BIOMASS PRETREATMENT TOWARD EFFICIENT HYDROLYSIS FOR SUSTAINABLE BIOFUEL APPLICATIONS

A Dissertation
Presented to
The Academic Faculty

by

Yuzhi Kang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology
May 2015

BIOMASS PRETREATMENT TOWARD EFFICIENT HYDROLYSIS FOR SUSTAINABLE BIOFUEL APPLICATIONS

Approved by:

Dr. Andreas S. Bommarius, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Matthew J. Realff, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Jay H. Lee
Department of Chemical and Biomolecular Engineering, *Korea Advanced Institute of Science and Technology (KAIST)*, and School of Chemical and Biomolecular Engineering,
Georgia Institute of Technology

Dr. Pamela Peralta-Yahya
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Joy Doran-Peterson
Department of Microbiology
University of Georgia

Date Approved: December 3rd, 2014

To my parents and grandparents.

篤　明　慎　博
行　辨　思　学

Learn extensively, think carefully, distinguish wisely, practice earnestly.

ACKNOWLEDGEMENTS

I strongly believe the 5+ years stay at Georgia Institute of Technology is a life-changing event remarkably broaden my horizon and enable me to explore the unknown future with great confidence. I owe my thanks to many people who accompanied, encouraged and helped me one way or another during these years.

Firstly, I wish to extend my special thanks to my advisors Dr. Andreas S. Bommarius and Dr. Matthew J. Realff for their continuous support, inspiring input and valuable feedbacks throughout my PhD study at Georgia Tech. Without their guidance, I could never have successfully finished my thesis work. I also thank my committee members Dr. Jay H. Lee, Dr. Pamela Peralta-Yahya and Dr. Joy Doran-Peterson for their time, insights and precious input. Dr. Jay H. Lee and Dr. Joy Doran-Peterson, physically located in South Korea and Athens, GA respectively, have been extremely supportive to accommodate committee meetings in unusual hours. Dr. Arthur Ragauskas is thanked for his time and service on my thesis committee prior to his departure to Oak Ridge National Laboratory.

Dr. Mélanie Hall, a former postdoctoral researcher in the Bommarius lab, guided me through the beginning of lab research, helped me with different assays and analytical techniques and conducted valuable initial studies for this project. Dr. Prabuddha Bansal has been a great collaborator ready to share insights and contributed with his expertise in computational modeling. I would like to express my gratitude to both Dr. Mélanie Hall and Dr. Prabuddha Bansal. I am also grateful to Dr. Shin-geon Choi and Dr. Minjeong

Sohn, both were visiting scholars from South Korea, for their contributions in different aspects of this work.

I would also like to thank the undergraduate researchers who worked with me on this project and have been of great help for the past four years: Sang beom Kim, Bryan Vowell, Jessica Curtis, Alec Seco, Bethany Carnes and Christopher Hoffman.

Grateful thanks are also extended to all the current and former group members in the Bommarius lab, Bettina, Mick, Michael, Ryan, Sam, Aditi, Lizzette, Harrison, Minjeong and others, the moments we spent together inside and outside the lab will be an unforgettable time for me.

I would like to thank my dear friends Roxy, Sihong, Xiaodan, Elaine, Chunhong, Yanfang, Xue, Dieh, Peng, Renbo, Boyi, Chen, Wei, Huayu, Chengwei and many others for being great companions and supports through all these fun times and stressful moments in the United States, making my stay here enjoyable and memorable.

Lastly, I would like to express my deepest thanks and gratitude to my parents, my grandparents and my entire family for their endless love. Without their support, I would never have the chance to fly over the Pacific Ocean, arrived at the country full of challenges and opportunities to complete my thesis work.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

SUMMARY

The production of biofuels from non-edible plant biomass has been necessitated by the concern for the environmental consequences of fossil fuel use and the tightening of supply and demand for liquid fuels. In contrast to first generation biofuels which rely on crops used for food supplies, second generation biofuels, derived from lignin-containing feedstocks, completely eliminate the competition for food. The major challenges associated with second generation biofuels are both technical and economic. Due to the recalcitrant nature of the raw biomass materials to further biological conversion, their structural degradation often requires severe and costly pretreatment processes such as heat, physical and chemical treatments to disturb and fractionate the biomass. Significant research effort has been devoted to understanding the recalcitrant nature and to accelerate the commercialization process of second generation biofuels. In this thesis, three pretreatment methods that belong to different categories have been investigated to understand their impacts on cellulose and/or lignocellulose and the subsequent hydrolysis steps.

Physicochemical pretreatments, such as steam explosion, have been identified as one of the most effective and cost-efficient pretreatment methods for lignocelluosic materials. In Chapter 2, SO_2-catalyzed steam explosion will be discussed and the effect of pretreatment severity on the substrate characteristics and degradation efficiency is also elucidated. Although the crystallinity index (CrI) of cellulose decreases as the severity increases, significant non-specific degradation and low yield of cellulose was observed at

high severity. A new method for cellulose CrI determination has been developed with least squares curve fitting and validated with mechanically mixed cellulose samples.

Biological pretreatment is another pathway through which the biomass structure can be modified to obtain a more amenable state for enzymatic degradation. Cellulose-binding domain (CBD) originated from *Trichoderma reesei* Cel7A (i.e. *Tr* cellobiohydrolase I) has been discovered as a potential biological pretreatment agent which is capable of modifying cellulose crystal structure. An extensive study on the protein engineering, expression, purification and functionalities of Cel7A CBDs was carried out (Chapter 3). The target mutations were identified with a computational protein engineering method involving principal component analysis (PCA). Due to the lack of catalytic activity and high throughput screening method, the library size was limited to nine. The wild-type and mutated CBDs were compared for their adsorption behavior and decrystallization effect on cellulose. Resulting saccharification efficiency after CBD pretreatment were studied and a possible explanation for the rate enhancement was proposed.

In addition to physicochemical and biological pretreatment methods, chemical pretreatment is also a commonly employed method to overcome the recalcitrance of lignocellulosic materials. The most widely studied include dilute acid, alkaline, and organosolv processes. Inspired by the rapidly growing green solvent ionic liquid (IL) researches in biomass pretreatment, substituted imidazoles have been investigated in this thesis to assess their potential as pretreatment agents for lignocelluloses (Chapter 4). 1-Methylimidazole (MI), a precursor to some ILs, has been determined to be the most promising agent for lignocellulose pretreatment due to its exceptional delignification and

cellulose expansion efficiency. The chemical recovery and MI process development will also be discussed in Chapter 4.

In order to understand pretreatment effect, a semi-quantitative assay utilizing low molecular weight direct dyes and cellulases to estimate the accessibility and pore size distribution has been developed for application on pure cellulose substrates in Chapter 5. Finally, main conclusions as well as future perspectives for this work will be discussed in Chapter 6.

CHAPTER 1

INTRODUCTION

1.1 Overview of biofuel production

Fossil fuels, including petroleum, coal and natural gas, are the primary energy sources used by human society. Fossil fuels have supported the industrialization and economic growth of the world for the past century (Huber, Iborra et al. 2006). However, they are finite resources and the combustion of fossil fuels is a major cause of environmental problems and global climate change. According to the data published by Environmental Protection Agency, the burning of fossil fuels accounted for the vast majority of U.S. CO_2 emission for the period of 1990 – 2011. The contribution of different energy-related economic sectors to total CO_2 emission in 2011 is shown in **Figure 1**.

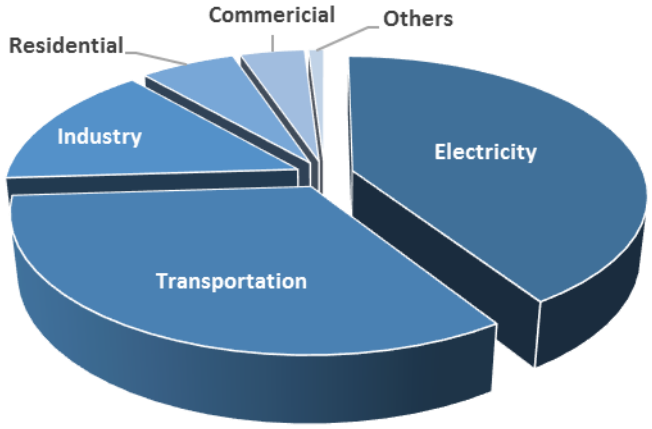


Figure 1. Total U.S. CO_2 emission by economic sector in 2011 (Data obtained from *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2011*)

One of the main greenhouse gases (GHG) emission contributors is transportation, which primarily comes from the burning of petroleum-based fuels such as gasoline and diesel. To alleviate the GHG emission from transportation and conserve the non-renewable fossil fuels, globally researchers have devoted significanteffort to find alternative renewable liquid transportation fuel sources.

Biofuels produced from renewable resources have shown promise. They are derived from the most abundant naturally-occurring carbon-rich resource on the earth – plant biomass. CO_2 in the atmosphere is fixed in the plant through photosynthesis in the form of natural polymers, thus no net CO_2 is released to the environment upon combustion. Therefore biofuels are considered to be essentially carbon neutral.

First generation biofuels are produced with starch or simple sugars from agricultural crops: corn, sugar cane and beets. These have been commercialized already and almost 50 billion liters are produced annually (Naik, Goud et al. 2010). The major concerns regarding the first generation biofuels are two-fold: adverse environmental and socioeconomic impact, with the food versus fuel debate being the most emblematic one (Gasparatos, Stromberg et al. 2013). Due to the large scale production of these biofuels, food price and security have become pressing concerns. As a result, a more sustainable alternative biofuel is desired.

Second generation biofuels, derived from non-food lignocellulosic biomass, emerged in response to the concerns regarding the food-based biofuels. They eliminate the food competition and are perceived as environmentally benign. The plant cell wall of lignocellulosic biomass is comprised of three carbon-rich main components: cellulose, hemicellulose and lignin. The main function of the cell wall is to confer the plant

resistance, rigidity and protection to exotic stress and attack (Ochoa-Villarreal, Aispuro-Hernández et al. 2012). Cellulose is by far the most abundant bio-renewable polymer on earth. It is consisted of linear, non-branched chains of several hundred to ten thousand of β-1,4 linked glucose. The cellulose chains aggregate together, mainly through H-bonding and van der Waal forces, to form compact, highly-ordered microfibrils with high crystallinity and resistance to enzymatic attack. Hemicellulose, another plant polymer, is comprised of a variety of branched and esterified polysaccharides formed with xylose, mannose, arabinose, galactose and glucose linked by various glycosidic bonds (Donohoe, Decker et al. 2008, Yarbrough, Himmel et al. 2009). The low molecular weight heterogeneous polysaccharides bind on the surface of cellulose to prevent direct contact among microfibrils (Ochoa-Villarreal, Aispuro-Hernández et al. 2012). In addition, hemicelluloses are also found in associated with lignin to form lignin-carbohydrate complexes (LCC) which create barrier for enzymatic degradation of plant cell wall (Wallace, Chesson et al. 1991). Lignin is the second most abundant renewable polymer; it is a three-dimensional phenolic compound based on paracoumaryl alcohol, coniferyl alcohol, and sinapyl alcohol (Thakur, Thakur et al. 2014). The existence of lignin reinforces the plant cell wall structure and renders lignocellulosic materials resistant to water and biological degradation (Davin and Lewis 2005).

Since cellulose is surrounded by the amorphous plant polymer matrix comprised of hemicelluloses and lignin (**Figure 2**), its accessibility to cellulolytic enzymes is very poor. Combined with the crystalline characteristics of cellulose, the lignocellulosic biomass is considered to be extremely recalcitrant to enzymatic attack. Therefore, a

process step which aids in the disintegration of the complex 3D cell wall structure and expose cellulose microfibril is required prior to enzymatic degradation.

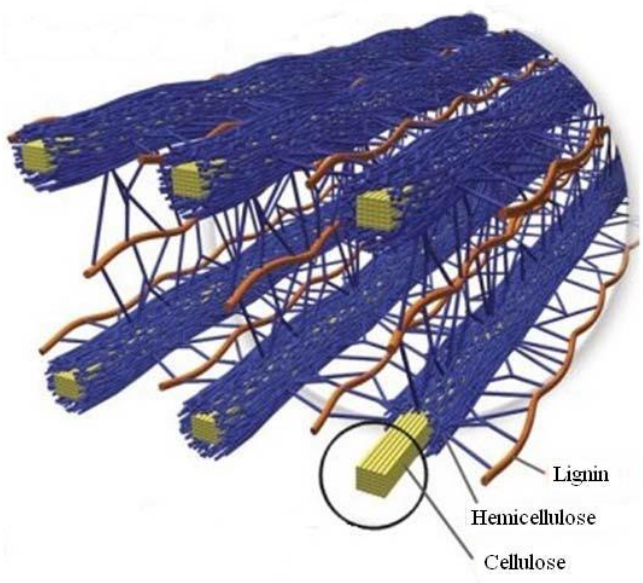


Figure 2. Cellulose microfibril embedded in hemicellulose and lignin matrix (Doherty, Mousavioun et al. 2011).

1.2 Various routes to accelerate biofuel commercialization

Biomass recalcitrance is largely responsible for the high cost of second generation biofuel production by biological routes (Himmel, Ding et al. 2007). Although biomass processing for biofuel purposes has received substantial attention over the past decades, full scale commercial production does not exist to date. A recent review paper summarized the technology development challenges as following (Sims, Mabee et al. 2010):

i) Improve understanding of feedstocks

ii) Technology improvement for the biochemical routes

iii) Co-products and process integration

In order to address the long-existing issues regarding the process efficiency and fundamental understanding of biomass and its degradation, a variety of research routes are being pursued: new technique development for characterization and imaging at nanometer scale (Himmel, Ding et al. 2007, Igarashi, Koivula et al. 2009, Dagel, Liu et al. 2010), discovery and engineering of the cellulolytic enzymes, feedstock engineering and process integration. In this chapter, we will mainly focus on biomass pretreatment techniques.

Typically, the production of liquid biofuels from lignin-containing biomass includes four key unit operations: pretreatment, enzymatic hydrolysis, fermentation and production separation/purification (**Figure 3**) (Mosier, Wyman et al. 2005).

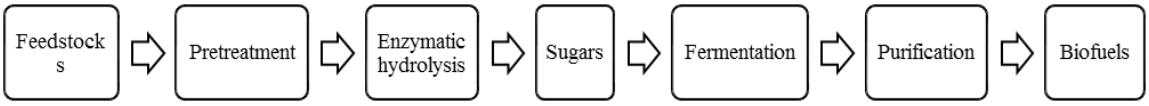


Figure 3. Key unit operations involved in second generation biofuel production

Pretreatment is a crucial step to overcome biomass recalcitrance, however, it is usually inefficient and costly. Pretreatment is considered to be the most expensive processing step which accounts for at least 20% of the total cost (Yang and Wyman 2008), but it has major influence on facilitating and boosting the efficiency of downstream processes such as hydrolysis and fermentation. The goals for all pretreatment techniques include but are not limited to the dissolution of lignin and hemicellulose, reduction in cellulose crystallinity, improvement in porosity and cellulose accessibility. Overall, the goal is to enhance the cellulose hydrolysis yield for conversion to fuels

(Kumar, Barrett et al. 2009). Promising pretreatment techniques should have the following features:

i) Environmentally and economically friendly

ii) Effective and efficient

iii) Selective and minimum sugar degradation

A huge variety of pretreatment techniques have been developed to overcome biomass recalcitrance to achieve amenable substrates. Nonetheless, the majority of pretreatment techniques cannot efficiently disrupt the biomass structure at a reasonable cost with minimal to moderate environmental impact and high selectivity. The commonly utilized and extensively studied pretreatment techniques can be divided into four categories: physical, physicochemical, chemical and biological, as shown in **Table 1**.

Table 1. Pretreatment techniques and categories.

Category	Pretreatment technique
Physical	Ball milling
Physicochemical	Steam explosion (SE), Ammonia fiber explosion (AFEX)
Chemical	Organosolv, Acid, Alkali
Biological	White-/brown-rot fungi

1.3 Overview of biomass pretreatment techniques

1.3.1 Physical pretreatment

Ball milling

The most intuitive route to improve digestibility of biomass is through size reduction. For the heterogeneous enzymatic reaction system, the soluble cellulolytic enzymes needs to adsorb onto the insoluble biomass/cellulose surface to perform hydrolytic reaction. Thus, size reduction is desired to create larger surface area for enzymes to attach to and perform catalysis. Ball milling is the most widely applied physical pretreatment method for biomass. Effective size reduction and cellulose decrystallization can be achieved with chipping and ball milling for the woody biomass, with typical particle size of 10-30 mm and 0.2-2 mm respectively (Kumar, Barrett et al. 2009). However, milling efficiency is the major concern for this pretreatment as it is energy-intensive and costly for process scale application.

