of cellulose. *Biotechnol* 3:722-726

141 Zhang Y-IP, Lynd LR (2004) Toward an aggregate understanding of enzymatic hydrolysis of cellulose: non-complexed cellulase systems. *Biotechnol Bioeng* 88: 797-824

142 Väljamäe P, Sild V, Nutt A, Petterson G, Johansson G, 1999. Acid hydrolysis of bacterial cellulose reveals different modes of synergistic action between cellobiohydrolases I and endoglucanase I. *Eur J Biochem* 266:327-334

143 Kadam KL, Rydholm EC, McMillan JD, 2004. Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol Prog* 20:698-705

144 Kyriacou A, MacKenzie R, Neufeld RJ, 1987. Detection and characterization of the specific and nonspecific endoglucanases of Trichoderma reesei: evidence demonstrating endoglucanase activity by cellobiohydrolase II. *Enz Microbiol Tech* 9(1):25–32

145 Nutt A, Sild V, Pettersson G, Johansson G, 1998. Progress curves—a mean for functional classification of cellulases. *Eur J Biochem* 258(1):200–206

146 Nummi M, Niku-Paavola ML, Lappalainen A, Enari TM, Raunio V, 1983. Cellobiohydrolase from *Trichoderma reesei*. *Biochemical J* 215:(3):677-683

147 Jalak J, Valjamae P, 2010. Mechanism of initial rapid rate retardation in cellobiohydrolase catalyzed cellulose hydrolysis. *Biotechnol Bioeng* 106(6): 871-883

148 Dadi AP, Schall CA, Varanasi S, 2007. Mitigation of cellulose recalcitrance to enzymatic hydrolysis by ionic liquid pretreatment. *Appl Biochem Biotechnol* 137-140(1-12):407-421

149 Nidetzky B, Zachariae W, Gercken G, Hayn M, Steiner W, 1994. Hydrolysis of cellooligosaccharides by *Trichoderma reesei* cellobiohydrolases: experiemtnal data and kinetic modelling. *Enz Microbiol Tech* 16:43-52.

150 Zhang Y-HP, Lynd LR, 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Non-complexed cellulase systems. *Biotechnol Bioeng* 88:797-824

151 Zhang Y-HP, Lynd LR, 2006. A functionally based model for hydrolysis of cellulose by fungal cellulase. *Biotech Bioeng* 94(5):888 - 898

152 Jalak J, Kurašin M, Teugjas H, Väljamäe P, 2012. Endo-exo synergism in cellulose hydrolysis revisited. *J Biol Chem* 287(34):28802-28815

153 Kurašin M, Väljamäe P, 2011. Processivity of cellobiohydrolases is limited by the substrate. *J Biol Chem* 286(1):169-177

154 van den Brink J, de Vries RP, 2011. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol* 91:1477–1492

155 Woodward J, Wiseman A, 1982. Fungal and other β-glucosidases—their properties and applications. *Enzyme Microb Technol* 4:73–79

156 Lo Leggio L, Welner D, De Maria L, 2013. A structural overview of GH61 proteins – fungal cellulose degrading polysaccharide monooxygenases. *Comput Struct Biotechnol J* 2(3): e201209019

157 Vaaje-Kolstad G, Horn SJ, van Aalten DM, *et al., 2005.* The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation*. J Biol Chem* 280: 28492-28497

158 Vaaje-Kolstad G, Houston DR, Riemen AH, *et al.,* 2005*.* Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* 280:11313-1131

159 Vaaje-Kolstad G, Bøhle LA, Gåseidnes S, *et al.,* 2012. Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high resolution structure of its oxidative CBM33 enzyme. *J Mol Biol* 416(2):239-254

160 Langston JA, Shaghasi T, Abbate E, *et al*., 2011. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microb* 77:7007–7015

161 Vaaje-Kolstad G, Westereng B, Horn SJ, *et al.,* 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330:219-222

162 Beeson WT, Phillips CM, Cate JHD, *et al.,* 2012. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *JACS* 134:890–892

163 Cannella D, Hsieh C-W, Felby C, *et al*., 2012. Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol Biofuels* 5:26

164 Bey M, Zhou S, Poidevin L, *et al.,* 2012. Comparison of two lytic polysaccharide monooxygenases (GH61) from *Podospora anserina* reveals differences upon cello-oligosaccharides oxidation. *Appl Environ Microbiol* 10.1128/AEM.02942-12AEM.02942-12

165 Sweeney MD, Vlasenko E, Abbate E, 2009. Methods for increasing hydrolysis of cellulosic material in the presence of cellobiose dehydrogenase. *United States Patent Application* 20100159536 A1

166 Sweeney MD, Vlasenko E, *2009*. Methods for determining cellulolytic enhancing activity of a polypeptide. *United States Patent Application* 20100159494 A1

167 Horn SJ, Vaaje-Kolstad G, Westereng B, et al., 2012. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels* 5:45

168 McQueen-Mason SJ, Fry SC, Durachko DM, Cosgrove DJ, 1993. The relationship between xyloglucan endotransglycosylase and in-vitro cell wall extension in cucumber hypocotyls. *Planta* 190:327-331

169 Li L-C, Cosgrove DJ, 2001. Grass group I pollen allergens (β-expansins) lack proteinase activity and do not cause wall loosening via proteolysis. *Eur J Biochem* 268: 4217–4226

170 Cosgrove DJ, *2001*. Enhancement of accessibility of cellulose by expansins. *US Patent* 6326470 B1

171 McQueen-Mason S, Cosgrove DJ, 1994. Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *PNAS USA* 91(14):6574-6578

172 Zenoni S, Reale L, Tornielli GB, *et al.*, 2004. Downregulation of the Petunia hybrida α-expansin gene PhEXP1 reduces the amount of crystalline cellulose in cell walls and leads to phenotypic changes in petal limbs. *Plant Cell* 16:295–308

173 Saloheimo M, Paloheimo M, Hakola S, *et al*., 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur J Biochem* 269(17):4202–4211

174 Wang M, Cai J, Huang L, *et al.,* 2010. High-Level Expression and Efficient Purification of Bioactive Swollenin in *Aspergillus oryzae. Appl Biochem Biotechnol* 162:2027–2036

175 Berlin A, Maximenko V, Gilkes N, Saddler J, 2007. Optimization of Enzyme Complexes for Lignocellulose Hydrolysis, Biotechnol. *Bioeng* 97(2):287-296

176 Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Aslam N, Walton JD, 2010. Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. *Biotechnol Bioeng* 106(5):707-720

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

178 de Vries RP, Visser J, 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev* 65(4):497–522,

179 Rättö M, Siika-aho M, Buchert J, Valkeajävi A, Viikari L, 1993. Enzymatic hydrolysis of isolated and fibre-bound galactoglucomannans from pine-wood and pine kraft pulp. *Applied Microbiol Biotechnol* 40(2-3):449-454

180 Henrissat B, Bairoch A, 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 293(3):781–788

181 Polizeli ML, Rizzatti AC, Monti R, Terenzi HF, Jorge JA, Amorim DS, 2005. Xylanases from fungi: properties and industrial applications. *Appl Microbiol Biotechnol* 67(5):577–591

182 Pollet A, Delcour JA, Courtin CM, 2010. Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families. *Crit Rev Biotechnol* 30(3):176–191

183 Biely P, Vrsanska M, Tenkanen M, Kluepfel D, 1997. Endo-beta-1,4-xylanase families: differences in catalytic properties. *J Biotechnol* 57(1–3):151–166

184 Dodd D, Cann IKO, 2009. Enzymatic deconstruction of xylan for biofuels production. *GCB Bioenergy* 1:2–17

185 Sun J, Tian C, Diamond S, Glass NL, 2012. Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in neurospora crassa. *Eukaryotic Cell* 11(4):482 – 493

186 Alba-Lois L, Segal-Kischinevzky C, 2010. Yeast fermentation and the making of beer and wine. *Nature Educ* 3(9):17

187 Pasteur L, 1857. Mémoire sur la fermentation alcoolique.*Comptes Rendus Séances de l'Academie des Sciences* 45 (913–916) : 1032–1036

188 Barnett JA, 1998. A history of research on yeast 1: Work by chemists and biologists, 1789–1850. *Yeast* 14:1439–1451

189 Barnett, JA, 2000. A history of research on yeast 2: Louis Pasteur and his contemporaries, 1850–1880. *Yeast* 16:755–771

190 Soleimani M, Tabil L, Panigrahi S, 2006. Bio-production of a Polyalcohol (xylitol) from Lignocellulosic Resources: A Review. *Canadian Society Bioeng* 06-106: 1-19

191 Kuhad RC, Gupta R, Khasa YP, Singh A, Zhang P Y-H, 2011. Bioethanol production from pentose sugars: Current status and future prospects. *Renewable Sustain Energy Rev* 15(9):4950–4962

192 Rosenberg SL, 1980. Fermentation of pentose sugars to ethanol and other neutral products by microorganisms. *Enzyme Microbial Technol* 2(3):185–193

193 Oreb M, Dietz H, Farwick A, Boles E, 2012. Novel strategies to improve co-fermentation of pentoses with D-glucose by recombinant yeast strains in lignocellulosic hydrolysates. *Bioengineered* 3(6):347-351

194 Fernandes S, Murray P, 2010. Metabolic engineering for improved microbial pentose fermentation. *Bioeng Bugs* 1:6:424-428

195 Kim JH, Block DE, Mills DA, 2010. Simultaneous consumption of pentose and hexose sugars: an optimal microbial phenotype for efficient fermentation of lignocellulosic biomass. *Appl Microbiol Biotechnol* 88:1077–1085

196 Wen F, Sun J, Zhao H, 2010. Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol. *Appl Environmental microbial* 76(4):1251–1260

197 Ilmén M, den Haan R, Brevnova E, McBride J, Wiswall E, Froehlich A, Koivula A, Voutilainen SP, *et al.*, 2011. High level secretion of cellobiohydrolases by Saccharomyces cerevisiae. *Biotechnol Biofuels* 4:30

198 Fan L-H, Zhang Z-J, Yu X-Y, Xue Y-X, Tan T-W, 2012. Self-surface assembly of cellulosomes with two miniscaffoldins on Saccharomyces cerevisiae for cellulosic ethanol production. *PNAS* 109(33):13260–13265

199 Hasunuma T, Kondo A, 2012. Consolidated bioprocessing andsimultaneous saccharification andfermentation of lignocellulose to ethanolwith thermotolerant yeast strains. *Process Biochem* 47(9): 1287-1294

200 Park EY, Naruse K, Kato T, 2012. One-pot bioethanol production from cellulose by co-culture of Acremonium cellulolyticus and Saccharomyces cerevisiae. Biotechnol *Biofuels* 5:64

201 Zuroff TR, Xiques SB, Curtis WR, 2013. Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium phytofermentans*/yeast co-culture. *Biotechnol Biofuels* 6:59

202 Margeot A, Hahn-Hagerdal B, Edlund M, Slade R, Monot F, 2009. New improvements for lignocellulosic ethanol. *Curr Opin Biotechnol* 20:1–9

203 Hendriks ATWM, Zeeman G, 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource Technology* 100:10–18

204 Wyman CE, Dale BE, Elander RT, *et al.,* 2009. Comparative sugar recovery and fermentation data following pretreatment of poplar wood by leading technologies. *Biotechnol Prog* 25:333–339

205 Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW, 2012. The Challenge of Enzyme Cost in the Production of Lignocellulosic Biofuels. *Biotechnol Bioeng* 109(4): 1-5

206 Merino S, Cherry J, 2007. Progress and challenges in enzyme development for biomass utilization. *Biofuels* 108:95-120

207 Simmons BA, Loqué D, Ralph J, 2010. Advances in modifying lignin for enhanced biofuel production. *Current Opin Plant Biol* 13: 312–319

208 US DOE Genomic Science Program, 2006. Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda. DOE/SC/EE-0095. Available at: http://genomicscience.energy.gov/biofuels (Last accessed January 2014)

209 Wen F, Nair NU, Zhao H, 2009. Protein engineering in designing tailored enzymes and microorganisms for biofuels production. *Curr Opin Biotechnol* 20:412–419

210 Ambavaram MMR, Krishnan A, Trijatmiko KR, Pereira A, 2011. Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. *Plant Physiol* 155:916–931,

211 Li X, Weng JK, Chapple C, 2008. Improvement of biomass through lignin modification. *Plant J* 54: 569–581

212 Jakob K, Zhou FS, Paterson A. 2009, Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. *In Vitro Cell Dev Biol Plant* 45: 291–305

213 Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW, 2010. Technoeconomic analysis of biofuels: a wiki-based platform for lignocellulosic biorefineries. *Biomass Bioenergy* 34:1914–1921

214 Blanch HW, Simmons BA, Klein-Marcuschamer D, 2011. Biomass deconstruction to sugars. *Biotechnol J* 6:1086–1102

215 Yang F, Mitra P, Zhang L, Prak L, Verhertbruggen Y, Kim J-S, Sun L, Zheng K, Tang K, Auer M, Scheller HV, Loque D, 2013. Engineering secondary cell wall deposition in plants. *Plant Biotechnol J* 11:325–335

216 Chou H, Dai Z, Hsieh C, Ku M, 2011.High level expression of Acidothermus cellulolyticus beta-1, 4-endoglucanase in transgenic rice enhances the hydrolysis of its straw by cultured cow gastric fluid. *Biotechnol Biofuels* 4:58

217 Klose H, Roder J, Girfoglio M, Fischer R, Commandeur U, 2012. Hyperthermophilic endoglucanase for in planta lignocellulose conversion. *Biotechnol Biofuels* 5:63

218 Brunecky R, Selig M, Vinzant T, Himmel M, Lee D, Blaylock M, Decker S, 2011. In planta expression of A. cellulolyticus Cel5A endocellulase reduces cell wall recalcitrance in tobacco and maize. *Biotechnol Biofuels* 4:1

219 Dai Z, Hooker B, Anderson DB, Quesenberry R, Thomas S, 2000. Expression of Acidothermus cellulolyticus endoglucanase E1 in transgenic tobacco: biochemical characteristics and physiological effects. *Transgenic Res* 9:43-45

220 Li X, Weng JK, Chapple C, 2008. Improvement of biomass through lignin modification. *Plant J* 54: 569–581

221 Taylor LE 2nd, Dai Z, Decker SR, Brunecky R, Adney WS, Ding SY, Himmel ME, 2008. Heterologous expression of glycosyl hydrolases in planta: a new departure for biofuels**.** *Trends Biotechnol* 26**:**413-424

222 Petersen K, Bock R, 2011. High-level expression of a suite of thermostable cell wall-degrading enzymes from the chloroplast genome. *Plant Mol Biol* 76(3-5):311-321

223 Himmel ME, Ding S-Y, Johnson DK, et al., 2007. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* 315:804

224 Voutilainen SP, Murray PG, Tuohy MG, *et al.,* 2010. Expression of *Talaromyces emersonii* cellobiohydrolase Cel7A in Saccharomyces cerevisiae and rational mutagenesis to improve its thermostability and activity. *Protein Eng Des Sel* 23(2):69-79

225 Lee CY, Yu KO, Kim SW, *et al.,* 2010. Enhancement of the thermostability and activity of mesophilic *Clostridium cellulovorans* EngD by in vitro DNA recombination with *Clostridium thermocellum* CelE. *J Biosci Bioeng* 109(4):331-336

226 Kim DM, Umetsu M, Takai K, *et al.,* 2011. Enhancement of cellulolytic enzyme activity by clustering cellulose binding domains on nanoscaffolds. *Small* 7(5):656-664

227 Seiboth B, Ivanova C, Seidl-Seiboth V, 2011. *Trichoderma reesei*: A fungal enzyme producer for cellulosic biofuels, biofuel production-recent developments and prospects, Ed; Dr. Marco Aurelio Dos Santos Bernardes, ISBN: 978-953-307-478-8, InTech

228 Reese ET, 1976. History of the cellulase program at the U.S. Natick Development Center. *Biotechnol Bioeng Symposium* 6:9-20

229 Bailey MJ, Nevalainene KMH, 1981. Induction, isolation and testing of stable *Trichoderma reesei* mutants with improved production of solubilizing cellulose. *Enzyme Microb Technol* 3: 153–157

230 Cherry JR, Fidantsef AL, 2003. Directed evolution of industrial enzymes: an update. *Curr Opin Biotech* 14:438–443

231 Mendels M, Weber J, Parizek R, 1971. Enhanced cellulase production by a mutant of *Trichoderma viride*. *Appl Microbiol* 21:152–154

232 Jun H, Bing Y, Keying Z, Xuemei D, Daiwen C, 2009. Strain improvement of *Trichoderma reesei* Rut C-30 for increased cellulase production. *Indian J Microbiol* 49:188–195