Conventional ball milling is primarily conducted in dry state or water medium. Recently, advances have been made to increase the milling efficiency via combination with other pretreatment methods or chemicals. Inoue et al. reported a combined ball milling (20 min), hot compressed water (160 ℃, 30 min) pretreatment on eucalyptus wood chips without addition of chemicals to be capable of yielding 70% total sugar with cellulase loading as low as 4 FPU/g (Inoue, Yano et al. 2008). While near complete hydrolysis was observed for the same feedstock pretreated by ball milling combined with ethanol/acetic acid hydrothermal treatment (Teramoto, Tanaka et al. 2008). For both cases, enzyme and energy loading can be significantly reduced. The ball milling time could also be shortened dramatically without affecting subsequent hydrolysis yield by simply change the medium from water to dilute alkali due to the delignification effect

(Lin, Huang et al. 2010). Based on these findings and through rational experimental design, the energy consumption of ball milling is likely to be optimized further.

1.3.2 Physicochemical pretreatment

Un-catalyzed and catalyzed steam explosion

Steam explosion (SE) is a conventional and one of the most effective pretreatment technique for a variety of biomass. It is usually performed at high pressure and temperature (160 - 240℃) for a short period of residence time (< 1 hour), a sudden pressure drop caused explosive discharge of the residual solids terminates the reaction and disrupt the integrity and rigidity of biomass. The easily solubilized hemicellulose is dissolved in the liquid fraction, whereas cellulose and lignin mainly retained in solid fraction. Due to the rapid pressure release, the cellulose microfibrils structure is disrupted, accompanied with accessibility enhancement and favorable enzymatic activity. The selectivity of SE varies with pretreatment condition. Generally speaking, more severe conditions result in stronger degradation and less retention of cellulose.

Depending on the origin of biomass of interest, SE can be carried out with or without catalyst. Un-catalyzed SE is also termed as autohydrolysis; no additional chemical is needed in the process, and it is thought to be one of the most cost-effective pretreatment techniques. Hemicellulose is hydrolyzed by acetic acids formed from the release of acetyl groups during the pretreatment process (Zheng, Pan et al. 2009). Catalyzed SE employs external acid catalysts, SO_2 or H_2SO_4, to impregnate lignin-containing biomass before SE. The catalyzed or un-catalyzed method selection is largely dependent on substrates. Similarly, SE efficiency and processing optimum is also highly substrate dependent (Kaar, Gutierrez et al. 1998).

For hardwood and herbaceous biomass, especially agricultural residues, un-catalyzed SE is frequently found to render satisfactory hemicellulose dissolution and cellulose digestibility improvement (Dekker and Wallis 1983, Ballesteros, Negro et al. 2006, Liu, Qin et al. 2013). The addition of acid catalysts is particularly beneficial for woody biomass, which shows higher lignin content and more complex structure, the catalysts could substantially improve the survival of C5 sugars and boost enzymatic digestibility of water-insoluble cellulose (Mackie, Brownell et al. 1985, Brownell, Yu et al. 1986, Shevchenko, Chang et al. 2000). Nonetheless, the generation of inhibitory compounds, such as acetic acid, furfural and hydroxymethylfurfural, is inevitable under typical SE conditions (Cantarella, Cantarella et al. 2004), advances have been made to address this issue.

Two-step SE, with a low severity 1st step and a higher severity 2nd step, has been exercised to optimize the hydrolysis of hemicellulose, selectivity of cellulose and subsequent fermentation step for a variety of feedstocks (Söderström, Pilcher et al. 2002, Söderström, Pilcher et al. 2003, Zhang, Fu et al. 2012). Moreover, some studies explored the possibility of combing SE with biological pretreatment to reduce the SE severity and further enhance the effectiveness of the combined pretreatment techniques (Sawada, Nakamura et al. 1995, Zhang, Li et al. 2008, Qiu and Chen 2012). However, these combined techniques are time-consuming. The pretreatment throughput is largely limited by the biological step.

Ammonia fiber explosion (AFEX)

Ammonia fiber/freeze explosion (AFEX) is conceptually similar to SE (Kumar, Barrett et al. 2009, Zheng, Pan et al. 2009). Liquid anhydrous ammonia (1-2 kg/kg dry

biomass) is used to pretreat biomass materials under moderate temperature (< 90 ℃) for certain residence time (< 30 min) (Zheng, Pan et al. 2009). The biomass is subjected to pretreatment with hot liquid ammonia, upon sudden pressure drop, the explosive discharge creates disruption of the cellulose crystal structure and substantial enhancement in cellulose degradation and fermentation rate (Kumar, Barrett et al. 2009). Due to the mild pretreatment temperature and ammonolysis reaction (Lau, Gunawan et al. 2009), the resulting inhibition is much less than other leading pretreatment technique such as SE and dilute acid. The main inhibitors from AFEX pretreatment that influence fermentation step are acetamide, feruloyl amide, coumaroyl amide (Chundawat, Vismeh et al. 2010). Wild-type and engineered *Saccharomyces cerevisiae* strains were found to tolerate these inhibitors, opening the chance to conduct simultaneous saccharification and co-fermentation (SSCF) of the AFEX-pretreated solid residues in one reactor (Jin, Sarks et al. 2013).

An additional benefit of the AFEX pretreatment is the minimal solid loss. All three major components cellulose, hemicellulose and lignin are largely retained in the solid residue with the lignin-carbohydrate bonds cleaved (Brodeur, Yau et al. 2011). Both loosely associated hemicellulose and cellulose are available to be further processed to produce biofuels and chemicals, rendering the AFEX technique with greater potential to be fully commercialized.

Nevertheless, there is a significant drawback of AFEX; although it is effective on crops, grass and agricultural residues (Gollapalli, Dale et al. 2002, Alizadeh, Teymouri et al. 2005, Murnen, Balan et al. 2007), the effect does not hold for woody biomass with higher lignin content (Holtzapple, Jun et al. 1991).

1.3.3 Chemical pretreatment

Organic solvents

The organic solvent pretreatment technique is possibly the fastest growing one in this category. Tremendous effort has been made to discover and identify novel solvents that are capable of dissolving lignocellulosic biomass or its components. Based on their different solubilization properties, the organic solvents used for biomass pretreatment can be divided into two groups: cellulose solvents and non-cellulose solvents. The application of organic solvents for biofuel purpose is mainly based on the solvent fractionation of biomass (da Costa Sousa, Chundawat et al. 2009).

Organic solvent: N-methyl morpholine N-oxide (NMMO)

Complete dissolution of cellulose has been studied extensively due to the need for analysis, regeneration and homogeneous chemical modification in polymer industry (Liebert 2010). The cellulose solvents for biofuel applications can sometimes be adapted directly from those used in the field of polymer production, one example is N-methyl morpholine N-oxide (NMMO). NMMO is known as an industrial solvent in the Lyocell process for Tencel® fiber production (Rosenau, Potthast et al. 2001). It dissolves cellulose by interrupting the hydrogen bond network with its high polarity N− O bond (Sathitsuksanoh, George et al. 2013). It is an environmentally friendly technique which can directly solubilize cellulose without derivatization (Adorjan, Sjöberg et al. 2004). As a result of its excellent cellulose-dissolving property, NMMO can be readily applied to lignocellulosic biomass pretreatment toward enhanced hydrolysis and fermentation efficiency. NMMO pretreatment effectiveness has been investigated on a variety of biomass species including agricultural residues, herbaceous materials, hardwood and

softwood; effective biomass dissolution and subsequent enzyme degradation and/or fermentation were observed for the regenerated biomass (Kuo and Lee 2009, Shafiei, Karimi et al. 2010). Interestingly, a recent publication studied *in situ* enzymatic hydrolysis of NMMO-pretreated dissolving pulp without the regeneration and washing steps and observed superior reaction rate for the initial 5h (Ramakrishnan, Collier et al. 2010), indicating mild enzyme inactivation induced by NMMO.

Organic solvent: Ionic liquid (IL)

Another rapidly developing organic solvent pretreatment technique is ionic liquid (IL), which is defined as molten organic salt with melting point below 100℃ (Gericke, Fardim et al. 2012). IL is known as green solvents whose physical and chemical properties can be fine-tuned by altering the cation-anion combinations (Tadesse and Luque 2011). The first work revealing the dissolution of cellulose in ILs was published in 2002; the results demonstrated imidazolium-based ILs with anions which are strong hydrogen bond acceptors could be efficient cellulose solvents (Swatloski, Spear et al. 2002). Since then, ILs have been extensively studied for their suitability as biomass pretreatment agents. The frequently employed ILs for biomass dissolution and pretreatment are shown in **Figure 4** (Tadesse and Luque 2011).

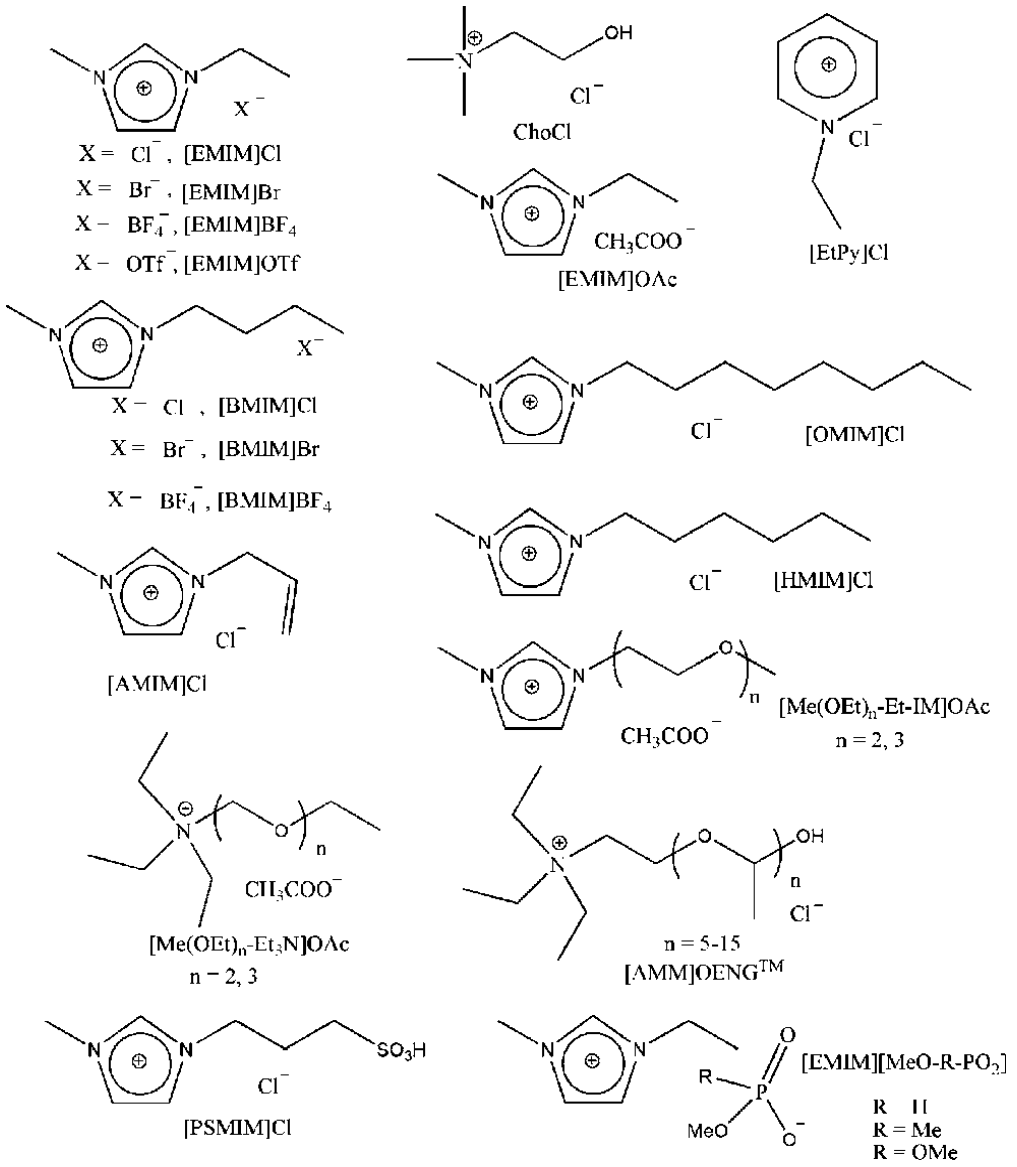


Figure 4. Commonly used ILs for biomass dissolution and pretreatment (Tadesse and Luque 2011). EMIM: 1-ethyl-3-methylimidazolium; BMIM: 1-butyl-3-methylimidazolium; AMIM: 1-allyl-3-methylimidazolium; HMIM: 1-hexyl-3-methylimidazolium; PSMIM: 1-methyl-3-propane sulfonic-imidazolium.

A wide variety of biomass, including herbaceous and woody, were subjected to IL pretreatment and resulted in complete dissolution under appropriate conditions (Brodeur, Yau et al. 2011). Upon regeneration, the crystal structure is almost completely disrupted, leaving behind amorphous cellulose rich solids ready to be hydrolyzed. ILs can inhibitive

to both cellulolytic enzymes and fermentative microorganisms, thus extensive wash step is required to prevent failure in subsequent processing. Recent research shown the IL-tolerant cellulase, which remain active in up to 40 v/v% various ILs, could be formulated and allowing the reduction in water consumption (Gladden, Park et al. 2014, Xu, He et al. 2014). A strain of IL-tolerant, thermophilic fermentative microorganism *Bacillus coagulans* was isolated and demonstrated to maintain growth at up to 4 wt% 1-ethyl-3-methylimidazolium based ILs (Simmons, Reddy et al. 2014). The discovery of IL-tolerant enzymes and microorganisms has the potential to greatly simplify the process and reduce the cost.

The most significant drawback associated with the IL pretreatment technique is the solvent cost and recovery. The solvent usage can be reduced by recycling; since ILs exhibit extremely low vapor pressure, more than 99% solvents could be recovered theoretically (Brodeur, Yau et al. 2011). Some low-cost ILs have been synthesized and studied for their effects on biomass pretreatment (Chen, Sharifzadeh et al. 2014, Verdía, Brandt et al. 2014). The $[HSO_4^-]$-based ILs were found to be cost competitive with regular organic solvents (Chen, Sharifzadeh et al. 2014); upon the successful application of the low-cost IL in large scale, the commercialization of IL-based technology can be greatly promoted.

Organic solvent: Other solvents

The abovementioned organic solvent pretreatment techniques are cellulose solvent-based. In addition, there are many other solvents or solvent systems that were found to be beneficial in the biomass-derived biofuel or biochemical applications.

Monoethanolamine (MEA) and gamma-lactone were reported to be suitable for utilization in organosolv pulping processes to yield cellulose-rich solids by dissolving lignin (Claus, Kordsachia et al. 2004, Petrus and Petrus-Hoogenbosch 2006). Furthermore, gamma-valerolactone, which is derived from cellulose, could be used as a green solvent for lignocellulosics; upon combining with dilute acid, the solvent system allows direct nonenzymatic production of sugar from biomass (Luterbacher, Rand et al. 2014). Another co-solvent system using tetrahydrofuran (THF) was observed to promote direct production of hydrocarbon fuel precursors such as furfural and hydroxymethylfurfural with high yield from woody biomass (Cai, Zhang et al. 2013).

Acid

Both concentrated and diluted acid were employed to destroy the biomass structure to facilitate subsequent process steps. Dilute acid (DA) pretreatment is considered to be one of the most promising pretreatment techniques owing to its low cost and high efficiency (Nguyen, Choi et al. 2009). The most prevalent acid used in the DA process is sulfuric acid, usually with concentration less than 4 wt% (Kumar, Barrett et al. 2009).

The DA process is a cost-efficient method to release hemicellulose and create void volume in the biomass for enzymatic attack. Depending on the severity, the hemicellulose can be hydrolyzed to oligomeric, monomeric sugar and further degraded to produce furfural and hydroxymethylfurfural (Mosier, Wyman et al. 2005). Higher DA severity is favored for cellulose digestibility, especially for woody biomass, however, more sugar degradation products will form at this condition and an extensive washing is required to detoxify the residue solids. Two-step DA pretreatment was studied to

maximize the recovery of both hemicellulose and cellulose sugars from the more recalcitrant biomass (Söderström, Pilcher et al. 2003, Sannigrahi, Ragauskas et al. 2008). The first step is performed at lower severity to remove a significant fraction of hemicellulose without extensive degradation, followed by a second step at higher severity to render the cellulose high amenability.

Compared to DA, concentrated acid pretreatment is much less popular, probably due to its high cost and corrosive, hazardous nature. Recovery of the acid is essential to make the process economically feasible. The most common concentrated acid pretreatment techniques employ sulfuric acid and phosphoric acid. A pretreatment technique named cellulose solvent- and organic solvent-based lignocellulosics fractionation (COSLIF) could be a valuable pretreatment method which combines concentrated phosphoric acid (85%) and ethanol to fractionate the biomass at moderate temperature based on different solubility of cell wall components in these solvents (Zhang, Zhu et al. 2010). The COSLIF-pretreated cellulose-rich residues achieve high digestibility (> 90% within 24 hours) at low enzyme loading (5 FPU/g glucan) for various biomass including corn stover, switch grass, hemp hurds and hybrid poplar (Zhang, Zhu et al. 2010). However, the recovery of hemicellulose sugars and phosphoric acid from the black liquor is crucial for its applicability in large scale.

Alkali

Alkali pretreatment employs bases, most popularly sodium hydroxide and lime (calcium hydroxide) to alter the biomass structure. It primarily dissolves lignin and hemicellulose to allow for high cellulose accessibility to hydrolytic enzymes (Zheng, Pan et al. 2009). It mainly targets the ester bonds and glycosidic side chains to impart

reduction in cellulose degree of polymerization, crystallinity, cause swelling and partial dissolution of hemicellulose and lignin (Brodeur, Yau et al. 2011). Contrary to other pretreatment methods, alkali processes are usually performed at lower temperature and pressure for a long period of time, ranging from hours to days (Mosier, Wyman et al. 2005).

Lime was recognized as one of the most promising biomass pretreatment techniques by various publications (Mosier, Wyman et al. 2005, Wyman, Dale et al. 2005, Yang and Wyman 2008), owing to its low cost, mild pretreatment conditions and high efficiency. It has been reported to effectively pretreat corn stover (Kaar and Holtzapple 2000), wheat straw (Chang, Nagwani et al. 1998) and switchgrass (Chang, Burr et al. 1997). When it comes to higher lignin content substrates such as woody biomass and newspaper, oxygen can be utilized with lime pretreatment to yield oxidative lime process to promote lignin removal and ensure high cellulose digestibility (Chang, Nagwani et al. 2001).

Figure 5 summarizes all the chemical pretreatment techniques discussed in this chapter.

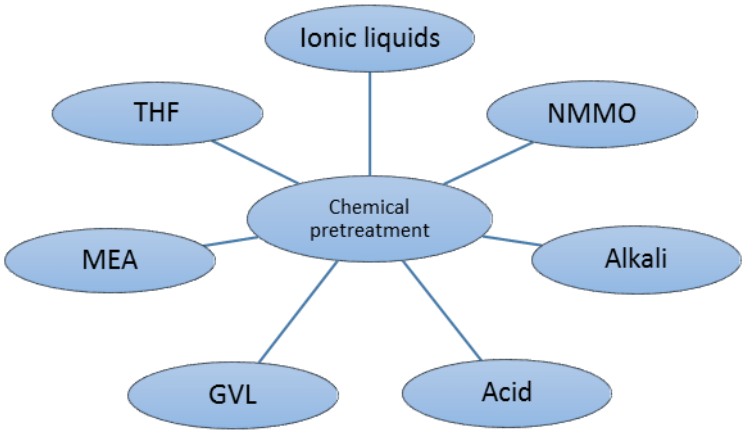


Figure 5. Various chemical pretreatment technologies discussed in this chapter

1.3.4 Biological pretreatment

White- and brown-rot fungi

Typical biological pretreatment technologies utilize various wood-degrading rot fungi to remove certain wood components in order to enhance the accessibility of cellulose for the following steps. It does not require expensive equipment or intensive energy inputs. The most well-known biological pretreatment agents are white- and brown-rot fungi (Worrall, Anagnost et al. 1997). **Table 2** shows examples of the two types of fungi studied in published work.

The brown-rot fungi were found to mainly attack structural carbohydrates with limited lignin degradation (Kim and Newman 1995), due to non-existence of lignin degrading enzymes (Goodell, Nicholas et al. 2003). Two types of white-rot fungi were reported: selective and non-selective. The selective white-rot fungi primarily degrade

lignin and hemicellulose, exposing cellulose microfibrils (Pandey and Pitman 2003); to the contrary, non-selective fungi simultaneously digest all three major wood components at similar rates to generate homogeneously degraded plant cell wall (Blanchette, Otjen et al. 1985, Pandey and Pitman 2003). The lignin degradation capacity of white-rot fungi is a result of the production of lignin decay enzymes such as peroxidases and laccase from the microorganisms (Lee, Gwak et al. 2007). White-rot fungi decay efficiency was found to be dependent on the ratio of syringyl to guaiacyl (S:G ratio), cell types and cell wall layers (Blanchette 1995).

To date, the most studied wood-decay microorganisms belong to white-rot fungi and they are thought to be the most promising biological pretreatment technique due to their lignin degrading activity (Zheng, Pan et al. 2009). However, the disadvantages of biological pretreatment technique is very evident: i) the slow digestion rate; ii) the requirement for growth condition control; iii) the low selectivity (Lee, Gwak et al. 2007). Biological pretreatment by itself is less commercially attractive, recently, researchers are trying to combine it with other pretreatment techniques to reduce the overall energy input, while no large scale application is reported so far.

Table 2. White- and brown-rot fungi studied in early publications.

	Microorganism
White-rot fungi	*Phanerochaete chrysosporium, Coriolus versicolor, Ceriporiopsis subvermispora, Echinodontium taxodii, Trametes versicolor*
Brown-rot fungi	*Coniophora puteana, L. cyathiformis*

1.4 Summary of biomass pretreatment techniques

Lignocellulosic biomass is the most abundant renewable carbon-rich resource on the planet. It has great potential to be utilized as a promising feedstock to produce liquid transportation fuels and reduce the world's reliance on fossil fuels. The biofuel research is a multi-disciplinary field currently conducted through the collaboration between academia, industry and government. The U.S. Department of Energy has been supporting three Bioenergy research centers since 2007: BioEnergy Research Center (BESC), Great Lakes Bioenergy Research Center (GLBRC) and Joint BioEnergy Institute (JBEI).

Although significant research progress for the production of lignocelluosic biofuel has been made, there are still several technical and economic challenges to address. Biomass pretreatment is a key unit operation to overcome its recalcitrant nature and impart the feedstock amenability to subsequent processes. To date, only a few pretreatment technologies have been brought to demonstration scale. None of the current pretreatment techniques could be considered as the perfect process for biofuel production. **Table 3** summarizes the pretreatment conditions, outcomes and disadvantages of the various techniques introduced in this chapter.

In addition to the existing issues regarding current pretreatment routes, it is tough to compare the pretreatment results from different research groups based on published data. Since a wide range of variables could change, such as the pretreatment condition, feedstock, characterization method, analytical method and data interpretation, it is almost impossible to evaluate a newly developed pretreatment technique by direct comparison with existing data. Comparative studies are often carried out to aid in the selection of

appropriate pretreatment for specific biomass (Wyman, Dale et al. 2005, Lau, Gunawan et al. 2009, Zhu, Sathitsuksanoh et al. 2009, Li, Knierim et al. 2010).

To further advance biomass pretreatment techniques, attention should be given to at least the following aspects: i) exploration of potentially commercially feasible new techniques ii) optimization of existing leading techniques toward large scale application iii) improve fundamental understanding of the biomass cell wall structure to enable rationally design and development of pretreatment strategies.

Table 3. Comparison of biomass pretreatment techniques.

Pretreatment technique	**Pretreatment condition**	**Pretreatment outcome**	**Disadvantage**
Ball milling	Ambient temperature, hours to days	Decrystallization, accessibility	Energy-intensive
Steam explosion	160-240 ℃, < 1h, high pressure	Hemicellulose removal, accessibility	Sugar degradation
AFEX	< 90 ℃, < 30 min, high pressure	Decrystallization, accessibility	Substrate specific
Organic solvent	Solvent-dependent	Decrystallization, lignin removal	High solvent cost, Solvent recovery required
Dilute acid	130-210 ℃, minutes to hours	Hemicellulose removal, accessibility	Sugar degradation
Alkali	Moderate temperature, hours to days	Lignin removal, accessibility	Slow rate
White-rot fungi	Near ambient temperature, weeks to months	Lignin and hemicellulose removal	Slow rate, low selectivity
Brown-rot fungi	Near ambient temperature, weeks to months	Cellulose degradation	Slow rate, cellulose degradation

1.5 Dissertation map

This thesis work mainly focuses on the investigation of three different pretreatment techniques for biofuel applications. Chapter 2 studies the effect of SO_2-catalyzed steam explosion and its severity on the substrate characteristics. Chapter 3 discusses a non-conventional biological pretreatment technology utilizing a family 1 cellulose-binding domain and the protein engineering of this peptide. Chapter 4 describes the development of a novel chemical pretreatment method based on substituted imidazoles for both cellulose and lignocellulosic biomass. Chapter 5 introduces a cellulose accessibility evaluation technique with low molecular weight direct dyes and its potential to assess pretreatment effectiveness. Chapter 6 summarizes the main findings of previous chapters.

CHAPTER 2

PHYSICOCHEMICAL PRETREATMENT OF LIGNOCELLULOSIC BIOMASS

2.1 Introduction

Lignocellulosic biomass is generally agreed to be one of the most promising feedstocks for biofuel production (Alzate and Toro 2006). Lignocellulosics are the most abundant natural resource which can be hydrolyzed to sugars and further fermented to biofuel by various microbes. However, the three main components of lignocellulosics include fermentable hexose/pentose-based cellulose and hemicellulose, as well as lignin which is both a physical and a chemical barrier for the hydrolysis process. These three complex plant polymers are held together tightly by covalent and hydrogen bonds, along with van der Waals forces. The tight packing creates difficulties in the fractionation and degradation of the raw biomass (Overend and Chornet 1987, Himmel, Ding et al. 2007). One of the most critical problems associated with the production of second-generation biofuels is overcoming the recalcitrance of the natural-occurring plant polymer complex. Recalcitrance is a combined effect of several factors such as the presence of lignin and hemicellulose, reactivity, surface area, pore volume, accessibility and cellulose crystallinity (Zhang, Ding et al. 2007, Sierra, Smith et al. 2008). Substantial research effort has been devoted to developing effective and economically-viable pretreatment methods to disrupt the physical hindrance (such as hemicelluloses and lignin) and to enhance surface area and accessibility of cellulose to increase yield and attain a more

amenable feedstock (Mosier, Wyman et al. 2005). Steam explosion is one of the most commonly used pretreatment methods, as it is suitable for a variety of feedstocks, and is regarded as one of the most efficient processes (Kumar, Barrett et al. 2009, Zhu and Pan 2010). Steam explosion significantly lowers both environmental impact and capital investment (Cara, Ruiz et al. 2006); it can be operated in either batch or continuous manner as a one- or two-step process to fractionate biomass into its three main constituents (Soderstrom, Galbe et al. 2004).

Softwood is considered to be one of the worst-case scenarios as a feedstock for the bioconversion process due to its high lignin recalcitrance (Lu, Yang et al. 2002, Pan, Xie et al. 2005, Zhao, Wang et al. 2008). Its pretreatment at lower-severity conditions optimally liberates easily-hydrolysable hemicelluloses, but generates a solid residue that is not readily amenable to the hydrolysis by cellulase (Lu, Yang et al. 2002). Many studies have focused on the optimization of these pretreatment conditions for various feedstocks (Boussaid, Robinson et al. 1999, Bura, Bothast et al. 2003, Cara, Ruiz et al. 2008, Carrasco, Baudel et al. 2011) to maximize sugar recovery while comparatively less work has been performed to elucidate the effects of severity on the characteristics of biomass and subsequent enzymatic hydrolysis (Boussaid, Robinson et al. 1999, Ibrahim, Agblevor et al. 2010, Kumar, Chandra et al. 2011).

This chapter focuses on the effect of severity on digestibility, accessibility, and crystallinity of loblolly pine (LP), the dominant tree species in the Southeastern United States. A simple analysis method was utilized to accurately and quickly determine the CrI of both untreated and steam-exploded lignocellulosic biomass based on X-ray diffraction (XRD) spectra. The goal of this study is to determine how the severity of steam explosion

pretreatment can affect the digestibility and accessibility of pretreated lignocellulosics as well as the interaction between cellulose and lignin. Furthermore, the significance of CrI in the enzymatic hydrolysis of pretreated lignocellulosics will be elucidated.

2.2 Materials and Methods

Materials

Avicel PH-101 (11365) purchased from Sigma-Aldrich (St Louis, MO, USA) was employed as standard crystalline cellulose in this study. Three differently-pretreated SELP samples were generously provided by Dr. John Muzzy from the Georgia Institute of Technology, Atlanta, GA, USA. Cellulase from *Trichoderma reesei* (Celluclast®, 159 FPU/ml) and β-glucosidase from almonds (5 U/mg) were purchased from Sigma-Aldrich and phosphoric acid (85%) was from EMD (Gibbstown, NJ, USA). The BCA protein assay kit was purchased from Thermo Fischer Scientific (Rockford, IL, USA).

Sample preparation

The acidic SELP samples as received were extensively washed with DI water and sodium acetate (NaOAc) buffer (50mM, pH 5.0) until pH 5.0 was reached. Lyophilization at -55℃ and 0.05 mbar was used to dry the washed samples overnight, and the absence of soluble reducing sugars in lyophilized samples was confirmed by dinitrosalicylic acid (DNS) assay.

Enzymatic hydrolysis

Batch hydrolysis reactions were carried out in 1.5 ml eppendorf tubes with 2% substrate consistency. Substrates were pre-incubated for 1 h in NaOAc buffer (50 mM, pH 5.0) at 45℃ with stirring. Following pre-incubation, enzymatic hydrolysis was

carried out for a desired time with varying enzyme loadings (0.0048 – 7.6 FPU/ml) at the same conditions. β-Glucosidase was added at 15 kU/L. The total reaction volume was 1 ml unless otherwise stated. At the desired time point, the tubes containing reaction mixture were immediately chilled on ice and centrifuged at 14,000 rpm for 3 min. The reducing sugar concentration in the supernatant was measured by DNS assay with glucose standards. All the enzymatic hydrolyses were carried out in duplicate. Initial rate was defined as the average glucose yield for first 10 min of enzymatic hydrolysis.

Steam explosion

The LP sample was mixed with the desired concentration of gaseous SO_2 (2.9%, 3% or 5%) for at least 1 hour in a sealed plastic bag. The material was then pretreated with steam in a 10-L pretreatment reactor at the desired temperature (214 ℃, 215 ℃ or 224 ℃) for a specified residence time (5 or 10 min). Detailed steam explosion conditions utilized in this work are tabulated in **Table 4**. After steam pretreatment, a pneumatic ball valve was opened rapidly and the contents were explosively discharged into a 100-L accumulator at atmospheric pressure. Solid residue was collected from the accumulator and stored at -20 ℃ for further analysis.

Adsorption studies

Substrates (2% consistency) were preincubated in 1 ml NaOAc buffer (50 mM, pH 5.0) at 45 ℃, 1000 rpm for 1 h followed by cooling on ice. Varying amounts (0.0048 – 7.6 FPU/ml) of cellulase were added and further agitated for 30 min at 4 ℃ to minimize the influence from hydrolysis. Following 30 min adsorption, all samples were immediately centrifuged for 3 min at 4 ℃ and 18,625 g.

Protein assay

A modified BCA assay with precipitation was utilized to determine protein concentrations in the supernatant after cellulase adsorption in order to eliminate the interference with both carbohydrates and soluble lignin. Control experiments were performed lacking substrate. This procedure has been previously published by R. Brown et al (Brown, Jarvis et al. 1989). The adsorbed (bound) cellulase was calculated as the total enzyme added minus free enzyme in the supernatant.

X-ray diffraction (XRD)

XRD patterns of lyophilized samples were recorded with a X'pert PRO X-ray diffractometer (PANanalytical BV, Almelo, the Netherlands) using Cu/Kα_1 irradiation (1.54 Å) at 45 kV and 40 mA. The scattering angle (2θ) ranged from 10 ° to 40 ° with a scan speed 0.021425 s^{-1} and step size 0.0167 °.

Phosphoric acid (PA) treatment

PA treatment of SELP was carried out following the procedure described by M. Hall et al (Hall, Bansal et al. 2010). Thirty mL of diluted ice-cold PA was used to pretreat 1 g slightly hydrated SELP 2 for 40 min with occational stirring. Thirty ml of ice-cold acetone was used for regeneration of cellulose. The resulting solid was washed twice with 30 ml ice-cold acetone, and four times with 100 ml DI water. The washed solid fraction was used for initial rate experiments directly, and the lyophilized powder was prepared for XRD measurement. The acid-insoluble lignin (AIL) content in acid-treated SELPs was measured by a two-stage acid hydrolysis following a published procedure (Foston, Hubbell et al. 2011). In brief, the 1st stage utilizes severe pH at low temperature (72 vol.% H_2SO_4 at 30 ℃ for 1 h) followed by a 2nd stage at low acid concentration and higher temperature (3 vol.% H_2SO_4 at 121 ℃ for 1 h). After cooling to room temperature,

the mixture was filtered using 934-AH glass microfiber filter (Whatman, PA, USA) and the insoluble solids were dried at 105 ± 3 ℃ in an oven overnight. The insoluble solids were considered as AIL.

2.3 Results and Discussion

2.3.1 Hydrolysis of SELPs

Steam explosion is considered to be one of the most effective and relatively low-cost pretreatment methods for hardwood as well as softwood. The additional step of SO_2 impregnation aims for an improved recovery of both cellulose and hemicelluloses (Taherzadeh and Karimi 2008). After pretreatment, the solid biomass residue contains most of cellulose and lignin, while the majority of hemicellulose is degraded and solubilized into the liquid biomass fraction.