233 Portnoy T, Margeot A, Seidl-Seiboth V, Le Crom S, Ben Chaabane F, Linke R, Seiboth B, Kubicek CP, 2011. Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. *Eukaryotic Cell* 10(2): 262-271

234 Cherry JR, Fidantsef AL, 2003. Directed evolution of industrial enzymes: an update. *Curr Opinion Biotechnol* 14(4):438-443

235 Aro N, Pakula T, Penttilä M, 2005. Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS Microbiol Reviews* 29(4):719-739

236 Kubicek CP, Mikus M, Schuster A, Schmoll M, Seiboth B, 2009. Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. *Biotechnol Biofuels* 2:19

237 Carroll A, Somerville C, 2009. Cellulosic biofuels. *Annual Rev Plant Biol* 60:165-182

238 Wilson DB, 2009. Cellulases and biofuels. *Curr Opinion Biotechnol* 20(3):295-299

239 Banerjee G, Scott-Craig JS, Walton JD, 2010. Improving enzymes for biomass conversion: a basic research perspective. *Bioenerg Res* 3:82-92

240 Lynd LR, Laser MS, Bransby D, Dale BE, Davison B, Hamilton R, Himmel M, Keller M, McMillan JD, Sheehan J, Wyman CE, 2008. How biotech can transform biofuels. *Nature Biotechnol* 26:169-172

241 Rosgaard L, Pedersen S, Cherry JR, Harris P, Meyer AS, 2006. Efficiency of new fungal cellulase systems in boosting enzymatic degradation of barley straw lignocellulose. *Biotechnol Prog* 22:493 – 498

242 Tengborg C, Galbe M, Zacchi G, 2001. Influence of enzyme loading and physical parameters on the enzymatic hydrolysis of steam-pretreated softwood. *Biotechnol Prog* 17:110 – 117

243 Jørgensen H, Vibe-Pedersen J, Larsen J, Felby C, 2007. Liquefaction of lignocellulose at high-solids concentrations. *Biotechnol Bioeng* 96(5):862-870

244 Larsen J, Petersen MØ, Thirup L, Li HW, Iversen FK, 2008. The IBUS Process – Lignocellulosic bioethanol close to a commercial reality. *Chem Eng Technol* 31(5): 765–772

245 Wingren A, Galbe M, Zacchi G, 2003. Techno-economic evaluation of producing ethanol from softwood: comparison of SSF and SHF and identification of bottlenecks. *Biotechnol Prog* 19(4):1109-17

246 Viikari L, Alapuranen M, Puranen T, Vehmaanperä J, Siika-aho M, 2007. Thermostable Enzymes in Lignocellulose Hydrolysis. *Adv Biochem Engin Biotechnol* 108: 121–145

247 Turner P, Mamo G, Karlsson EN, 2007. Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb Cell Factories* 6:9

248 Heinzelman P, Snow CD, Smith MA, Yu X, Kannan A, Boulware K, Villalobos A, Govindarajan S, Minshull J, Arnold FH, 2009. SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. *J Biol Chem* 284(39):26229–26233

249 Murashima K, Kosugi A, Doi RH, 2002. Thermostabilization of cellulosomal endoglucanase EngB from *Clostridium cellulovorans* by in vitro DNA recombination with non-cellulosomal endoglucanase *EngD. Mol Microbiol* 45(3):617–626

250 van den Brink J, de Vries RP, 2011. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol* 91:1477-1492

251 Banerjee G, Scott-Craig JS, Walton JD, 2010. Improving enzymes for biomass conversion: A basic research perspective. *Bioenerg Res* 3:82–92

252 Koike H, Aerts A, La Butti K, GrigorievIgor V, Baker SE, 2013. Comparative genomics analysis of Trichoderma reesei strains. *Ind Biotechnol* 9(6): 352-367

253 Alper H, Stephanopoulous G, 2009. Engineering for Biofuel: Exploiting innate microbial capacity or importing biosynthetic potential? *Nat Rev Microbiol* 7:715–723

254 Barnard, Casanueva A, Tuffin M, Cowan D, 2010. Extremophiles in biofuel synthesis. *Environ Technol* 31(8–9): 871–888

255 van den Brink J, van Muiswinkel GCJ, Theelen B, Hinz SWA, de Vries RP, 2013. Efficient plant biomass degradation by thermophilic fungus *Myceliophthora heterothallica. Applied Env Microbiol* 79(4):1316–1324

256 Roy SK, Dey SK, Raha SK, Chakrabarty SL, 1990. Purification and properties of an extracellular endoglucanase from Myceliophthora thermophilavD-14 (ATCC 48104). *J Gen Microbiol* 136:1967–1971

257 Maijala P, Kango N, Szijarto N, Viikari L, 2012. Characterization of hemicellulases from thermophilic fungi. *Antonie Van Leeuwenhoek* 101:905–917

258 van den Brink J, Samson RA, Hagen F, Boekhout T, de Vries RP, 2012. Phylogeny of the industrial relevant, thermophilic genera *Myceliophthora* and *Corynascus*. *Fungal Diversity* 52:10.

259 Karnaouri AC, Topakas E, Christakopoulos P, 2013. Cloning, expression and characterization of a thermostable GH7 endoglucanase from *Myceliophthora thermophile* capable of high-consistency enzymatic liquefaction. *Appl Microbiol Biotechnol* DOI 10.1007/s00253-013-4895-9

260 King BC, Waxman KD, Nenni NV, Walker LP, Bergstrom GC, Gibson DM, 2011. Arsenal of plant cell wall degrading enzymes reflects host preference among plant pathogenic fungi. *Biotechnol Biofuels* 4:4

261 Voutilainen SP, Puranen T, Siika-aho M, Lappalainen A, Alapuranen M, Kallio J, Hooman S, Viikri L, Vehmaanpera J, Koivula A, 2008. Cloning, expression and characterization of novel thermostable family 7 cellobiohydrolases. *Biotechnol Bioeng* 101(3): 515-528

262 Li YL, Li H, Li AN, Li DC, 2009. Cloning of a gene encoding thermostable cellobiohydrolase from the thermophilic fungus *Chaetomium thermophilum* and its expression in *Pichia pastoris. J Appl Microbiol* 106:1867–1875

263 Meyer AS, Rosgaard L, Sorensen HR, 2009. The minimal enzyme cocktail concept for biomass processing. *J Cereal Science* 50:337-344

264 Gibson DM, King BC, Hayes ML, Bergstrom GC, 2011. Plant pathogens as a source of diverse enzymes for lignocellulose digestion. *Curr Opinion Microbiol* 14:264–270

265 Hong JTH, Yamamoto K, Kumagai H, 2003. Cloning of a gene encoding a thermostable endo-β-1,4-glucasnase from *Thermoascus aurantiacus* and its expression in yeast. *Biotechnol Lett* 25:657–661

266 Murray PG, Aro N, Collins CM, Grassick A, Penttila M, Saloheimo M, Tuohy MG, 2004. Expression in *Trichoderma reesei* and characterisation of a thermostable family 3-β-glucosidase from the moderately thermophilic fungus *Talaromyces emersonii*. *Protein Expression Purif* 38:248–257

267 Szijártó N, Horan E, Zhang J, Puranen T, Siika-aho M, Viikari L, 2011. Thermostable endoglucanases in the liquefaction of hydrothermally pretreated wheat straw. *Biotechnol Biofuels* 4:2

268 Kumar R, Wyman CE, 2008. An improved method to directly estimate cellulase adsorption on biomass solids *Enzyme Microbial Technol* 42(5): 426-433

269 Riou C, Salmon J-M, Vallier M-J, Nata ZG, Barre P, 1998. Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from Aspergillus oryzae. *Appl Environ Microbiol* 64(10):3607–3614

270 Barnard, Casanueva A, Tuffin M, Cowan D, 2010. Extremophiles in biofuel synthesis. *Environ Technol* 31(8–9): 871–888

271 Gao D, Uppugundla N, Chundawat SPS, Yu X, Hermanson S, Gowda K, Brumm P, Mead D, Balan V, Dale BE, 2011. Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnol Biofuels* 4:5

272 Scott BR, Hill C, Tomashek J, Liu C, 2009. Enzymatic hydrolysis of lignocellulosic feedstocks using accessory enzymes. *US Patent Application* 2009/0061484

273 Sweeney MD, Xu F, 2012. Biomass converting enzymes as industrial biocatalysts for fuels and chemicals: Recent developments. *Catalysts* 2:244–263

274 Morais S, Barak Y, Caspi J, Hadar Y, Lamed R, Shoham Y, Wilson DB, Bayer EA, 2010. Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *mBio* 2010, 1:e00285-00210

275 16. Kumar R, Singh S, Singh OV, 2008. Bioconversion of lignocellulosic biomass: Biochemical and molecular perspective. *J Ind Microbiol Biotechnol* 35:377–391

276 Sweeney MD, Xu F, 2012. Biomass converting enzymes as industrial biocatalysts for fuels and chemicals: Recent developments. *Catalysts* 2:244–263

277 Wilson DB, 2009. Cellulases and biofuels. *Curr Opin Biotechnol* 20:295–299

278 Himmel M, Xu Q, Luo Y, Ding S, Lamed R, Bayer E, 2010. Microbial enzyme systems for biomass conversion: Emerging paradigms. *Biofuels* 1:323–341

279 Gilbert HJ, 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. *Plant Physiology* 153(2):444-455

280 Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B, 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* 37(suppl 1):D233-D238

281 ovanovic I, Magnuson J, Collart F, Robbertse B, Adney W, Himmel M, Baker S, 2009. Fungal glycoside hydrolases for saccharification of lignocellulose: outlook for new discoveries fueled by genomics and functional studies. *Cellulose* 16(4):687-697

282 Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P *et al*., 2009. Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci USA* 106(6):1954-1959

283 Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJ, Culley D, Thykaer J, Frisvad JC, Nielsen KF, Albang R, Albermann K *et al*., 2011. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome research* 21(6):885-897

284 Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS *et al*, 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715-1719

285 Flipphi MJ, van Heuvel M, van der Veen P, Visser J, de Graaff LH, 1993. Cloning and characterization of the abfB gene coding for the major alpha-L-arabinofuranosidase (ABF B) of *Aspergillus niger*. *Curr Genet* 24(6):525-532

286 Ishikawa E, Sakai T, Ikemura H, Matsumoto K, Abe H, 2005. Identification, cloning, and characterization of a Sporobolomyces singularis beta-galactosidase-like enzyme involved in galacto-oligosaccharide production. *J Biosci Bioeng* 99(4):331-339

287 Mertens JA, Burdick RC, Rooney AP, 2008. Identification, biochemical characterization, and evolution of the Rhizopus oryzae 99-880 polygalacturonase gene family. *Fungal Genet Biol* 45(12):1616-1624

288 Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR, 2006. Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc Natl Acad Sci USA* 103(30):11417-11422

289 Martens-Uzunova ES, Zandleven JS, Benen JA, Awad H, Kools HJ, Beldman G, Voragen AG, Van den Berg JA, Schaap PJ, 2006. A new group of exo-acting family 28 glycoside hydrolases of Aspergillus niger that are involved in pectin degradation. *Biochem J* 400(1):43-52

290 Murphy C, Powlowski J, Wu M, Butler G, Tsang A, 2011. Curation of characterized glycoside hydrolases of fungal origin. *Database (Oxford)* 2011:bar020

291 Cui Z, Mochizuki D, Matsuno Y, Nakamura T, Liu Y, Hatano T, Fukui S, Miyakawa T, 1992. Cloning and molecular analysis of cDNA encoding a carboxymethylcellulase of the yeast *Cryptococcus flavus*. *Biosci Biotechnol Biochem* 56(8):1230-1235

292 Zimmermann CR, Johnson SM, Martens GW, White AG, Pappagianis D, 1996. Cloning and expression of the complement fixation antigen-chitinase of Coccidioides immitis. *Infect Immun* 64(12):4967-4975

293 Margolles-Clark E, Tenkanen M, Nakari-Setala T, Penttila M, 1996. Cloning of genes encoding alpha-L-arabinofuranosidase and beta-xylosidase from Trichoderma reesei by expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 62(10):3840-3846

294 Saloheimo A, Henrissat B, Hoffren AM, Teleman O, Penttila M, 1994. A novel, small endoglucanase gene, egl5, from *Trichoderma reesei* isolated by expression in yeast. *Mol Microbiol* 13(2):219-228

295 Zhang ZG, Yi ZL, Pei XQ, Wu ZL, 2010. Improving the thermostability of *Geobacillus stearothermophilus* xylanase XT6 by directed evolution and site-directed mutagenesis. *Bioresour Technol* 101(23):9272-9278

296 Tung JY, Chang MD, Chou WI, Liu YY, Yeh YH, Chang FY, Lin SC, Qiu ZL, Sun YJ, 2008. Crystal structures of the starch-binding domain from *Rhizopus oryzae* glucoamylase reveal a polysaccharide-binding path. *Biochem J* 416(1):27-36

297 Nakazawa H, Okada K, Onodera T, Ogasawara W, Okada H, Morikawa Y, 2009. Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. *Appl Microbiol Biotechnol* 83(4):649-657

298 van Zyl WH, Lynd LR, den Haan R, McBride JE, 2007. Consolidated bioprocessing for bioethanol production using Saccharomyces cerevisiae. *Adv Biochem Eng Biotechnol* 108:205-235

299 Wilde C, Gold ND, Bawa N, Tambor JH, Mougharbel L, Storms R, Martin VJ, 2012. Expression of a library of fungal beta-glucosidases in *Saccharomyces cerevisiae* for the development of a biomass fermenting strain. *Appl Microbiol Biotechnol* 95(3):647-659

300 Ward M, Wilson LJ, Kodama KH, Rey MW, Berka RM, 1990. Improved production of chymosin in Aspergillus by expression as a glucoamylase-chymosin fusion. *Biotechnology (NY)* 8(5):435-440

301 van Dijck PW, Selten GC, Hempenius RA, 2003. On the safety of a new generation of DSM Aspergillus niger enzyme production strains. *Regul Toxicol Pharmacol* 38(1):27-35

302 Schuster E, Dunn-Coleman N, Frisvad JC, Van Dijck PW, 2002. On the safety of *Aspergillus niger*--a review. *Appl Microbiol Biotechnol* 59(4-5):426-435

303 Nielsen BR, Lehmbeck J, Frandsen TP, 2002. Cloning, heterologous expression, and enzymatic characterization of a thermostable glucoamylase from *Talaromyces emersonii*. *Protein Expr Purif* 26(1):1-8

304 Bussink HJ, Brouwer KB, de Graaff LH, Kester HC, Visser J, 1991. Identification and characterization of a second polygalacturonase gene of *Aspergillus niger*. *Curr Genet* 20(4):301-307

305 Ademark P, de Vries RP, Hagglund P, Stalbrand H, Visser J, 2001. Cloning and characterization of *Aspergillus niger* genes encoding an alpha-galactosidase and a beta-mannosidase involved in galactomannan degradation. *Eur J Biochem* 268(10):2982-2990

306 Master ER, Zheng Y, Storms R, Tsang A, Powlowski J, 2008. A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization. *Biochem J* 411(1):161-170

307 Kim JY, Kavas M, Fouad WM, Nong G, Preston JF, Altpeter F, 2011. Production of hyperthermostable GH10 xylanase Xyl10B from *Thermotoga maritima* in transplastomic plants enables complete hydrolysis of methylglucuronoxylan to fermentable sugars for biofuel production. *Plant Mol Biol* 76(3-5):357-69

308 Verma D, Kanagaraj A, Jin S, Singh ND, Kolattukudy PE, Daniell H, 2010. Chloroplast-derived enzyme cocktails hydrolyse lignocellulosic biomass and release fermentable sugars. *Plant Biotechnol J* 8(3):332-350

309 Kolotilin I, Kaldis A, Pereira EO, Laberge S, Menassa R, 2013. Optimization of transplastomic production of hemicellulases in tobacco: effects of expression cassette configuration and tobacco cultivar used as production platform on recombinant protein yields. *Biotechnol biofuels* 6(1):65

310 Goosen C, Yuan XL, van Munster JM, Ram AF, van der Maarel MJ, Dijkhuizen L, 2007. Molecular and biochemical characterization of a novel intracellular invertase from *Aspergillus nig*er with transfructosylating activity. *Eukaryot Cell* 6(4):674-681

311 Yuan XL, Goosen C, Kools H, van der Maarel MJ, van den Hondel CA, Dijkhuizen L, Ram AF, 2006. Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. *Microbiol* 152(10):3061-3073

312 Tang CM, Waterman LD, Smith MH, Thurston CF, 2001. The cel4 gene of *Agaricus bisporus* encodes a beta-mannanase. *Appl Environ Microbiol* 67(5):2298-2303