In this work, the effects of steam explosion on the solid residue are investigated. Three SELPs under different pretreatment severities were extensively washed with DI water and sodium acetate buffer to adjust the pH value to 5.0 and remove soluble substances and inhibitors. This washing will remove any soluble sugars and degradation products from the solid sample and hence the amount of cellulose that remains in each of the solid samples will be different. The samples were designated as SELP 1, SELP 2 and SELP 3 in order of ascending severity. The pretreatment conditions and resulting severities are shown in **Table 4**. The severity index (R) was calculated according to the following empirical Equation 1 (Nguyen, Tucker et al. 2000):

$$R = \log_{10}\left(te^{\frac{T-100}{14.75}}\right) - pH \quad (1)$$

where t is residence time (min) and T is temperature (℃).

Table 4. Steam explosion conditions and corresponding severities for the SELP samples utilized in this study.

	SO_2 (%)	Temperature T (℃)	Residence time t (min)	Severity R
SELP 1	2.9	215	5	3.1
SELP 2	5	214	5	3.2
SELP 3	3	224	10	3.6

The compositional analysis results are shown in **Table 5**. Compared to non-pretreated LP, SELPs shown negligible hemicelluloses content. The most severely steam-exploded SELP 3 has the lowest glucan content, which confirmed a decrease in steam explosion selectivity as severity increases. It has previously been shown that during the steam explosion process more cellulose is nonspecifically hydrolyzed and solubilized into the liquid fraction with increasing severity, thus increasing the corresponding percentage of lignin in the solid residue. After normalization to total cellulose content of the solid fraction (**Figure 6**), SELP 3 shows the highest glucose yield among the three samples, despite containing a much lower beginning cellulose content. Solubilized sugar concentrations in hydrolysates of SELP 1 and 2 are similar throughout the 42 hour period. The sugar concentration reached saturation at 18 h for SELP 1 and 2, while the saturation

was achieved much earlier for SELP 3 at around 2 h. Thus, there is a compromise between the pretreatment selectivity, which measures the relative removal of hemicelluloses and cellulose during steam explosion, and resulting cellulose hydrolysability. High steam explosion severity results in a more degradable cellulose fraction, while the high selectivity of SO_2-catalyzed acid hydrolysis during pretreatment is not maintained (Ramos 2003). In addition to the decreased cellulose concentration, the presence of lignin will inhibit the enzymatic reaction by two mechanisms: forming a physical barrier which reduces the accessibility of cellulose and non-productive binding with cellulosic enzymes to impair their efficiency (Zhu, Pan et al. 2010). Thus, the combined, but countervailing, effects of increased lignin content and enhanced cellulose accessibility determine the hydrolysis rate and extent of differently steam-pretreated LP.

Table 5. Composition of untreated LP and SELP under different conditions

	Arabinan wt.%	**Galactan wt.%**	**Glucan wt.%**	**Xylan wt.%**	**Mannan wt.%**	**AIL wt.%**
LP	1.01	2.34	40.54	5.85	8.70	30.61
SELP 1	<0.01	<0.01	47.62	0.46	<0.01	49.99
SELP 2	<0.01	0.08	42.21	0.42	0.76	52.85
SELP 3	<0.01	0.06	18.64	<0.01	<0.01	76.58

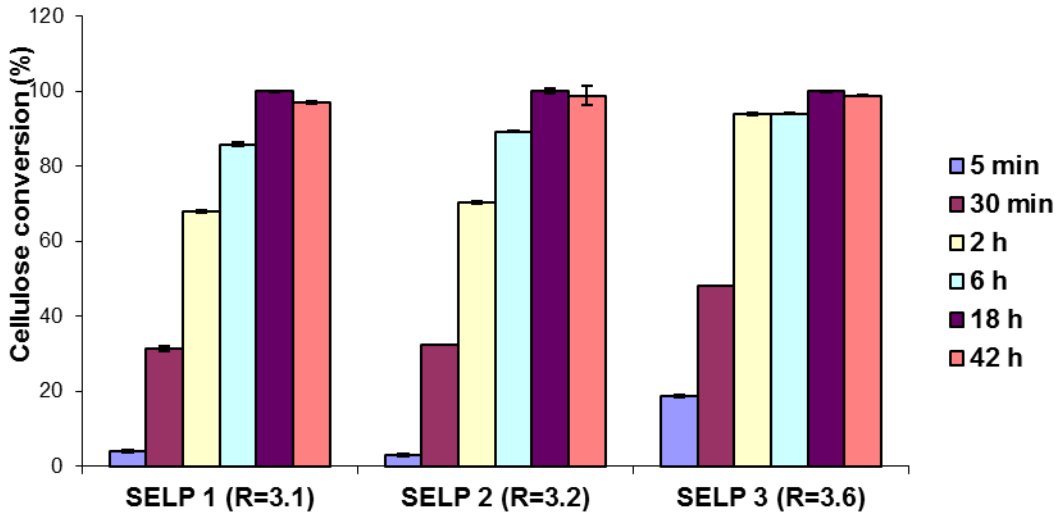


Figure 6. Comparison of enzymatic hydrolysis efficiency of three SELP samples.

2.3.2 Adsorption and initial rate studies

Due to the heterogeneous nature of cellulase action, adsorption is a prerequisite step for the catalytic reaction to occur. The adsorption isotherm was obtained at 4^{o}C to minimize carbohydrate degradation during adsorption. Carbohydrate degradation reduces the number of available active binding sites and introduces error in adsorption studies (Hall, Bansal et al. 2010). The adsorption of cellulolytic enzyme onto substrates is an effective process which usually requires 2-5 min to achieve equilibrium; 30 min was allowed in this study (Dijkerman, Vervuren et al. 1996, Tsao 1999). **Figure 7** shows the SELPs adsorption isotherms. The substrate accessibility can be assessed by the adsorbed cellulase (the difference between total enzyme loaded and free enzyme remaining in supernatant) (Bansal, Vowell et al. 2012). The experimental adsorption data was fitted to Langmuir isotherm having the following form:

$$B = B_{max} \frac{K_{ad}\, C}{1 + K_{ad}\, C} \quad (2)$$

where B is the adsorbed cellulase amount (μg/mg substrate), B_{max} is the maximum adsorption amount (μg/mg substrate), K_{ad} is adsorption equilibrium constant (L/g), and C is free cellulase concentration (g/L).

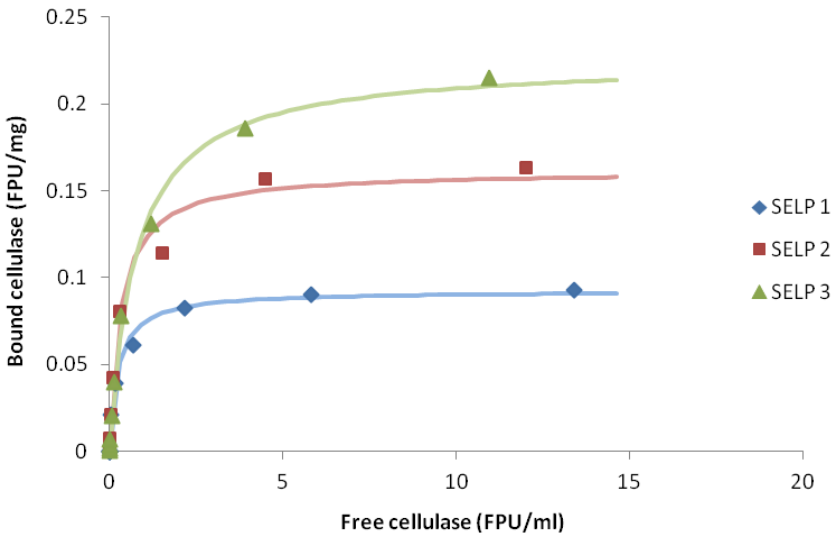


Figure 7. Adsorption isotherm for three SELP samples. The adsorption experiments were conducted at 4℃ for 30 min to minimize impact from hydrolysis. The curves are a nonlinear fit of experimental data to a Langmuir isotherm.

All three SELPs have similar affinity to cellulase enzymes at low enzyme loadings (< 1.0 FPU/ml); noticeable differences begin to appear beyond 1.0 FPU/ml. From the nonlinear fitting results, cellulase adsorption on all SELPs tested follows a Langmuir isotherm with R^2 = 0.971 or higher. The parameters B_{max} and K_{ad} determined

from each Langmuir isotherm are listed in **Table 6**. Among the three samples, SELP 3 showed the highest B_{max} value of 184.4 mg/g, which indicates the highest accessibility to enzyme. Lu et al. studied the adsorption behavior of steam-exploded Douglas-fir, which can also be fitted to a Langmuir isotherm, and found the maximum adsorption capacity to be 171.3 mg/g substrate (Lu, Yang et al. 2002). The relative magnitude of accessibility of all three substrates investigated are SELP 3 > SELP 2 > SELP 1. The accessibility of substrates is positively correlated with the pretreatment severity. The adsorption equilibrium constants K_{ad} (**Table 6**) showed a decreasing trend as pretreatment severity increases, which indicates that cellulolytic enzymes are more readily adsorbed on less severely pretreated samples.

Table 6. Adsorption parameters determined from Langmuir isotherm

	R^2 (%)	B_{max} (μg/mg)	K_{ad} (ml/mg)
SELP 1	99.6	75.7	0.005
SELP 2	97.1	132.3	0.0037
SELP 3	99.6	184.4	0.0016

The binding capacity B_{max} is a reflection of the accessible substrate surface which includes cellulose, lignin and trace amounts of hemicellulose. Binding capacity can be ascribed to cellulase adsorption on both cellulose and lignin, the latter of which has been reported to have a high affinity toward cellulolytic enzymes (Tu, Pan et al. 2009, Nakagame, Chandra et al. 2011). There is an apparent difference in the adsorption capacity of SELP 1 and 2, which likely resulted from the different SO_2 concentrations

used for LP impregnation prior to pretreatment. The pretreatment conditions for the two SELPs differ in both SO_2% and temperature. The significant difference in adsorption behavior can likely be attributed to different SO_2% as temperature varies by only 1°C. While SELP 1 and 2 show similar compositions (**Table 5**), we surmise that higher SO_2% could increase both the mobility of lignin, leading to potential redistribution within the lignocellulosic structure, and higher affinity toward cellulolytic enzymes (Excoffier, Toussaint et al. 1991). No obvious difference in adsorption will be observed unless high enzyme loadings are employed, at which point the available active sites on the substrate become limiting.

The subsequent initial rate experiments result in plots of initial rate as a function of total adsorption capacity (**Figure 8**). A wide range of enzyme concentrations (0.0048 – 7.6 FPU/ml) was tested and a linear relationship was found for SELP 1 throughout the range. As for SELP 2 and 3, a "phase transition" was observed, which is indicative of a substrate reactivity change over the enzyme concentration range. At low enzyme loadings, the slope of SELP 3 is the highest, followed by SELP 2; this phenomenon is probably due to the difference in cellulose morphology after steam explosion. On the contrary, at higher enzyme loadings, the slope order is SELP 1 > SELP 2 > SELP 3 (0.1361, 0.0575, 0.0504, respectively); the reversed order can be explained by the limited number of available binding sites on reactive fraction of the substrate as enzyme loading increases. As a result, an increasing portion of bound cellulase ended up non-productively bound onto lignin. The slope is an indicator of the reactivity of substrate; a smaller slope implies more cellulase adsorption is required to obtain the same initial rate at the enzyme concentrations investigated. Although both SELP 1 and SELP 2 show similar initial rates,

the substrate affinity is markedly different. SELP 2 is more accessible to cellulase compared to SELP 1, while the cellulase adsorption on SELP 1 is much more productive than SELP 2 due to a similar sugar yield despite the difference in adsorption capacity. The difference in initial rate between SELP 3 and the other two SELPs is more evident at relatively low enzyme loadings because of its exceptional high reactivity at extremely low enzyme concentration. Significant lignin and/or cellulose structural change, which resulted in distinct adsorption behavior, must have occurred for the three SELPs during the process of steam explosion.

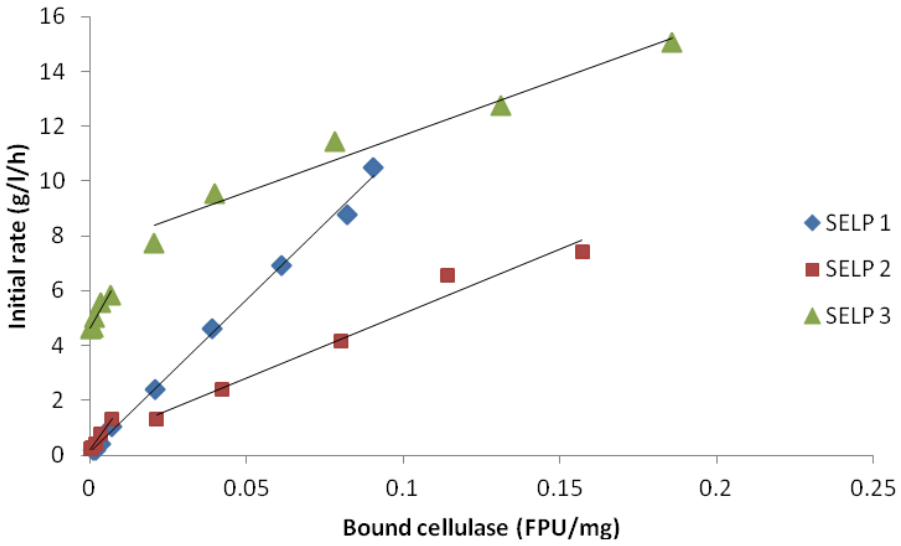


Figure 8. The dependence of initial rate on cellulose adsorption for enzyme loadings ranging from 0.0048 to 7.6 FPU/ml.

2.3.3 CrI measurement of SELPs

The determination of CrI for the cellulose fraction in pretreated wood samples can be tricky because of the existence of many interfering factors such as the presence of

lignin, hemicellulose, residual extractives, and the interaction between them. To quantitatively determine CrI, a simple mathematical method, least-squares estimation (LSE), was utilized to estimate CrI with reasonable accuracy for complicated lignocellulosic substrates.

Due to the ease of use, about 70-85% published studies calculated the CrI of cellulose/lignocellulosics based on the peak height method; this method was developed in 1962 by Segal et al. as an empirical method to obtain the CrI from an XRD pattern (Segal, Creely et al. 1962, Park, Baker et al. 2010). The peak height method is a semi-quantitative CrI measurement based on the reflected intensities of the 200 phase and amorphous cellulose ($2\theta = 18°$) with Equation 3. It is a convenient time-efficient method that is useful for the measure of relative crystallinity.

$$CrI\% = \frac{I_{200}}{I_{200}+I_{am}} \times 100\% \quad (3)$$

However, the peak height method tends to overestimate the crystallinity of varies cellulose-containing substrates (Park, Baker et al. 2010). To attain a more quantitative measure of CrI, the entire XRD spectrum needs to be taken into account to minimize loss of information. In this work, the normalized spectrum of a sample is expressed as a linear combination of the normalized spectra of three standards: crystalline cellulose, amorphous cellulose, and lignin (Andersson, Serimaa et al. 2003, Bansal, Hall et al. 2010). Since a large portion of the easily-hydrolysable hemicellulose was degraded and

dissolved by steam explosion and subsequent washing steps, the remaining trace amount was ignored in the calculation of CrI. Equation 4 shows the linear combination:

$$I(2\theta) = f_a I_a(2\theta) + f_l I_l(2\theta) + f_c I_c(2\theta) + \varepsilon \quad (4)$$

$$f_a + f_l + f_c = 1 \quad (5)$$

where $I_a(2\theta)$ is the intensity of amorphous cellulose at 2θ, $I_l(2\theta)$ is the intensity of lignin at 2θ, $I_c(2\theta)$ is the intensity of crystalline cellulose at 2θ, f_a is the contribution of amorphous cellulose to the spectra, f_l is the contribution of lignin, f_c is the contribution of crystalline cellulose, and ε is random error. The summation of the above three factors equals 1, as it is assumed the SELP samples are mainly comprised of cellulose and lignin (Equation 4).

Model crystalline cellulose (Avicel), amorphous cellulose (phosphoric acid swollen cellulose or PASC), and lignin are used as standards. LSE was performed to estimate the two parameters f_a and f_l with the spectrum of the sample and the standards. The accuracy of the curve-fitting was evaluated via R^2; comparing how well the original spectra are likely to be predicted by the reconstructed spectra (**Table 8**). Since neither the amorphous cellulose nor lignin contains crystalline structure, the CrI was calculated by Equation (6):

$$CrI = (1 - f_a - f_l)CrI_c \quad (6)$$

where CrI_c is the known CrI of Avicel which was determined by both ^{13}C-NMR and XRD to be 60% (Hall, Bansal et al. 2010).

To validate the newly developed CrI measurement method, three mechanical mixtures of microcrystalline cellulose (Avicel) and amorphous lignin with various ratios were prepared. **Figure 9** – **Figure 11** shown the normalized original and reconstructed diffraction patterns for the different mixtures. As the Avicel to lignin ratio increases from 1:4 to 4:1, the diffracted intensity of the characteristic peak around 22.5° which correspond to the 200 phase of cellulose I increases as expected. The overall CrI determined with the LSE-based curve-fitting method increased from 12.7% to 47% (**Table 7**). The crystallinity of the cellulose fraction (CrI_{cel}, **Table 7**) can be calculated based on the linear combination assumption with Equation 7. The resulting CrI_{cel} for all three mixtures agrees well with the crystallinity of Avicel (i.e. 60%).

$$CrI_{cel} = \frac{CrI}{F_c} \tag{7}$$

where F_c is the percentage content of cellulose in the sample.

Table 7. The overall CrI and CrI_{cel} of mechanically mixed samples with different Avicel to lignin ratio

Avicel:lignin	1:4	1:1	4:1
Overall CrI (%)	12.7	29.9	47.0
CrI_{cel} (%)	63.5	59.8	58.8

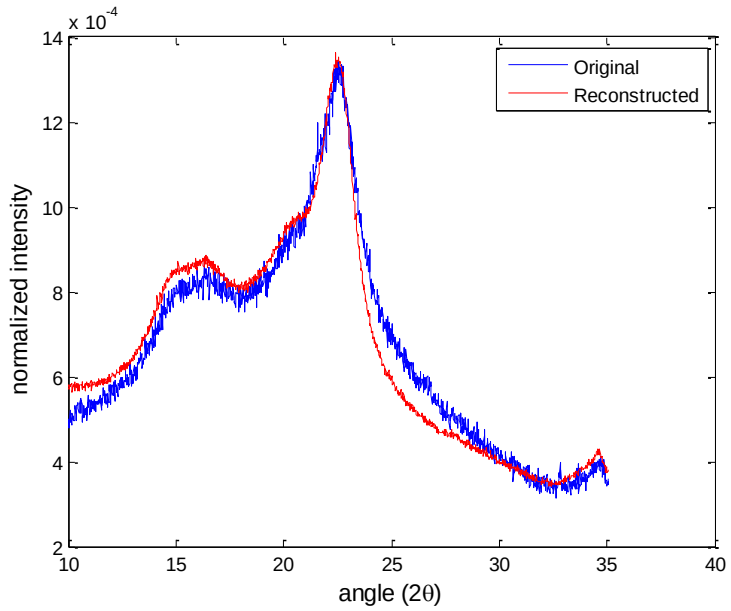


Figure 9. Reconstructed and original XRD spectrum of the mechanical mixture Avicel:lignin = 1:4.

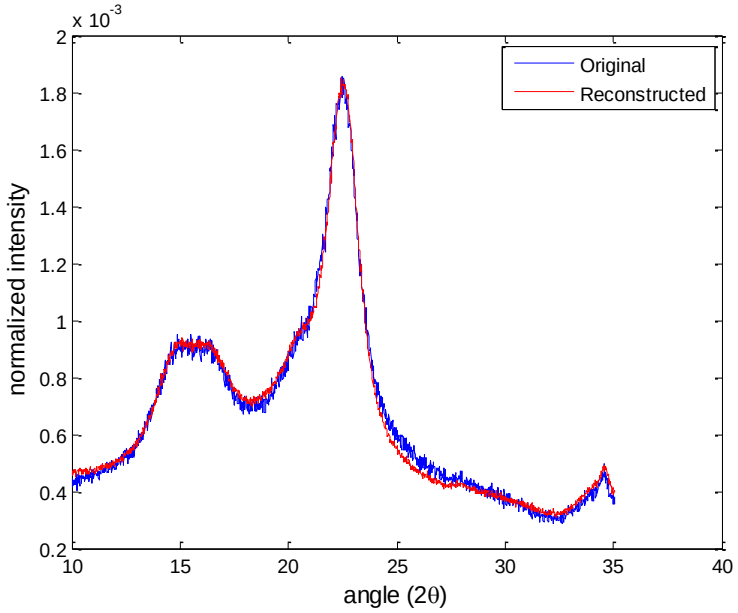


Figure 10. Reconstructed and original XRD spectrum of the mechanical mixture Avicel:lignin = 1:1.

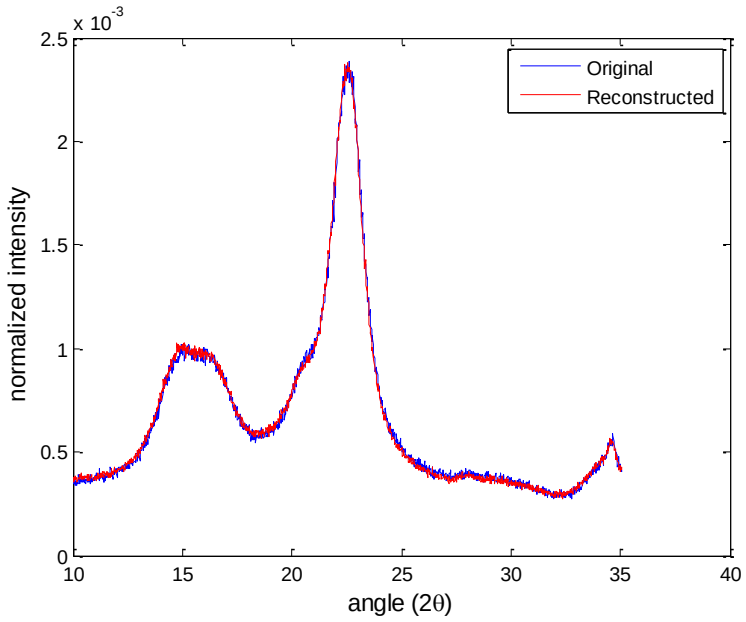


Figure 11. Reconstructed and original XRD spectrum of the mechanical mixture Avicel:lignin = 4:1.

Table 8. Estimated CrI, AIL and CrI_{cel} for untreated LP and SELPs

	CrI (%)	**R^2 (%)**	**AIL (%)**	**CrI_{cel} (%)**
LP	26.2 ± 1.8	0.98	30.61	64.6 ± 4.4
SELP 1	41.9 ± 1.9	0.98	49.99	88.0 ± 4.0
SELP 2	31.2 ± 2.0	0.97	52.85	73.9 ± 4.7
SELP 3	12.9 ± 3.1	0.89	76.58	69.2 ± 16.6

The curve fitting results in **Table 8** lead to the conclusion that the overall CrI of SELP 1, 2 and 3 decreases with rising pretreatment severity, but is increased in comparison to the untreated wood. The decrease in CrI is a combined result of the removal of amorphous components and the structural change of cellulose crystal.

The 95% confidence interval for each CrI value is also indicated in **Table 8**. The change in CrI_{cel} shows the same trend as the overall CrI of SELPs. Based on the compositional analysis result in **Table 5**, the untreated LP contains 40.54% cellulose, which is very close to the published result in earlier work (Sannigrahi, Miller et al. 2010); the CrI_{cel} was then calculated using Equation 7 to be 64.6% ± 4.4%. The CrI_{cel} of untreated LP obtained with the curve-fitting method developed in this work agrees well with previous study which measured the CrI of extracted LP cellulose to be 62.5% with ^{13}C-NMR (Sannigrahi, Ragauskas et al. 2008). Compared to untreated LP, SELP 1 (the lowest severity) showed an increase in the cellulose CrI, and thus also in CrI_{cel}. Further increases in severity resulted in a decrease of cellulose CrI and CrI_{cel} in LP samples. The change in CrI from untreated LP to SELP 1 is in agreement with published results (Yamashiki, Matsui et al. 1990, Cherian, Leao et al. 2010) and the increase can be ascribed to the preferential hydrolysis of amorphous cellulose during steam explosion (Wang, Jiang et al. 2009).

The finding that further increase in severity generates a LP sample with more amorphous-like cellulose is somewhat surprising. We postulate that the decrystallization effect of steam explosion occurs at a certain severity and may be accompanied by nonselective carbohydrate degradation. At low severity, the increase in CrI is mainly due to the hydrolysis of the amorphous fraction including hemicelluloses and possibly a small amount of amorphous cellulose. The dramatic increase in CrI_{cel} from untreated LP to SELP 1 (**Table 8**) is mainly a result of the removal of easily acid-hydrolysable amorphous cellulose. As severity further increases, the effect of steam explosion will not be limited to amorphous region but also extend to the crystalline portion. The

decrystallization effect of steam explosion can be clearly demonstrated by comparing the calculated CrI_{cel} of the SELPs in **Table 8**, where a decreasing trend was observed.

Some researchers isolated cellulose from steam-exploded agricultural residues by alkali extraction and bleaching and concluded CrI increases as severity increases (Ibrahim, Agblevor et al. 2010). In addition, the effect of steam explosion on pure cellulose isolated from wood and cotton was studied, and the same trend was confirmed (Yamashiki, Matsui et al. 1990). The effect of steam explosion on isolated pure cellulose and cellulose which is embedded in a complex plant polymer matrix will likely differ; moreover, previously published work (Yamashiki, Matsui et al. 1990) did not investigate a severity high enough to be comparable with this study's SELP 3. In assessing the actual impact of steam explosion on substrate CrI, the possibility that additional isolation/extraction steps introduce further change in CrI and/or cellulose crystal structure requires caution when choosing extraction conditions. By studying steam explosion severity from low to extremely high levels and avoiding additional isolation steps of biomass constituents, the true effect of steam explosion itself can be revealed. The results from this present work provide evidence that there might exist a critical severity point below which the CrI_{cel} shows positive correlation with severity while the opposite is true when exceeding this point.

2.3.4 Phosphoric acid treatment of SELP

The CrI of pure cellulosic substrate (e.g. Avicel) is a crucial factor in the hydrolysis process (Fan, Lee et al. 1980, Taherzadeh and Karimi 2008, Hall, Bansal et al. 2010). To investigate the role of CrI in the hydrolysis rate of pretreated lignocellulosics

rather than Avicel, the interference from compositional difference needs to be ruled out. Thus, phosphoric acid (PA) treatment was performed on SELP 2, and its crystalline structure was analyzed via X-ray diffractometry before and after PA treatment. PA has been used as a standard method to create amorphous cellulose with a high level of accessibility to enzymes (Den Haan, Rose et al. 2007). The XRD spectra of PA-treated SELP 2 (PASELP) are shown in **Figure 12**. A series of PA concentrations ranging from 76% to 82% was employed to generate PASELPs with different CrI, calculated by the above-mentioned curve fitting method. It is observed from **Figure 12** that PA preferentially attacks the crystal planes 200 and 110; a sufficiently concentrated PA can completely disrupt those two planes resulting in a completely amorphous SELP substrate. As PA concentration increases from 76% to 82%, the CrI of the resulting PASELP decreases linearly (diamond points in **Figure 13**). The CrI of PASELP is quite sensitive to the PA concentration change and only a narrow range has detectable impact on CrI. PA of a concentration lower than 76% is not capable of decrystallizing steam-pretreated lignocellulosics. The presence of critical PA concentrations was also reported by previous studies to be 75% and 83% for Avicel and industrial hemp, respectively (Moxley, Zhu et al. 2008, Hall, Bansal et al. 2010).

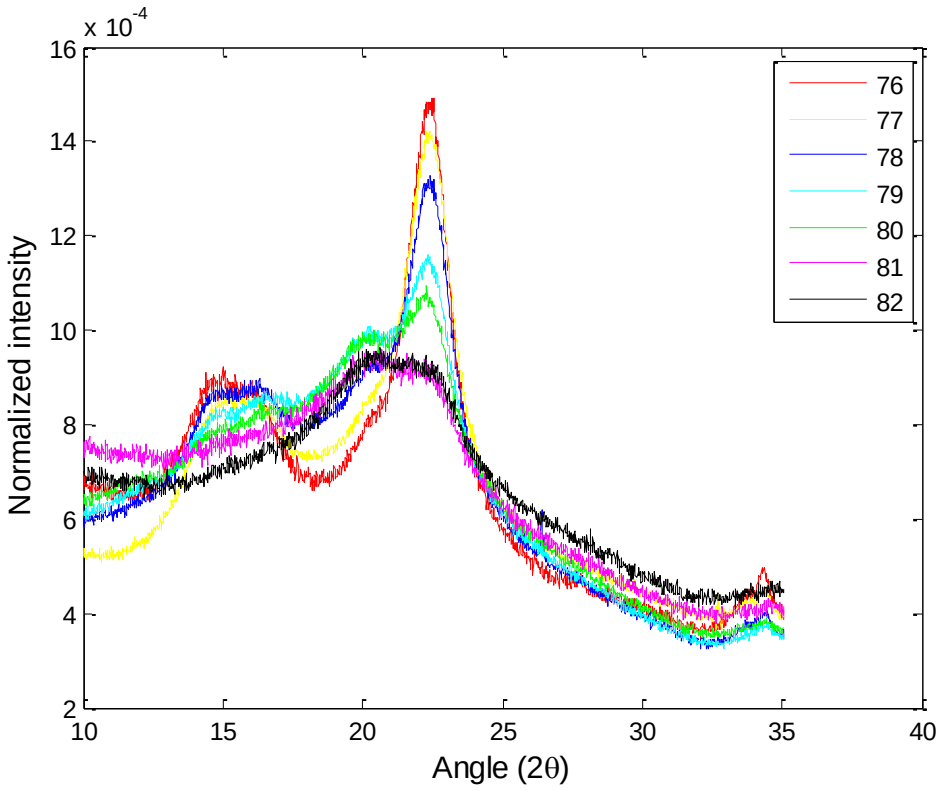


Figure 12. XRD patterns of PA-pretreated SELP 2. Pretreatment condition: 45 min with ice-cold PA concentration ranging from 76 – 82%.

Initial rate experiments were carried out to determine the hydrolysability of PASELP. A CrI change of ~30%-points increases the initial rate fivefold (square points in **Figure 13**) and a linear relationship was found between initial rate and CrI_{cel} (**Figure 14**). Although crystallinity is always coupled with other properties of cellulosic/lignocellulosic biomass (Park, Baker et al. 2010) such as accessibility (Bansal, Vowell et al. 2012), the calculation of CrI is independent of other properties and capable of excluding any interference from other factors. Previous studies have shown the PA treatment does not change the degree of polymerization (DP) significantly (Zhang and Lynd 2005, Jeoh, Ishizawa et al. 2007), which ruled out DP as the main factor affecting initial hydrolysis rate. Although acetone, which was used for PASELP preparation, is

known to dissolve lignin, the AIL measurement of PASELPs indicates the lignin contents in all samples decreased from 52.85% (AIL content of SELP2) to 42.5% ± 3.0%. No systematic trend was observed for AIL contents with respect to PA concentration. The CrI_{cel} of PASELPs were calculated according to the actual AIL contents in samples after PA treatment. The linear dependence of initial hydrolysis rates on independently calculated CrI_{cel} enables us to conclude that CrI plays an important role on the enzymatic hydrolysis rate of steam-pretreated lignocellulosics.

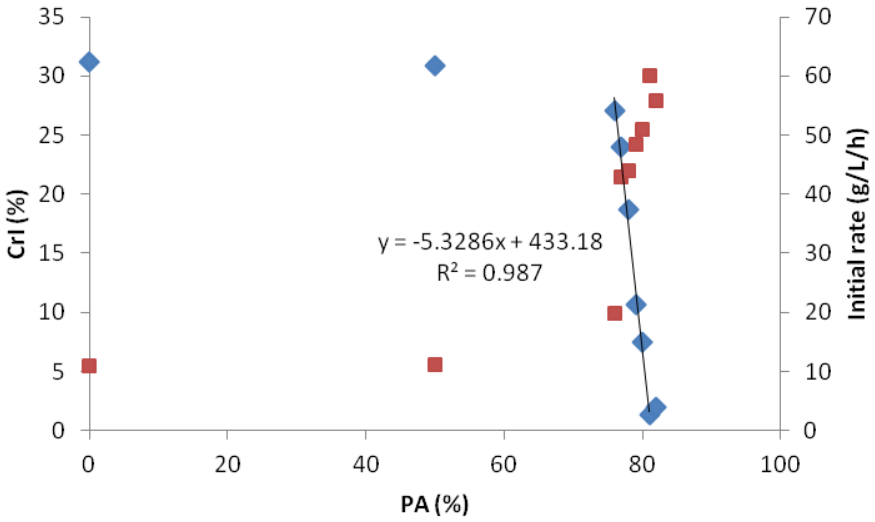


Figure 13. The effects of PA treatment on CrI and initial rate of PASELP. The diamond and square points show the dependence of CrI and initial rate on PA concentration respectively.