313 Skory CD, Freer SN, Bothast RJ, 1996. Expression and secretion of the *Candida wickerhamii* extracellular beta-glucosidase gene, bglB, in *Saccharomyces cerevisiae*. *Curr Genet* 30(5):417-422

314 Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, Darvill A, 1999. A xyloglucan-specific endo-beta-1,4-glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification and characterization of the recombinant enzyme. *Glycobiology* 9(1):93-100

315 Skjot M, Kauppinen S, Kofod LV, Fuglsang C, Pauly M, Dalboge H, Andersen LN, 2001. Functional cloning of an endo-arabinanase from *Aspergillus aculeatus* and its heterologous expression in A. or oryzae and tobacco. *Mol Genet Genomics* 265(5):913-921

316 Roca H, Garcia B, Rodriguez E, Mateu D, Coroas L, Cremata J, Garcia R, Pons T, Delgado J, 1996. Cloning of the *Penicillium minioluteum* gene encoding dextranase and its expression in *Pichia pastoris*. *Yeast* 12(12):1187-1200

317 Duruksu G, Ozturk B, Biely P, Bakir U, Ogel ZB, 2009. Cloning, expression and characterization of endo-beta-1,4-mannanase from *Aspergillus fumigatus* in *Aspergillus sojae* and *Pichia pastoris*. *Biotechnol Prog* 25(1):271-276

318 Wang L, Aryal UK, Dai Z, Mason AC, Monroe ME, Tian ZX, Zhou JY, Su D, Weitz KK, Liu T *et al*., 2012. Mapping N-linked glycosylation sites in the secretome and whole cells of *Aspergillus niger* using hydrazide chemistry and mass spectrometry. *JProteome Research* 11(1):143-156

319 Liu Y, Yoshida M, Kurakata Y, Miyazaki T, Igarashi K, Samejima M, Fukuda K, Nishikawa A, Tonozuka T, 2010. Crystal structure of a glycoside hydrolase family 6 enzyme, CcCel6C, a cellulase constitutively produced by Coprinopsis cinerea. *FEBS J* 277(6):1532-1542

320 Carapito R, Imberty A, Jeltsch JM, Byrns SC, Tam PH, Lowary TL, Varrot A, Phalip V, 2009. Molecular basis of arabinobio-hydrolase activity in phytopathogenic fungi: crystal

structure and catalytic mechanism of *Fusarium graminearum* GH93 exo-alpha-L-arabinanase. *J Biological Chem* 284(18):12285-12296

321 Goncalves AM, Pedro AQ, Maia C, Sousa F, Queiroz JA, Passarinha LA, 2013. Pichia pastoris: A Recombinant Microfactory for Antibodies and Human Membrane Proteins. *J Microbiol Biotechnol* 23(5):587-601

322 Poidevin L, Feliu J, Doan A, Berrin JG, Bey M, Coutinho PM, Henrissat B, Record E, Heiss-Blanquet S, 2013. Insights into exo- and endoglucanase activities of family 6 glycoside hydrolases from *Podospora anserina*. *Appl Environ Microbiol* 79(14):4220-4229

323 Zeitler B, Bernhard A, Meyer H, Sattler M, Koop HU, Lindermayr C, 2013. Production of a de-novo designed antimicrobial peptide in *Nicotiana benthamiana*. *Plant Mol Biol* 81(3):259-272

324 Fernandez-del-Carmen A, Juarez P, Presa S, Granell A, Orzaez D, 2013. Recombinant jacalin-like plant lectins are produced at high levels in *Nicotiana benthamiana* and retain agglutination activity and sugar specificity. *J biotechnol* 163(4):391-400

325 Beckham GT, Dai Z, Matthews JF, Momany M, Payne CM, Adney WS, Baker SE, Himmel ME, 2012. Harnessing glycosylation to improve cellulase activity. *Curr opin biotechnol* 23(3):338-345

326 Deng P, Li D, Cao Y, Lu W, Wang C, 2006. Cloning of a gene encoding an acidophilic endo-[beta]-1,4-xylanase obtained from *Aspergillus niger* CGMCC1067 and constitutive expression in *Pichia pastoris*. *Enzyme Microb Technol* 39(5):1096-1102

327 Ito K, Ikemasu T, Ishikawa T, 1992. Cloning and Sequencing of the Xyna-Gene Encoding Xylanase-a of *Aspergillus-Kawachii*. *Biosci Biotech Bioch* 56(6):906-912

328 Structural Genomics C, China Structural Genomics C, Northeast Structural Genomics C, Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S *et al*., 2008. Protein production and purification. *Nature methods* 5(2):135-146

329 Emanuelsson O, Brunak S, von Heijne G, Nielsen H, 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2:953-971

330 Davies WL, Carvalho LS, Hunt DM, 2007. SPLICE: A technique for generating in vitro spliced coding sequences from genomic DNA. *Bio Techniques* 43:785-789

331 Sambrook J, Russell DW, 2001. Molecular cloning : a laboratory manual, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

332 Yoon JR, Laible PD, Gu M, Scott HN, Collart FR, 2002. Express primer tool for high-throughput gene cloning and expression. *Biotechniques* 33:1328-1333

333 Zerbs S, Frank AM, Collart FR, 2009. Bacterial systems for production of heterologous proteins. *Methods Enzymol* 463:149-168

334 Wu S, Letchworth GJ, 2004. High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *BioTechniques* 36:152-154

335 Storms R, Zheng Y, Li H, Sillaots S, Martinez-Perez A, Tsang A, 2005. Plasmid vectors for protein production, gene expression and molecular manipulations in *Aspergillus niger*. *Plasmid* 53(3):191-204

336 Debets AJM, Bos CJ, 1986. Isolation of small protoplasts from *Aspergillus niger*. *Fungal Genet Newsl* 33:44

337 Kafer E, 1977. Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. *Adv Genet* 19:33-131

338 Harris LJ, Gleddie SC, 2001. A modified Rpl3 gene from rice confers tolerance of the Fusarium graminearum mycotoxin deoxynivalenol to transgenic tobacco. *Physiol Mol Plant Pathol* 58(4):173-181

339 Wu K, Malik K, Tian L, Hu M, Martin T, Foster E, Brown D, Miki B, 2001. Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, tCUP. *MGG Mol Genetics Genomics* 265(5):763-770

340 Matsuoka M, Yamamoto N, Kano-Murakami Y, Tanaka Y, Ozeki Y, Hirano H, Kagawa H, Oshima M, Ohashi Y, 1987. Classification and structural comparison of full-length complementary dna for pathogenesis-related proteins. *Plant Physiol (Rockville)* 85(4):942-946

341 Hood EE, Gelvin SB, Melchers LS, Hoekema A, 1993. New Agrobacterium helper plasmids for gene transfer to plants. *Transgenic Research* 2(4):208-218

342 Marillonnet S, Thoeringer C, Kandzia R, Klimyuk V, Gleba Y, 2005. Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nature Biotechnol* 23(6):718-723

343 Silhavy D, Molnar A, Lucioli A, Szittya G, Hornyik C, Tavazza M, Burgyan J, 2002. A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *Eur Mol Biol Organ J* 21(12):3070-3080

344 Bradford MM, 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254

345 Doner LW, Irwin PL, 1992. Assay of reducing end-groups in oligosaccharide homologues with 2,2'-bicinchoninate. *Anal Biochem* 202(1):50-53

346 Grishutin SG, Gusakov AV, Markov AV, Ustinov BB, M.V. S, Sinitsyn AP, 2004. Specific xyloglucanases as a new class of polysaccharide-degrading enzymes. *Biochim Biophys Acta* 1674:268-281

347 Zorov IN, Dubasova MY, Sinitsyn AP, Gusakov AV, Mitchenko AA, Baraznenok VA, Gutyerres B, Popova NN, 1997. Application of the bicinchoninic method of the assay for the reducing sugars to determine carboxymethylcellulase activity of cellulases using a microplate reader. *Biochemistry (Mosc)* 62:704–709

348 Jørgensen H, Eriksson T, Börjesson J, Tjerneld F, Olsson L, 2003. Purification and characterization of five cellulases and one xylanase from *Penicillium brasilianum* IBT 20888. *Enzyme Microb Technol* 32:851-861

349 Ausubel F, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, 1999. Current Protocols in Molecular Biology. New York: John Wiley & Sons

350 Krieg RC, Dong Y, Schwamborn K, Knuechel R, 2005. Protein quantification and its tolerance for different interfering reagents using the BCA-method with regard to 2D SDS PAGE. *J Biochem Biophys Methods* 65(1):13-19