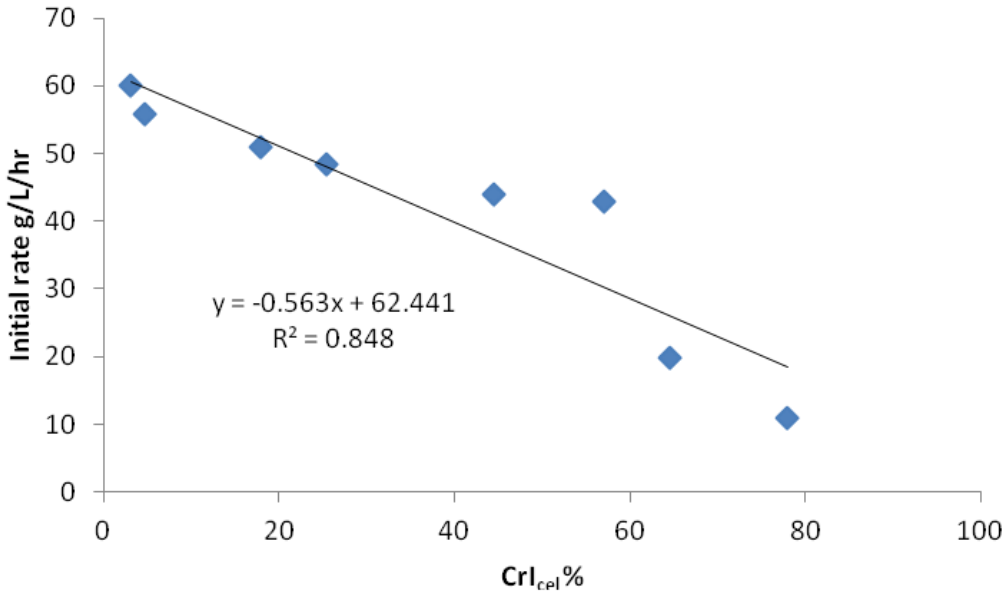


Figure 14. Initial hydrolysis rate of PASELPs as a function of the CrI_{cel}.

Furthermore, the correlation between initial rates and intensities was analyzed at all diffraction angles ranging from 10 ° to 35 ° in seven differently treated PASELPs. The correlation plot is shown in **Figure 15**. There is strong negative correlation (> 0.85) at diffraction angles around 110 and 200 phases, at 2θ = 15 ° and 22.5 ° respectively, which means the decrement of intensities in those regions has a positive contribution to the initial rate of samples. Although no physical meaning has been assigned to the interval between these two phases, the intensities were found to correlate positively with hydrolysis rates. This range of angles almost superimposes with that of the location of peaks of amorphous cellulose, thus, the finding that hydrolysis rate increases with increased amorphous fraction in samples can well explain the resulting correlation in this

region. The results shown in Figure 6 are in good agreement with previous findings for Avicel (Bansal, Hall et al. 2010).

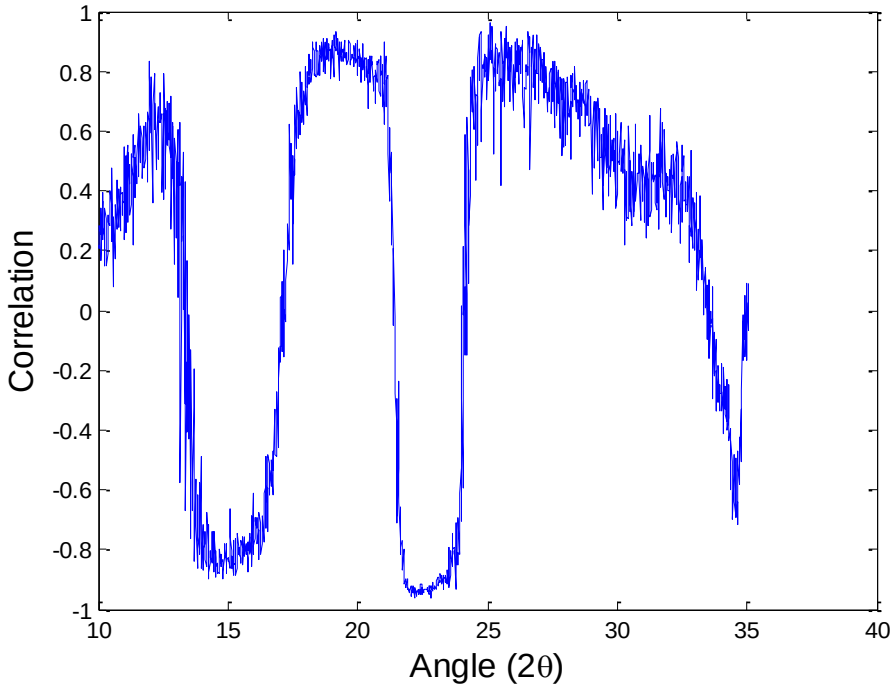


Figure 15. Correlation plot of initial hydrolysis rate and X-ray diffraction intensity at different diffraction angles for PASELPs.

2.4 Conclusions

Steam explosion at low severity preferentially removes the amorphous constituents of loblolly pine, however, further increases in severity generate a substrate with lower CrI; a combined effect of the removal of amorphous components and disruption of crystal structure. A critical severity point exists in steam explosion beyond which the degradation of crystalline cellulose dominates the process rather than the opening of a tight cellulose-lignin substructure. The significance of CrI in enzymatic hydrolysis of pretreated lignocellulosics was confirmed by a linear dependency of initial

rate on CrI. These findings lead us to believe that CrI is an indicator of hydrolysis rate and digestibility of pretreated lignocellulosics.

CHAPTER 3

CELLULOSE-BINDING DOMAIN – A NOVEL BIOLOGICAL PRETREATMENT AGENT FOR CELLULOSE

3.1 Introduction

The polysaccharide-recognizing domains, which mediate in the protein-carbohydrate interaction process, are defined as carbohydrate-binding modules (CBMs). Depending on the amino acid sequence similarity, CBMs can be classified into different families. As of September 2014, there are 71 CBM families based on the carbohydrate-active enzymes (CAZy) database (http://www.cazy.org/Carbohydrate-Binding-Modules.html). The CBMs display large diversity in their ligand specificity. The characterized CBMs are identified to bind to a variety of carbohydrates such as crystalline cellulose, amorphous cellulose, xylan, mannan, glucomannan, galactan, chitin, starch and lactose. Besides affinity binding, additional non-hydrolytic effects were identified and extensive mutational work has been carried out with a wide variety of CBMs to improve their functionalities.

Proximity and targeting effect: CBMs display no hydrolytic function and the widely accepted role of CBMs is their binding affinity to both soluble and insoluble carbohydrate substrates. Through addition or deletion of the CBM, ligand binding capacity can be altered. It has been reported that single-module polysaccharide hydrolases, such as cellobiohydrolase and mannanase, exhibit higher catalytic activity as well as thermostability when attached to heterologous CBMs (Hall, Rubin et al. 2011, Tang, Li et al. 2013, Thongekkaew, Ikeda et al. 2013, Voutilainen, Nurmi-Rantala et al.

2013). Upon fusion to other non-cellulolytic enzymes, the targeting effect of CBM can bring the chimera to within close proximity of the carbohydrate-containing substrate, which in turn improves the appended enzymes activity on their substrates in close association with carbohydrates. Ravalason et al. demonstrated that biobleaching efficiency can be drastically enhanced by forming a chimeric enzyme comprising laccase and family 1 CBM (Ravalason, Herpoel-Gimbert et al. 2009), through which the effective chimera concentration on insoluble lignin/cellulose surface can be enriched. However, higher ligand binding capacity/affinity does not guarantee an improved cellulase activity as the ultra-tight binding could become a restriction for the dynamic motion of the enzymes (Boraston, Bolam et al. 2004). Through the comparison study of a single- and double-CBM processive endoglucanase Cel9A, the same mutations introduced to the latter were found to be significantly less advantageous. This is due possibly to the existence of a strong wild-type binding domain in addition to the one being engineered (Li, Irwin et al. 2010). Besides the critical residues responsible for carbohydrate adsorption, the binding affinity of CBMs is modulated in part by the glycosylation pattern. Taylor et al. through computational investigation found a 140-fold increase in binding attributable to the addition of an artificial glycan on serine locates on the planar binding face of *Tr*Cel7A (Taylor, Talib et al. 2012). The improvement on binding and the ultimate effect on hydrolysis efficiency await experimental validation. Nonetheless, the molecular simulation results present an alternative direction for protein engineering of cellulases.

Non-productive binding to lignin: In addition to the positive ligand-recognition function, recently, the role of CBMs at high substrate consistency was investigated for

both cellulose and lignocelluloses. Even more than pure cellulose, the lignin-containing plant cell wall is highly recalcitrant and resistant to biological degradation. Cellulases deprived of CBMs exhibits similar or even higher hydrolysis yield at high substrate loadings compare to wild-type cellulase, especially for lignocelluloses. It was thus postulated CBMs partly contribute to non-productive binding to lignin (Le Costaouec, Pakarinen et al. 2013, Varnai, Siika-Aho et al. 2013). However, more direct evidences on CBMs' preferential adsorption onto lignin are needed to support the hypothesis and demonstrate the generic applicability for CBM-containing cellulases.

Processivity: The impact of CBMs on the mechanism of enzyme processivity is not completely understood (Li, Irwin et al. 2007). Among the recently published work, processivity is believed mainly to be controlled by the residues of the CDs (Igarashi, Koivula et al. 2009, Chiriac, Cadena et al. 2010, Li, Irwin et al. 2010). However, controversial results are also reported regarding the role of CBMs in processivity (Irwin, Shin et al. 1998). Recently, the CBM of a family 5 endoglucanases was identified to play a vital role in enzyme processivity through comparison with the CBM-less variant (Zheng and Ding 2013) and computational simulation results also provide evidence that family 1 CBMs are responsible for processivity (Beckham, Matthews et al. 2010). The *Tr*Cel7A CBD exhibits thermodynamically stable regions correspond to the length of a cellobiose unit, which is a critical length scale of the processive *Tr*Cel7A (Beckham, Matthews et al. 2010). Depending on the different modular structures of cellulases, the mode of processivity can be enzyme-specific. Nevertheless, the recognition of the protein-carbohydrate interaction face and critical residues for binding in the CBM could provide the most direct evidence for its role in processivity.

Besides the proximity, targeting and processivity effect, novel non-hydrolytic functions of CBMs have been discovered, such as hydrogen bond splitting, cellulose decrystallization and thermostabilization effects (Ciolacu, Kovac et al. 2010, Hall, Bansal et al. 2011, Hall, Rubin et al. 2011).

Cel7A is the most abundantly secreted cellulase in *T. reesei*, representing about 60% of the total extracellular protein (Shoemaker, Schweickart et al. 1983, Shoemaker, Watt et al. 1983). It is generally agreed to be one of the most efficient exoglucanases in degrading crystalline cellulose. Cel7A CBD is a family I CBD comprised of 36 amino acids, which preferentially binds to the surface of insoluble, highly crystalline cellulose. The three-dimensional structure was determined by NMR spectroscopy in 1989 (Kraulis, Clore et al. 1989). The wedge-shaped Cel7A CBD possesses a platform-like hydrophobic face which presumably bind to cellulose surface. The domain is comprised of three anti-parallel β sheets which are held together by two disulfide bonds.

Due to the identification of novel non-hydrolytic functions of Cel7A CBD, it could be a promising protein engineering target for improved performance. Protein engineering has been utilized broadly and extensively to enhance hydrolytic enzyme activity, thermostability and identify functionally important residues (Linder, Mattinen et al. 1995, Escovar-Kousen, Wilson et al. 2004, Nakazawa, Okada et al. 2009). The most widely applicable protein engineering strategies include rational design and directed evolution (Percival Zhang, Himmel et al. 2006, Wilson 2009). The rational design of proteins requires foreknowledge of the structure-function relationship (Percival Zhang, Himmel et al. 2006). Random mutagenesis is a traditional and successfully-applied protein engineering strategy to improve the performance of a variety of target proteins

(Lin, Meng et al. 2009, Nakazawa, Okada et al. 2009, Zhang and Zhang 2011). However, the low hit rate and requirement of high throughput screening assay usually hinders the generic application of this approach. Computational protein engineering is a popular tool that has been leading to considerable success. The most well-known computational tools all require structural knowledge such as SCHEMA, CASTing and Rosetta (Bommarius, Blum et al. 2011).

For CBD engineering, the aforementioned strategies are not feasible owing to the lack of catalytic activity, high throughput screening method and limited structural knowledge. In this work, the top nine CBD variants were identified by Principal Component Analysis (PCA), a data-driven protein engineering method, with the concept of consensus (Bansal 2011). Both wild-type CBD and the identified mutants were expressed as GFPuv fusion proteins in bacterial and yeast expression systems to study the effect of glycosylation on CBD performance.

3.2 Materials and methods

3.2.1 Strains, vectors and media

Escherichia coli strain BL21, SHuffle® T7 express *lysY* and *Pichia pastoris* strain KM71H with Mut^S phenotype were used for heterologous protein expression. pET21c (kindly provided by Dr. Shin-geon Choi at Kangwon National University, South Korea) and pET28a-sumo (kindly provided by Dr. Bettina Bommarius at Georgia Tech, USA) were used for cloning and expression for *E. coli*. pPICZα-A (Invitrogen, Carlsbad, CA) was used for cloning and expression for *P. pastoris*. *E. coli* strain XL1-Blue was used for all plasmid construction and storage. *E. coli* was grown in LB medium and *P. pastoris* was grown in BMGY medium containing 20 g/L peptone, 10 g/L yeast extract, 13.4 g/L yeast nitrogen base, 10 ml/L glycerol, 100 ml/L 1 M K_2HPO_4/KH_2PO_4, pH 6.0 and 0.4 mg/L biotin. MagicMedia™ *E. coli* expression medium (Invitrogen) and BMMY (20 g/L peptone, 10 g/L yeast extract, 13.4 g/L yeast nitrogen base, 5 ml/L 100% methanol, 100 ml/L 1 M K_2HPO_4/KH_2PO_4, pH 6.0 and 0.4 mg/L biotin) were utilized for *E. coli* and *P. pastoris* expression respectively. YPDS plates containing 20 g/L peptone, 10 g/L yeast extract, 20 g/L dextrose, 182.2 g/L sorbitol, 20 g/L agar and 100 μg/ml Zeocin™ (Invitrogen) were used for the selection of transformants.

3.2.2 Construction of the expression vectors for GFPuv, GL and GLC

Three genes coding for GFPuv, GFPuv-linker and GFPuv-linker-CBD were constructed. For *P. pastoris* expression, the GFPuv gene was amplified from pGFPuv (Clontech, Mountain View, CA) with forward primer

5’CGGGAATTCACCATG**CATCATCATCATCATCAT**AGTAAAGGAGAAGAACT and reverse primer 5’ GTAGGCGGCGGCCGCTTATTTGTAGAGCTCATCC, introducing the underlined restriction sites for *EcoRI* and *NotI* respectively. The added 6X His tag is shown in bold. DNA encoding the GFPuv-linker-CBD (GLC) was obtained in two steps: firstly, linker-CBD region was obtained by PCR amplification of codon-optimized Cel7A with forward primer 5’ATGGATGAGCTCTACAAAGGAAATCGTGGCACCACC and reverse primer 5’GTAGGCGGCGGCCGCTTACAGGCACTGACTGTA; secondly, the gene coding for GLC was obtained with forward primer 5’CGGGAATTCACCATGCATCATCATCATCATCATAGTAAAGGAGAAGAACT and reverse primer which is the product from the first step. The sequence of the final GLC product can be found in **Figure 58**. The gene coding for GFPuv-linker (GL) was amplified from the GLC PCR product with forward primer 5’CGGGAATTCACCATGCATCATCAT and reverse primer 5’GTAGGCGGCGGCCGCTTAAGGTCCGGGAGAGCTTCC.

After 6 hour restriction with *EcoRI* and *NotI*, the double digested genes were ligated to the same restriction enzymes digested pPICZα-A vector with Quick Ligation™ kit (New England Biolabs, Ipswich, MA). The three final ligation mixtures were transformed by heat shock into XL1-Blue competent cells separately. After selective growth on LB + Zeocin plates at 37℃ overnight, the select colonies were screened for the correct insert using PCR. Recombinant vectors were obtained and sequences were confirmed.

For *E. coli* strain SHuffle expression, the GFPuv, GL and GLC genes were obtained with PCR amplification which introduces the restriction sites for NdeI and HindIII. After 6 hour restriction with NdeI and HindIII, the double digested genes were ligated into the same restriction enzymes digested pET28a-sumo vector with Quick Ligation kit and transformed into XL1-Blue competent cells. After selective growth on LB + Kanamycin plates at 37℃ overnight, the select colonies were screened for the correct insert using PCR. Recombinant vectors were obtained and sequences were confirmed.

3.2.3 Construction of the expression vectors for GLC variants

For *P. pastoris* expression

To introduce mutation to the CBD-containing gene GLC, QuikChange™ site-directed mutagenesis protocol was employed. The designed mutagenic primer pairs including forward (F) and backward (B) primers are shown in **Table 9**. The mutagenized positions for all listed primers are underlined. Backward primers were designed to be reverse complements of the corresponding forward primers. After PCR cycles, DpnI restriction enzyme was added to digest the non-mutated template DNA and the digestion mixture was transformed into XL1-Blue competent cells using heat shock. The colonies containing the target inserts were selectively grown on either LB + Ampicillin or LB + Zeocin plates at 37℃ overnight. The sequences of all mutant-containing genes for GLC variants were confirmed.