351 REN21, 2010. *Renewables*: Global Status Report

352 Szklo A, Schaeffer R, Delgado F, 2007. Can One Say Ethanol Is a Real Threat to Gasoline? *Energy Policy,* 35:5411-542

353 Davies G, Henrissat B, 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* 3:85.3-859

354 Cos O, Ramón R, Montesinos JL, Valero F, 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. *Microb Cell Fact*; 5:1-20

355 Zhang X, Ekwe E, Conley AJ, Kaldis A, Zerbs S, Collart F, Magnuson J, Baker SE, Menassa R, Powlowski J, Tsang A, 2013. A comprehensive comparison of fungal glycoside hydrolases produced in bacterial, yeast, filamentous fungal and plant expression hosts. *Microb Cell Fact* In Press

356 Center for Biological Sequence Analysis, Technical University of Denmark (DTU) [http://www.cbs.dtu.dk/services/NetNGlyc/]

357 Center for Biological Sequence Analysis, Technical University of Denmark (DTU) [http://www.cbs.dtu.dk/services/NetOGlyc/]

358 Roth J, 2002. Protein N-glycosylationalong the secretory pathway: relationship to organelle topography and function, protein quality and cell interaction. *Chem Rev 1*02:285-303

359 Walton JD, 1994. Deconstructing the cell wall. *Plant Physiol* 104: 1113-1118

360 Kawai T, Nakazawa H, Ida N, *et al.* Analysis of the saccharification capability of high-functional cellulase JN11 for various pretreated biomasses through a comparison with commercially available counterparts. *J Ind Microbiol Biotechnol* 2012;39(12): 1741-1749.

361 Nieves RA, Ehrman CI, Adney WS, *et al*. Survey and analysis of commercial cellulase preparations suitable for biomass conversion to ethanol. *World J Microbiol Biotechnol* 1997;14(2):301-304

362 Vinzant TB, Adney WS, Decker SR, *et al.* Fingerprinting *Trichoderma reesei* hydrolases in a commercial cellulase preparation. *Appl Biochem Biotechnol* 2001;91-93:99-107

363 Duan C-J, Feng J-X, 2010. Mining metagenomes for novel cellulase genes. *Biotechnol Lett* 32:1765–1775

364 Niehaus F, Bertoldo C, KaÈhler M, Antranikian G, 1999. Extremophiles as a source of novel enzymes for industrial application. *Appl Microbiol Biotechnol* 51(6): 711-729

365 Mewis K, Taupp M, Hallam SJ, 2011. A high throughput screen for biomining cellulase activity from metagenomic libraries. JoVE. 48. http://www.jove.com/details.php?id=2461, doi: 10.3791/2461

366 King BC, Donnelly MK, Bergstrom GC, Walker LP, Gibson DM, 2009. An optimized microplate assay system for quantitative evaluation of plant cell wall-degrading enzyme activity of fungal culture extracts. *Biotechnol. Bioeng.* 102(4) : 1033-1044

367 Studer MH, DeMartini JD, Brethauer S, McKenzie HL, 2010. Engineering of a high-throughput screening system to identify cellulosic biomass, pretreatments and enzyme formulations that enhance sugar release, *Biotechnol. Bioeng.* 105(2) : 231-238

368 Mohagheghi A, Tucker M, Grohman K, Wyman CE, 1992. High solid simultaneous saccharification and fermentation of pretreated wheat straw to ethanol. *Appl. Biochem. Biotechnol*. 33:67-81.

369 Stenberg K, Bollok M, Reczey K, Galbe M, Zacchi G, 2000. Effect of substrate and cellulase concentration on simultaneous saccharification and fermentation of steampretreated softwood for ethanol production. *Biotechnol Bioeng.* 68(2):204-210

370 Kristensen JB, Felby C, Jørgensen H, 2009. Yield-determining factors in high-solids enzymatic hydrolysis of lignocellulose*. Biotechnol Biofuels* 2:11 doi: 10.1186/1754-6834-2-11

371 Palonen H, Tjerneld F, Zacchi G, Tenkanen M, 2004. Adsorption of *Trichoderma reesei* CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *J Biotechnol* 107: 65-72

372 Spindler DD, Wyman CE Mohagheghi A, Grohman K, 1988. Thermotolerant yeast for simultaneous saccharification and fermentation of cellulose to ethanol. *Appl Biochem Biotechnol* 17(1-3):279–293

373 Peterson GL, 1977. A simplification of the protein assay method of Lowry et al. which is more generally applicable. *Anal Biochem* 83(2): 346-356

374 Xiao Z, Storms R, Tsang A, 2004. Microplated-based filter paper assay to measure total cellulase activity. *Biotechnol Bioeng* 88: 832-7

375 Decker SR, Adney WS, Jennings E, Vinzant TB and Himmel ME, 2003. Automated Filter Paper Assay for Determination of Cellulase Activity. *Appl Biochem Biotechnol* (105–108):689-703

376 Wood TM and Bhat KM, 1988. Methods for measuring cellulase activities. *Methods Enzymol* 160:87-112

377 Zorov IN, et al., 1997. Application of the bicinchoninic method of the assay for reducing sugars to determine carboxymethylcellulase activity of cellulases using a microplate reader. *Biochem (Mosc)* **62**:704-709

378 Gusakov AV, Sinitsyn AP, Salanovich TN, Bukhtojarov FE, Markov AV, Ustinov BB, van Zeijl C, Punt P, Burlingame R, 2005. Purification, cloning and characterisation of two forms of thermostable and highly active cellobiohydrolase (Cel7A) produced by the industrial strain of *Chrysosporium lucknowense*. *Enzyme Microb Technol* 36: 57-69

379 National Renewable Energy Laboratory, US DOE, 2008. Determination of Structural Carbohydrates and Lignin in Biomass. Laboratory Analytical Procedure (LAP); *Technical Report*, *NREL/TP-510-42618*

380 National Renewable Energy Laboratory, US DOE, 2008. Enzymatic saccharification of lignocellulosic biomass. Laboratory Analytical Procedure (LAP); *Technical Report NREL/TP-510-42629*. Ed; Selig M, Weiss N, Ji Y.

381 Moerman R, van den Doel LR, Picioreanu S, Frank J, Marijnissen JCM, van Dedem G, Hjelt KH, Vellekoop MJ, Sarro PM, Young IT, 1999. Micro-injections of *β*-D-glucose standards and Amplex Red reagent on Micro-arrays. *Proc SPIE-Intl Soc Opt Eng* 3606:119-119

382 Zhou M, Diwu Z, Panchuk-Voloshina N, Haugland RP, 1997. A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Anal Biochem* 253: 162-168

383 Mohanty JG, Jaffe JS, Schulman ES, Raible DG, 1997. A highly sensitive fluorescent micro-assay of H2O2 release from activated human leukocytes using a dihydroxyphenoxazine derivative. *J Immunol Methods* 202: 133-41

384 National Renewable Energy Laboratory , US DOE, 2008. SSF Experimental Protocols — Lignocellulosic Biomass Hydrolysis and Fermentation. Laboratory Analytical Procedure (LAP); Technical *Report NREL/TP-510-42630.* Ed; Dowe N, McMillan J.

385 Arantes V, Saddler JN, 2011. Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnol Biofuels* 4:3

386 McGovem JN and Wegner TH, 1991. Papermaking properties of aspen ultrahigh-yield mechanical pulp. *Tappi J* 74(8): 157-161

387 Ramos LP, Breuil C, Saddler JN, 1992. Comparison of steam pretreatment of Eucalyptus, Aspen and Spruce woodchips and their enzymatic hydrolysis. *Appl Biochem Biotechnol* 34(35): 37-47

388 Maekawa E, 1996. On available pretreatment for the enzymatic saccharification of lignocellulosic materials. *Wood Sci Technol* 30: 133-139

389 Väljamäe P, Kipper K, Pettersson G, Johansson G, 2003. Synergistic cellulose hydrolysis can be described in terms of fractal-like kinetics. *Biotechnol Bioeng* 84:254-257.

390 Filho, AZ, Siiko-Aho M, Ramos LP, 2006. Method for characterization of an enzyme profile and the determination of CBHI (Cel7a) core protein in *Trichoderma reesei* cellulase preparations. *World J Microbiol Biotechnol* 22: 821-825

391 Nidetzky B and Claeyssens M, 1994. Specific quantification of *Trichoderma reesei* cellulases in reconstituted mixtures anf its application to cellulase-cellulose binding studies. *Biotechnol Bioeng* 44: 961-966

392 Tolan J and Foody B, 1999. Cellulase from submerged fermentation. *Adc Biochem Eng Biotechnol* 65: 41-67

393 Tomme P, van Tilbeurgh H, Pettersson G, van Damme J, Vandekerckhove J, Knowles J, Teeri TT, Claeyssons M, 1988. Studies of the cellulytic system of *Trichoderma reesei* QM9414. *Eur J Biochem* 170: 575-581

394 Kipper K, Väljamäe P, Johansson G, 2005. Processive action of cellobiohydrolase Cel7A from Trichoderma reesei is revealed as 'burst' kinetics on fluorescent polymeric model substrates. *Biochem J* 385:527-535

395 Praestgaard E, Elmerdahl J, Murphy L, Nymand S, McFarland KC, Borch K, Westh P, 2011. A kinetic model for the burst phase of processive cellulases. *FEBS J* 278:1547-1560

396 Lynd LR, 1996. Overview and evaluation of fuel ethanol from cellulosic biomass: Technology, Economics, the Environment and Policy. *Annu Rev Energy Environ* 21:403–65

397 National Renewable Energy Laboratory [http://www.nrel.gov/biomass/biorefinery.html]

398 Cara C, Moya M, Ballesteros I, Negro MJ, Gonzalez A, Ruiz E, 2007. Influence of solid loading on enzymatic hydrolysis of steam exploded or liquid hot water pretreated olive tree biomass. *Process Biochem* 42:1003–1009

399 Lu Y, Wang Y, Xu G, Chu J, Zhuang Y, Zhang S, 2010. Influence of High Solid Concentration on Enzymatic Hydrolysis and Fermentation of Steam-Exploded Corn Stover Biomass. *Appl Biochem Biotechnol* 160:360–369

400 van den Brink J, van Muiswinkel GCJ, Theelen B, Hinz SWA, de Vries RP, 2003. Efficient plant biomass degradation by thermophilic fungus *Myceliophthora heterothallica. Applied Env Microbiol* 79(4):1316–1324

401 Gnansounou E, 2010. Production and use of lignocellulosic bioethanol in Europe: Current situation and perspectives. *Biores Technol* 101: 4842–4850

402 Stickel JJ, Knutsen JS, Liberatore MW, Luu W, Bousfield DW, Klingenberg DJ, Scott CT, Root TW, Ehrhardt MR, Monz TO, 2009. Rheology measurements of a biomass slurry: an inter-laboratory study. *Rheologica Acta* 48(9):1005-1015

403 Dasari RK, R. Berson E, 2007. The effect of particle size on hydrolysis reaction rates and rheological properties in cellulosic slurries. *Appl Biochem Biotechnol* 136–140: 289-299

404 Rosgaard L, Andric P, Dam-Johansen K, Pedersen S, Meyer AS, 2007. Effects of substrate loading on enzymatic hydrolysis and viscosity of pretreated barley straw. *Appl Biochem Biotechnol* 143(1):27 – 40

405 Ventura C, Blanco A, Negro C, Ferreira P, Garcia F, Rasteiro M, 2007. Modeling pulp fiber suspension rheology. *TAPPI J* 6(7):17-23

406 Blanco A, Negro C, Fuente E, Tijero J, 2007. Rotor selection for a searle-type device to study the rheology of paper pulp suspensions. *Chem Eng Process* 46(1):37-44

407 Badhan AK, Chadha BS, Kaur J, Saini HS, Bhat MK, 2007. Production of multiple xylanolytic and cellulolytic enzymes by thermophilic fungus *Myceliophthora sp*. IMI 387099. *Biores Technol* 98: 504–510

408 Badhan AK, Chadha BS, Sonia KG, Saini HS, Bhat MK, 2004. Functionally diverse multiple xylanases of thermophilic fungus *Myceliophthora sp*. IMI 387099. *Enzyme Microbial Technol* 35(5):460 – 466

409 Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, Bartels J, Visser J, Sinitsyn AP, Emalfarb MA, Verdoes JC, Wery J, 2011. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnol* 7(3): 214-223

410 Hult EL, Larsson PT, Iversen T, 2000. A comparative CP/MAS 13C-NMR study of cellulose structure in spruce wood and kraft pulp. *Cellulose* 7:35–55

411 Hong JTH, Yamamoto K, Kumagai H, 2003. Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacu*s and its expression in yeast. *Appl Microbiol Biotechnol* 63:42–50

412 Fan ZL, South C, Lyford K, Munsie J, van Walsum P, Lynd LR, 2003. Conversion of paper sludge to ethanol in a semicontinuous solids-fed reactor. *Bioprocess Biosyst Eng* 26:93–10

413 US DOE Joint Genome Institute [http://genome.jgi.doe.gov/Spoth2/Spoth2.home.html]

414 Aslanidis C, de Jong PJ, 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* 18(20):6069-6074

415Difco Laboratories,2003. Difco and BBL Manual of Microbiological Culture Media. Be cton, Dickinson and Company: Sparks Maryland

416 Center for Biological Sequence Analysis, Technical University of Denmark (DTU) [http://www.cbs.dtu.dk/services/SignalP/]

417 ExPASy: SIB Bioinformatics Resource Portal [http://web.expasy.org/protparam/]

418 Laemmli UK, 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227:680 - 685

419 Wasiak S, Legendre-Guillemin V, Puertollano R, Blondeau F, Girard M, de Heuvel E, Boismenu D, Bell AW, Bonifacino JS, McPherson PS, 2002. Enthoprotin: a novel clathrin-associated protein identified through subcellular proteomics. *J Cell Biol* 158 (5):855-62

420 Shinoda K, Tomita M, Ishihama Y, 2010. emPAI Calc—for the estimation of protein abundance from large-scale identification data by liquid chromatography-tandem mass spectrometry. *Bioinformatics* 26(4): 576–577

421 Ekwe E, Tsang A, Powlowski J, 2013. A microplate-based method for screening fungal enzymes on high consistency biomass. *Manuscript*

422 Grange JM, 1978. Fluorimetric assay of mycobacterial group-specific hydrolase enzymes. *J Clin Pathol* 31: 378-381

423 Ockerman PA, 1968. 4-Methyl-umbelliferyl-beta-glucosidase activity in human blood plasma. *Clin Chim Acta* 21(2):279-82

424 Fia G, Giovani G, Rosi I, 2005. Study of beta-glucosidase production by wine-related yeasts during alcoholic fermentation. A new rapid fluorimetric method to determine enzymatic activity. *J Appl Microbiol* 99(3):509-17

425 Currie HA, Perry CC, 2006. Resolution of complex monosaccharide mixtures from plant cell wall isolates by high pH anion exchange chromatography. *J Chromatogr A* 1128(1-2):90-6

426 Sherrad EC, Kressman FW, 1945. Review of processes in the United States prior to World War II. *Ind Eng Chem* 37(1):5–8

427 Humbird D, Mohagheghi A, Dowe N, *et al.,* 2010. Economic impact of total solids loading on enzymatic hydrolysis of dilute acid pretreated corn stover. *Biotechnol Prog* 26:1245-1251

428 Humbird D, Aden A, 2009. Biochemical production of ethanol from corn stover: National Renewable Energy Laboratory state of technology model. *NREL Technical Report* No. 2009/TP-510-46214, Golden, CO

429 van den Brink J, de Vries RP, 2011. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol* 91:1477-1492

430 Kim ES, Lee HJ, Bang W-G, *et al.,* 2009. Functional characterization of a bacterial expansin from Bacillus subtilis for enhanced enzymatic hydrolysis of Cellulose. *Biotechnol Bioeng* 102(5): 1342–1353

431 Chen X-I, Ishida N, Todaka N, *et al.,* 2010. Promotion of efficient Saccharification of crystalline cellulose by *Aspergillus fumigatus* Swo1. *Appl Environ Microbiol* 76(8):2556-2561

432 Harris PV, Welner D, McFarland KC, *et al.*, 2010. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochem* 49:3305-3316

433 Asztalos A, Daniels M, Sethi A, *et al.,* 2012. A coarse-grained model for synergistic action of multiple enzymes on cellulose. *Biotechnol Biofuels* 5(1):55

434 Li Y, McQueen-Mason S, 2003. Expansins and cell growth. *Curr Opin Plant Biol* 6:603–610

435 McQueen-Mason SJ, Durachko DM, Cosgrove DJ, 1992. Two endogenous proteins that induce cell wall extension in plants. *Plant Cell* 4:1425-1433

436 Shcherban TY, Shi J, Durachko DM, *et al.,* 1995. Molecular cloning and sequence analysis of expansins – a highly conserved, multigene family of proteins that mediate cell wall extension in plants. *PNAS* USA 92:9245–9249