For *E. coli* expression

The desired GLC variant gene sequences for *E. coli* expression were directly amplified from the recombinant vectors constructed for *P. pastoris* with PCR to introduce the restriction sites for NdeI and HindIII.

Table 9. Primer pairs for site-directed mutagenesis

Mut #	Mutagenic primer pairs
1	**F:** 5' CAATGCGGCGGTATTGGTTACACAGGACCTACTGTCTGTGCTAGT 3' **B:** 5' ACTAGCACAGACAGTAGGTCCTGTGTAACCAATACCGCCGCATTG 3'
2	**F**: 5' GGTTACAGCGGACCTACTACCTGTGCTAGTGGCACAAC 3' **B**: 5' GTTGTGCCACTAGCACAGGTAGTAGGTCCGCTGTAA 3'
3	**F**: 5' GACCTACTGTCTGTGCTAGTCCTACAACTTGTCAAGTACTGAATCC 3' **B**: 5' GGATTCAGTACTTGACAAGTTGTAGGACTAGCACAGACAGTAGGTC 3'
4	**F:** 5' GGGCAATGCGGCGGTATTGGTTGGAGCGGACCTACTGTCTGTGCT 3' **B:** 5' AGCACAGACAGTAGGTCCGCTCCAACCAATACCGCCGCATTGCCC 3'
5	**F**: 5' CCGGACCTACCCAATCTCATTGGGGGCAATGCGGCGGTATTG 3' **B**: 5' CAATACCGCCGCATTGCCCCCAATGAGATTGGGTAGGTCCGG 3'
6	**F**: 5' GTGCTAGTGGCACAACTTGTAAAGTACTGAATCCTTATTACAGTG 3' **B**: 5' CACTGTAATAAGGATTCAGTACTTTACAAGTTGTGCCACTAGCAC 3'
7	**F**: 5' CTGTCTGTGCTAGTGGCTACACTTGTCAAGTACTGAATCC 3' **B**: 5' GGATTCAGTACTTGACAAGTGTAGCCACTAGCACAGACAG 3'
8	**F**: 5' CTGTCTGTGCTAGTGGCGCGACTTGTCAAGTACTGAATCC 3' **B**: 5' GGATTCAGTACTTGACAAGTCGCGCCACTAGCACAGACAG 3'
9	**F**: 5' GGTTACAGCGGACCTACTGCGTGTGCTAGTGGCACAACTTG 3' **B**: 5' CAAGTTGTGCCACTAGCACACGCAGTAGGTCCGCTGTAACC 3'

3.2.4 Protein expression in *E. coli* and *P. pastoris*

For *E. coli* expression

SHuffle: Target gene-containing recombinant pET28a-sumo vectors were transformed into the expression host *E. coli* SHuffle strain using heat shock. Colony screening was performed to select for the colony with highest expression level. For larger scale expression, overnight pre-cultures were grown in LB medium at 30℃ and inoculated (2.5%) into the MagicMedia containing 25 μg/ml Kanamycin to select for the plasmid-containing cells. After 24 hour expression in MagicMedia, the cell pellets were collected and stored in -80°C for further processing.

BL21: Target gene-containing recombinant pET21c vectors were transformed into the expression host E. coli BL21 strain using heat shock. Overnight pre-cultures were incubated at 37℃ in LB medium and inoculated (1%) to fresh LB supplemented with 50 μg/ml Ampicillin. The cultures were incubated at 37℃ in shaking flasks and expression was induced by addition of 0.5 mM (final concentration) Isopropyl β-D-1-thiogalactopyranoside at OD_{600} of 0.5. Cells were harvested typically after 4 – 6 hours unless otherwise stated.

For *P. pastoris* expression

Sequence confirmed recombinant pPICZα-A were digested with restriction enzyme BstXI. 5 – 10 μg linearized expression vectors were transformed into the expression host *P. pastoris* strain KM71H by electroporation. After selective growth on YPDS + Zeocin plates, the best expression colony was selected by colony screening. For larger scale expression, overnight pre-cultures were prepared in buffered glycerol complex medium (BMGY) and inoculated into the growth medium BMGY at the 1% level. The growth medium was grown at 30°C and 105 rpm in a shaker incubator. The cells were collected by centrifugation at 2,000 X g for 5 min after the OD_{600} reached the

range of 2 – 6. The cell pellet was resuspended in 1/10 volume of growth medium with buffered methanol complex medium (BMMY, induction medium) which contains 0.5% methanol. 100% sterile HPLC grade methanol was added every 24 hours to keep the final concentration at 0.5%. Expression was induced for 3 days in a shaker incubator at 25°C and 250 rpm unless otherwise specified. After expression, culture supernatant was collected for further processing.

3.2.5 Protein purification

For *E. coli*, the cell pellets were thawed on ice and resuspended in Ni-NTA binding buffer to a final OD_{600} of 5 – 10. For each 1 ml cell suspension, the cells were lysed by 1 min sonication followed by 1 min incubation on ice. The clarified supernatant was obtained and applied onto Ni-NTA column for purification. All elution fractions were combined and sumoase Ulp1 digestion was conducted at protein loading of 400 U/mg at 4°C for 15 hours to achieve complete cleavage. Following sumoase digestion, the system was buffer exchanged to Ni-NTA binding buffer with ultrafiltration membrane and re-applied onto Ni-NTA column to collect the purified GLCs in the flow-through. To eliminate the interference from imidazole, the buffer is further exchanged to 50 mM pH 5.0 sodium acetate buffer.

For *P. pastoris*, the pH of the culture supernatant was adjusted to 7.4, centrifuge to obtain clarified supernatant and apply onto Ni-NTA column to purify the his tagged target protein. Combine the elution fractions and exchange to 50 mM pH 5.0 sodium acetate buffer with ultrafiltration membrane.

Figure 16 is a summary of the fusion protein GLC production from the two expression systems.

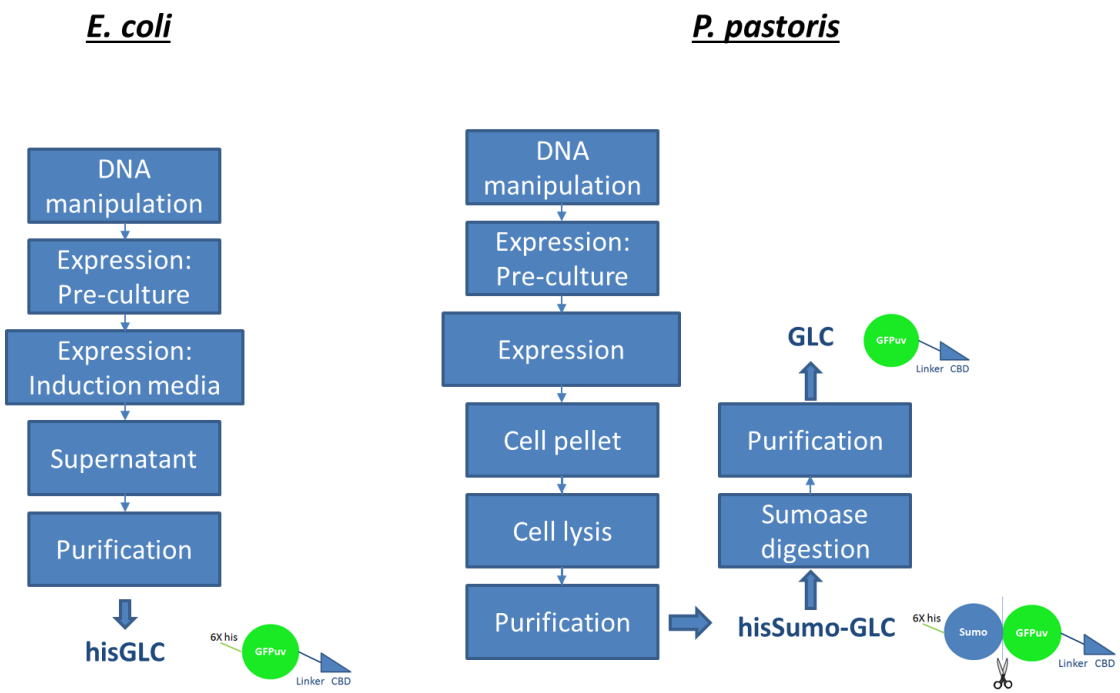


Figure 16. Summary of GLC production from *E. coli* and *P. pastoris*

3.2.6 Protein adsorption

All protein adsorption experiments including GFPuv, GL, GLC and mutation-containing GLCs were performed with Avicel loading 20 g/L at various protein concentrations. The adsorption mixture was agitated for 30 min at 45℃ to reach equilibrium. Following 30 min adsorption, all samples were immediately centrifuged for 3 min at 4℃ and 18,625 g force. The free protein concentration in supernatant was measured based on GFP fluorescence reading by BioTek Synergy H4 Multi-Mode plate reader (Winooski, VT) at excitation and emission wavelength of 385 nm and 508 nm respectively.

3.2.7 Cellulose pretreatment and hydrolysis

Buffer-exchanged GLCs were utilized for pretreatment experiments of model cellulose substrate Avicel. For most of the pretreatment, GLCs were added in various amounts (typically 1-5 mg/g) to Avicel in 50 mM, pH 5.0 sodium acetate buffer; the final Avicel concentration is 40 g/L. The mixture was incubated at 42-45 ℃ with agitation for 15 hours. Then, a dilution was performed with NaOAc buffer to lower the Avicel concentration to 20 g/L and allowed 30 min for the system to re-establish equilibrium. After that, the hydrolysis was initiated by adding 1.1 g/L Celluclast® mixture and 5 kU/L β-glucosidase to the reaction system. The glucose concentration was monitored by DNS assay at desired time points.

3.2.8 X-ray diffraction and crystallinity measurement

XRD patterns of lyophilized GLC pretreated Avicel samples were recorded with a X'pert PRO X-ray diffractometer (PANanalytical BV, Almelo, the Netherlands) using Cu/Kα_1 irradiation (1.54 Å) at 45 kV and 40 mA. The scattering angle (2θ) ranged from 10 °to 40 °with a scan speed 0.021425 s^{-1} and step size 0.0167 °. The crystallinity index of the pretreated Avicel samples was calculated from the XRD spectrum with an analytical method developed previously (Kang, Bansal et al. 2013).

3.3 Results and discussion

3.3.1 The expression of wild-type GFPuv, GL and GLC

The Cel7A CBD (referred to as CBD in the following context) is well-known for its high affinity binding on crystalline cellulose. Besides, it has been found that the CBD and partial linker peptide obtained upon papain digestion could impart the

microcrystalline cellulose decrystallization effect to facilitate enzymatic degradation (Hall, Bansal et al. 2011). Recently, the role of linker and its glycosylation pattern has also been explored for the potential structure-function relationship relevant to mechanistic study of the Cel7A enzyme and the computational simulation results revealed the glycosylated linker peptide, which is previously known as disordered, flexible region connecting CD and CBD, potentially contributes to binding (Payne, Resch et al. 2013). With the foreknowledge on CBD and linker, we first investigated the role of CBD, linker and glycosylation on the heterologous proteins binding behavior from two expression systems *E. coli* and *P. pastoris*.

Figure 17 is a schematic representation of the three GFPuv (referred to as GFP in the following context) fusion proteins constructed in this work. The goal of this study is to understand the role of CBD and linker in binding and other non-hydrolytic phenomena, thus the catalytic domain CD is excluded from the structures. However, the CBD itself is a small peptide comprises of 36 amino acids, to facilitate the target protein detection, localization, purification and subsequent adsorption assay, a GFP was added to the N terminal of the linker region. The combined effect of linker-CBD and linker alone are evaluated from the binding of GLC and GL respectively. A GFP with 6X his tag at the N terminal was also constructed and expressed as control. All three constructs were successfully expressed in three strains: *E. coli* strains BL21, SHuffle and *P. pastoris* strain KM71H. The GFP fluorescence can be utilized as a convenient tool for colony screening.

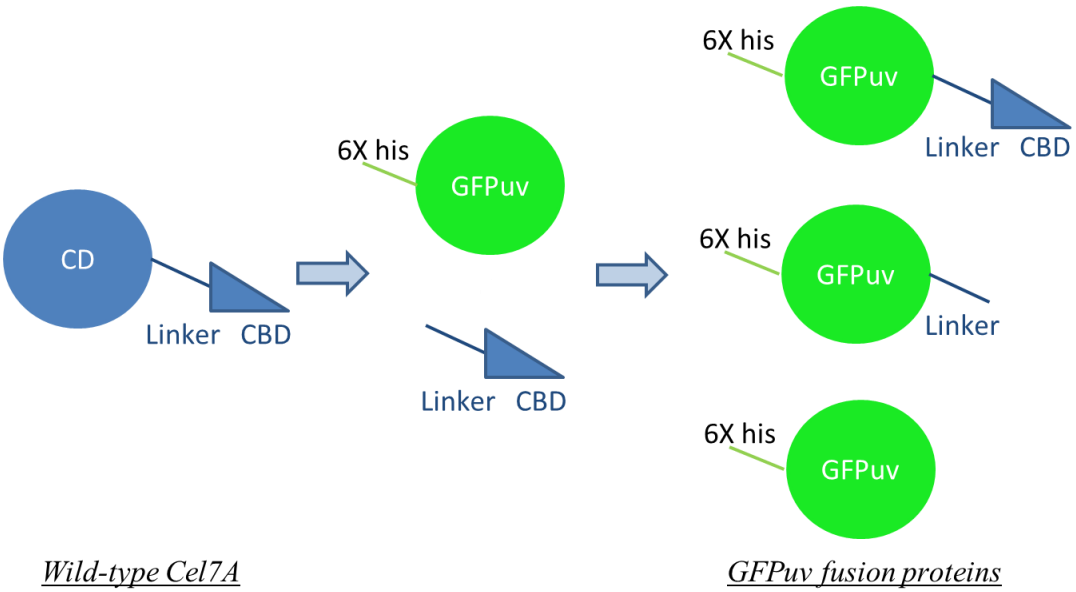


Figure 17. Schematic representation of GFPuv fusion proteins for bacteria and yeast expression systems.

E. coli expression of GFPuv, GL and GLC

The 3D structure of Cel7A CBD was held together and stabilized by three anti-parallel β sheets and two disulfide bonds. The crucial role of disulfide bond in CBD-polysaccharide interaction was demonstrated through the expression of target proteins in *E. coli* expression strain BL21. BL21 is not capable of promoting the formation of disulfide bond in cytoplasm due to the existence of reductases. The expression results was analyzed by SDS-PAGE (

Figure 18). All three target proteins GFP, GL and GLC were expressed intracellularly and exhibit the anticipated molecular weight. Although the three GFP-containing purified proteins all exhibit strong green fluorescence upon excitation by UV light, near zero or negative net adsorption was observed for all three proteins (**Figure 19**). Thus, the expression and binding results indicate the GFP moiety in all fusion proteins

are folded independently and correctly; however, the CBD domain in hisGLC is not correctly folded probably due to the lack of two disulfide bonds. The linker and correctly-folded GFP moieties by themselves or in combination do not contribute to binding (

Figure 19).

To obtain a functional CBD from *E. coli*, the strain SHuffle® T7 express *lysY* (referred to as SHuffle in the following context) was employed. The glutaredoxin reductase and thioredoxin reductase were silenced and disulfide bond isomerase DsbC was introduced to promote the correct S-S formation (Lobstein, Emrich et al. 2012). pET28a-sumo has been demonstrated to be capable of promoting target protein expression yield by forming sumo fusion protein to protect both the bacterial from toxicity of the peptide and the peptide from the host proteolytic enzymes (Bommarius, Jenssen et al. 2010).

With the SHuffle® strain the correctly-folded CBD was successfully expressed in the cytoplasm. Similar to BL21 expression, the GFP and GL expressed from SHuffle® did not show any binding affinity toward Avicel probably due to the lack of CBD domain (results not shown). Based on these observations, it can be inferred the linker region and GFP does not contribute to adsorption by themselves.

Although the new expression strain SHuffle promotes the correct folding of both GFP and CBD, it could not perform post-translational modification in the linker region which is heavily glycosylated in its native host *T. reesei*. The role of glycosylation on cellulase hydrolysis and cellulase-substrate interaction is still not clearly understood. Recently, the role of often overlooked linker region has been revisited by computational simulation. The simulation results pointed out the possibility of polysaccharide binding

affinity of the glycosylated Cel7A linker (Payne, Resch et al. 2013). To gain a better understanding of glycosylation on the fusion protein binding behavior, we conducted expression of GFPuv, GL and GLC in both *E. coli* and *P. pastoris*.