437 Kende H, Bradford KJ, Brummell DA, *et al.,* 2004. Nomenclature for members of the expansin superfamily of genes and proteins. *Plant Mol Biol* 55:311–314

438 Sampedro J, Cosgrove DJ, 2005. The expansin superfamily. *Genome Biol* 6:242

439 Cosgrove DJ, 2000. Loosening of plant cell walls by expansins. *Nature* 407(6802):321–326

440 Quiroz-Castañeda RE, Martínez-Anaya C, Cuervo-Soto LI, *et al.,* 2011. Loosenin, a novel protein with cellulosedisrupting activity from *Bjerkandera adusta. Microbial Cell Fact* 10:8

441 Georgelis N, Tabuchi A, Nikolaidis N *et al.,* 2011. Structure-function analysis of the bacterial expansin EXLX1. *J Biol Chem* 286(19):16814–16823

442 Yennawar NH, Li LC, Dudzinski DM, *et al.,* 2006. Crystal structure and activities of EXPB1 (Zea m 1), a β-expansin and group-1 pollen allergen from maize. *PNAS USA* 103(40):14664–14671

443 Kerff F, Amoroso A, Herman R, *et al.,* 2008. Crystal structure and activity of *Bacillus subtilis* YoaJ (EXLX1), a bacterial expansin that promotes root colonization. *PNAS USA* 105(44):16876–16881

444 Henrissat B, Teeri TT, Warren RAJ, 1998. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Lett* 425:352-354

445 Phillips CM, Beeson WT, Cate JH *et al.,* 2011. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol* 6:1399-1406

446 Westereng B, Ishida T, Vaaje-Kolstad G, *et al.,* 2011. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE* 6:e27807

447 Quinlan RJ, Sweeney MD, Lo Leggio L, *et al.,* 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *PNAS USA* 108:5079-15084

448 Boraston AB, Bolam DN, Gilbert HJ, *et al.,* 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* 382(3):769-781

449 Forsberg Z, Vaaje-Kolstad G, Westereng B, *et al.,* 2011. Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20:1479–1483

450 Zifcakova L, Baldrian P, 2012. Fungal polysaccharide monooxygenases: new players in the decomposition of cellulose. *Fungal ecology* 5:481-489

451 Purushotham P, Arun PVPS, Prakash JSS, *et al.,* 2012. Chitin binding proteins act synergistically with chitinases in *Serratia proteamaculans* 568. *PLoS ONE* 7(5): e36714. doi:10.1371/journal.pone.0036714

452 Häkkinen M, Arvas M, Oja M, *et al.,* 2012. Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. *Microb Cell Fact* 11:134 doi:10.1186/1475-2859-11-134

453 Berka RM, Grigoriev IV, Otillar R, *et al.,* 2011. Comparative genomic analysis of the thermophilic biomass degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nature Biotech* 29:922-927

454 Longoni P, Rodolfi M, Pantaleoni L, *et al.,* 2012. Degradation of cellulosic substrates by a *Chaetomium globosum* endophytic isolate: a functional analysis. *Appl Environ Microbiol* 78:3693–3705

455 Hara Y, Hinoki Y, Shimoi H, *et al.,* 2003. Cloning and sequence analysis of endoglucanase gene from an industrial fungus, *Aspergillus kawachii. Biosci Biotechnol Biochem* 67(9): 2010-2013

456 Georgelisa N, Yennawarb NH, Cosgrove DJ, 2012. Structural basis for entropy-driven cellulose binding by a type-A cellulose-binding module (CBM) and bacterial expansin. *PNAS* 109(37):14830–14835

457 Kim IJ, Ko H-J, Kim T-W, *et al.,* 2013. Characteristics of the binding of a bacterial expansin (*Bs*EXLX1) to microcrystalline cellulose. *Biotech Bioeng* 110(2):401-407

458 Li LC, Bedinger PA, Volk C, *et al.* 2003. Purification and characterization of four β-expansins (Zea m 1 isoforms) from maize pollen. *Plant Physiology* 132(4):2073-2085

459 Ayub R, Rombaldi C, Lucchetta L, *et al.,* 2008. Mechanisms of melon fruit ripening and development of sensory quality. In Pitrat M, editor. *Proceedings of the IXth EUCARPIA meeting on genetics and breeding of Cucurbitaceae*. *Montfavet (France):INRA* 241–248

460 Catala C, Rose JKC, Bennett AB, 2000. Auxin-regulated genes encoding cell wall-modifying proteins are expressed during early tomato fruit growth. *Plant Physiol*122:527–534

461 Catala C, Rose JKC, Bennett AB, 1997. Auxin regulation and spatial localization of an endo-1,4-beta-d-glucanase and a xyloglucan endotransglycosylase in expanding tomato hypocotyls. *Plant J* 12:417–426

462 Carpita NC, 1996. Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol* 47:445–476

463 Yao Q, Sun T-T, Liu W-F, *et al.,* 2008. Gene cloning and heterologous expression of a novel endoglucanase, swollenin, from *Trichoderma pseudokoningii* S38. *Biosc Biotechnol Biochem* 72(11):2799–2805

464 Henriksson, G, Johansson G, Petersson G, 2000. A critical review of cellobiose dehydrogenase. *J Biotechnol* 78:93–113

465 Cameron MD, Aust SD, 2001. Cellobiose dehydrogenase—an extracellular fungal flavocytochrome. *Enzyme Microb Technol* 28:129–138

466 Zamocky M, Ludwig R, Peterbauer C, *et al*., 2006. Cellobiose dehydrogenase—a flavocytochrome from wood-degrading, phytopathogenic and saprotropic fungi. *Curr Protein Pept Sci* 7**:**255–280

467 Langston JA, Brown K, Xu F, *et al*., 2012. Cloning, expression, and characterization of a cellobiose dehydrogenase from *Thielavia terrestris* induced under cellulose growth conditions. *BBA* 1824:802-812

468 Wymelenberg AV, Gaskell J, Mozuch M, *et al.,* 2010. Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. *Appl Environ Microbiol* 76:3599–3610

469 Suwannarangsee S, Bunterngsook B, Arnthong J, *et al*., 2012. Optimisation of synergistic biomass-degrading enzyme systems for efficient rice straw hydrolysis using an experimental mixture design. *Biores Technol* 119:252–261

470 Lee H, Lee S, Ko H-J, *et al*., 2010. An expansin-like protein from *Hahella chejuensis* binds cellulose and enhances cellulase activity. *Mol Cell* 29:379–385

471 Gao D, Chundawat SP, Krishnan C, *et al*., 2010. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresour Technol* 101(8):2770-2781

472 Kim IJ, Ko HJ, Kim TW, *et al*., 2012. Binding characteristics of a bacterial expansin (BsEXLX1) for various types of pretreated lignocellulose. *Appl Microbiol Biotechnol* DOI 10.1007/s00253-012-4412-4416

473 Baker JO, King MR, Adney WS, *et al*., 2000. Investigation of the cell–wall loosening protein expansin as a possible additive in the enzymatic saccharification of lignocellulosic biomass. *Appl Biochem* Biotechnol 84–86:217-223

474 Wei W, Yang C, Luo J, *et al*., 2010. Synergism between cucumber α-expansin, fungal endoglucanase and pectin lyase. *J Plant Physiol* 167: 1204–1210

475 Jäger G, Girfoglio M, Dollo F, *et al*., 2012. How recombinant swollenin from *Kluyveromyces lactis* affects cellulosic substrates and accelerates their hydrolysis. *Biotechnol Biofuels* 4:33

476 Wang Y, Tang R, Tao J, *et al*., 2011. Quantitative investigation of non-hydrolytic disruptive activity on crystalline cellulose and application to recombinant swollenin. *Appl Microbiol Biotechnol* 91:1353–1363

477 Dimarogona M, Topakas E, Olsson L *et al*., 2011. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile. Biores Technol* 110:480-487

478 Okamoto A, Okamoto H, 1995. Two proteins regulate the cell wall extensibility and the yield threshold in glycerinated hollow cylinders of cowpea hypocotyl. *Plant Cell Environ* 18:827–830

479 Inouhe M, Nevins D, 1997. Regulation of cell wall glucanase activities by non-enzymic proteins in maize coleoptiles. *Int J Biol Macromol* 21:15–20

480 Foreman PK, Brown D, Dankmeyer L, *et al*., 2003. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei. J Biol Chem* 278:31988-31997