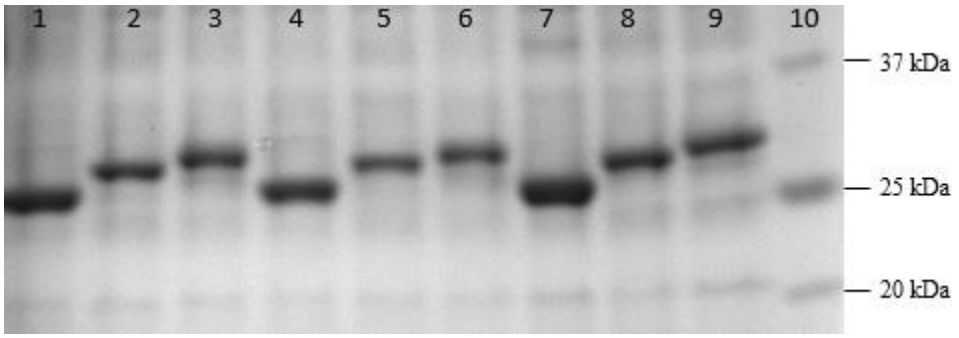


Figure 18. Purified intracellular GFP, GL and GLC from BL21 expression. Lane 1 – 3: GFPuv, GL and GLC after 24h expression; lane 4 – 6: GFPuv, GL and GLC after 48h expression; lane 7 – 9: GFPuv, GL and GLC after 72h expression; lane 10: protein ladder

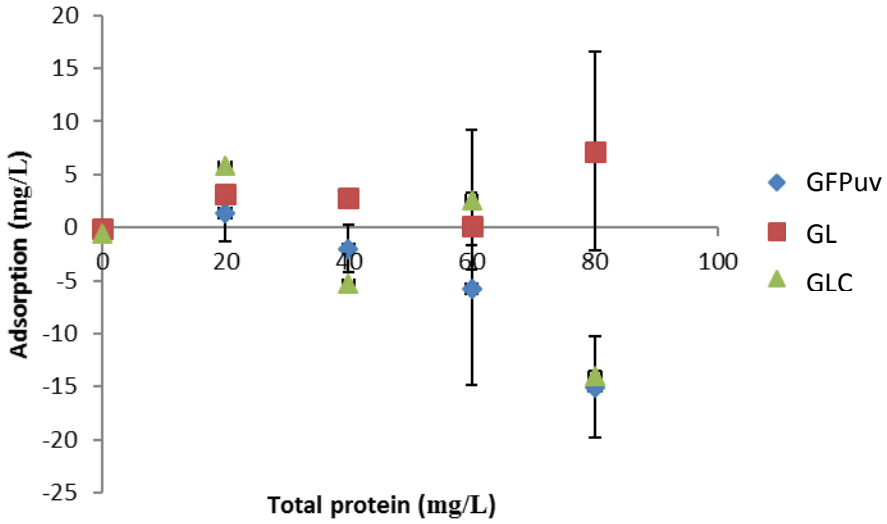


Figure 19. Avicel binding test of GFPuv, GL and GLC expressed from *E. coli* strain BL21.

***P. pastoris* expression of GFPuv, GL and GLC**

P. pastoris has been studied intensively for the past decades as a heterologous protein expression host for a variety of foreign proteins. It is a methylotrophic yeast which can utilize methanol as the sole carbon source. The increasing popularity of expression with *P. pastoris* system can be attributed to several factors (Cereghino and Cregg 2000):

i) Simplicity of techniques for the genetic manipulation

ii) Ability to produce foreign proteins at high levels either intra- or extra-cellularly

iii) Capability of performing eukaryotic post-translational modifications

In addition to *E. coli* expression introduced earlier, the fusion proteins GFPuv, GL and GLC were also expressed in *P. pastoris* to elucidate the effect of glycosylation on CBD functionalities (Dr. S.G. Choi, unpublished results). The adsorption behavior of the fusion proteins were compared qualitatively in **Figure 20**. After 30 min adsorption experiment and a wash step, only GLC remain bound on cellulose surface, which was illustrated by the green fluorescent cellulose pellet (**Figure 20**, 6^{th} tube). No green fluorescence was detected on the pellet from either GFPuv or GL system. Similar to the *E. coli* binding results, the GFP and glycosylated linker were demonstrated to be non-contributing factors in protein-cellulose recognition process.

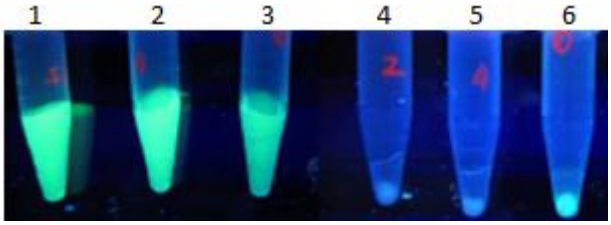


Figure 20. Adsorption of GFPuv, GL and GLC on Avicel. 1 – 3: Purified GFPuv, GL and GLC respectively (before adsorption). 4 – 6: Avicel pellets after adsorption and wash step for GFPuv, GL and GLC respectively. (Dr. S.G. Choi, unpublished results)

3.3.2 Identification of target CBD mutations – Principal Component Analysis (PCA)

The WT GFPuv, GL and GLC were successfully expressed in active form from both microorganisms. A computational analysis based on sequence consensus was carried out by Dr. Prabuddha Bansal to identify target CBD mutations that could potentially improve the CBD performance on cellulose binding and other non-hydrolytic functions (Bansal 2011). **Figure 21** illustrates the PCA-based approach for identification of the CBD mutations.

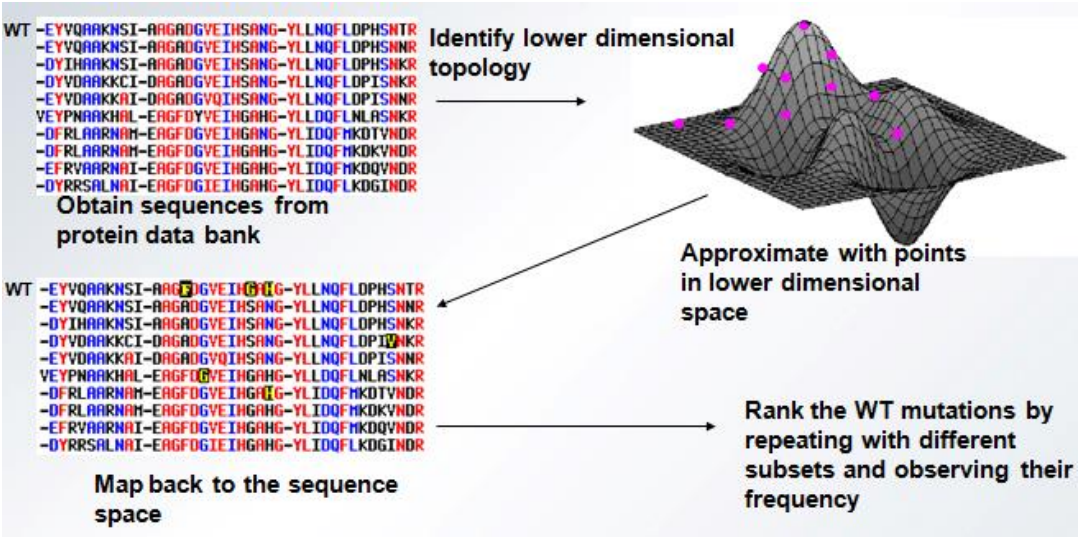


Figure 21. The PCA-based approach for target mutation identification (Bansal 2011)

Previously, the protein engineering of CBD was typically conducted exclusively for identification of functionally important residues toward improved binding on cellulose. The knowledge on structure-function relationship is largely limited to the hydrophobic flat binding face comprises of aromatic and polar amino acid residues for peptide-polysaccharide interactions (Linder, Lindeberg et al. 1995, Linder, Mattinen et al. 1995). Herein, a PCA-based method was employed to generate CBD variants according to the underlying sequence pattern to avoid suggesting mutations that could lead to significant conformational perturbation.

To obtain the family 1 CBDs with high sequence similarity for PCA analysis, a BLAST search was carried out. 41 out of 100 hits were collected. The aligned CBD sequences and corresponding organism origins are given in **Figure 55** and **Figure 56** respectively (Bansal 2011). An important feature of this consensus-based sequence analysis was that the suggested mutation for a given position must have appeared in at least one of remaining 40 CBD sequences. Due to the lack of catalytic activity and high throughput screening assay, the library size must be limited to a controllable number. The top nine mutations identified by PCA are listed in **Table 10**. The following list shows some of the key properties of the PCA method (Bansal 2011):

1) The method identifies the desired mutational positions and the target residues to mutate to, in order of their ranking.
2) PCA involves eigenvalue decomposition of the covariance matrix. Thus, covariations of the residues in the sequence space is accounted for.
3) The suggested target residues are within the library.
4) No structural or functional information is required.

Table 10. Top nine CBD mutations suggested by PCA (Bansal 2011)

Mutant #	WT residue	Position	Suggested residue
1	S	14	T
2	V	18	T
3	G	22	P
4	Y	13	W
5	Y	5	W
6	Q	26	K
7	T	23	Y
8	T	23	A
9	V	18	A

The functionally important aromatic and polar residues Y5, N29, Y31, Y32, Q34 located at the flat putative binding face were found to impart binding affinity to insoluble cellulose through π-π stacking and H-bonding (Linder, Lindeberg et al. 1995, Linder, Mattinen et al. 1995). These residues are highly conserved in the family 1 fungal CBDs; however, Y5 and Y31 in some cases are replaced by tryptophan which retains the aromatic side chain and the stacking potential (**Figure 55**). Y5 was selected by PCA as a target mutation spot ranked 5th in the list; the remaining suggested residues involve none of the critical residues. All nine mutations in **Table 10** will be introduced into CBD sequence as single mutations to study the potential effect on the functionalities.

3.3.3 The expression of GLC mutants

Most cellulolytic enzymes from filamentous fungi contain both N- and O-linked glycosylation, O-glycans usually contain fewer carbohydrate units (Taylor, Talib et al. 2012). The extent and heterogeneity of glycosylation often vary with factors such as expression host and growth condition (Hui, Lanthier et al. 2001). For the multi-modular enzyme *T. reesei* Cel7A, all three moieties CD, CBD and linker are glycosylated in their native expression host: CD contains three N-type, linker contains nine O-type and CBD contains two O-type glycosylation sites. **Figure 22** shows partial Cel7A sequence and its glycosylation patterns.

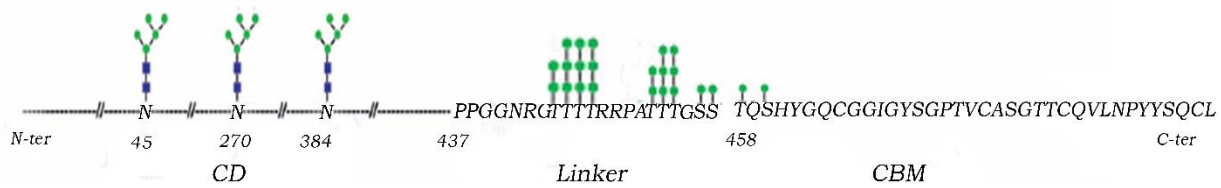


Figure 22. Partial Cel7A sequence and glycosylation pattern in native host. Adapted from (Beckham, Dai et al. 2012).

Based on the previous results in this work, both *E. coli* and *P. pastoris* were demonstrated to be capable of expressing active WT CBD with SHuffle strain and KM71H strain respectively. GFP and linker by themselves or in combination did not contribute to substrate binding. To examine the effect of glycosylation on the non-hydrolytic functions of WT and variant CBD-containing GLCs, the encoding genes were also expressed in the two systems.

Table 11 summarizes the GLC variants expression results from *E. coli* and *P. pastoris*. The identified CBD mutations were introduced into the CBD sequence as single mutations. The mutants are designated as M1 – M9 according to their ranking in PCA analysis.

Table 11. Summary of expression status of WT and variant GLCs.

	WT	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**
E. coli	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓
P. pastoris	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓

The variant M4 failed to be constructed for expression in both systems, possibly due to the low tolerance to mutation at position 13. Although Y13 is not exposed to the surface, this position is highly conserved with aromatic residues which might hold a structural role. M7 expressed well in *E. coli*, however, no M7-containing colony was identified after colony screening for *P. pastoris*. Except for M4 and M7, the remaining GLC variants were successfully constructed and expressed in both microorganisms. The expression levels of WT and all GLC single mutants are summarized in **Table 27** and

Table 28.

For *E. coli* expression, the GLCs were expressed intracellularly and fused to sumo which has been reported to promote fusion protein yield as well as protect target protein from proteolysis (Bommarius, Jenssen et al. 2010). The *P. pastoris* expression system expresses target proteins extracellularly and the proteins were conveniently purified from culture supernatant. Since the *P. pastoris* strain secrets very low levels of native proteins,

the secretion step could serve as an initial purification step and vast majority of the total protein in the media was comprised of heterologous protein (Cereghino and Cregg 2000).

Figure 23 shows the SDS-PAGE analysis of purified GLC variants expressed from two systems and visualized by different staining systems. Two protein gels loaded with the same proteins were performed in parallel, one was stained with coomassie blue and the other stained with glycoprotein staining kit to detect protein glycosylation. The magenta band appeared after glycoprotein staining and is indicative of protein glycosylation, otherwise no band will show up. As expected, completely blank lanes were observed after glycoprotein staining for *E. coli* side, indicating the host strain SHuffle was not capable of performing protein glycosylation. All the GLC bands expressed from *P. pastoris* stained magenta which stands for glycosylated protein from the eukaryotic host KM71H. Typically, overglycosylation phenomena of the intact Cel7A enzyme was observed for the eukaryotic heterologous hosts *P. pastoris* and *S. cerevisiae* at the N-glycosylation sites, the former was found to be less extensive and closer to the native pattern (Boer, Teeri et al. 2000). In this work, the calculated molecular weight for GFPuv, GL and GLC are 27.6, 29.7 and 33.4 kDa respectively. The potential glycosylation sites for both CBD and linker were O-type, thus no overglycosylation occurred judging from the apparent molecular weight on SDS-PAGE (**Figure 23**). The detailed glycosylation pattern at each site was not examined further.

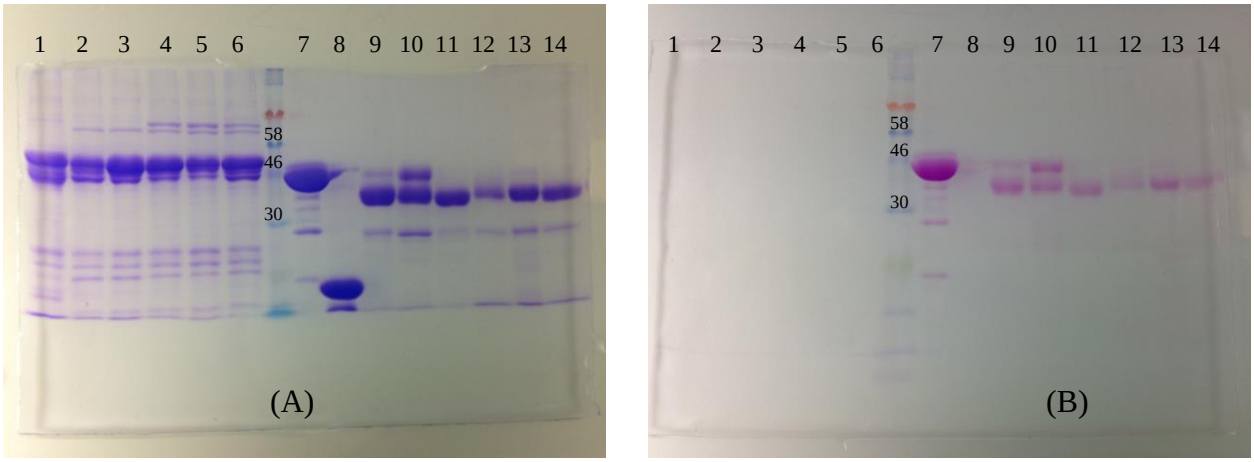


Figure 23. (A) SDS-PAGE stained with coomassie blue R-250 dye and (B) SDS-PAGE stained with glycoprotein staining kit. PC: positive control (glycoprotein); NC: negative control (aglycoprotein). For both gel, lane 1-6 are sumo-GLC mutants (calculated MW ~51 kDa) 1, 2, 3, 5, 6, 8 expressed from *E. coli* and lane 9-14 are the GLC mutants (calculated MW ~34 kDa) 1, 2, 3, 5, 6, 8 expressed from *P. pastoris*.

3.3.4 Adsorption isotherms

It was observed in the 1990s that single amino acid changes in the binding face of the wedge-shaped Cel7A CBD can destroy the interaction with crystalline cellulose (Reinikainen, Ruohonen et al. 1992). Nevertheless, the mutational studies mainly focused on Y5, Y31, Y32, N29 and Q34; drastic changes in these highly conserved positions probably can lead to either impaired structural integrity or loss of binding capacity. In this work, it was demonstrated the introduction of a single amino acid mutation at the non-crucial positions could also result in a dramatic change in the binding behavior. **Figure 24** and **Figure 25** show the binding isotherms of GLCs on microcrystalline cellulose (Avicel), from *E. coli* and *P. pastoris* respectively. All mutations, glycosylated or non-glycosylated, were found to be capable of binding on the microcrystalline cellulose, however, the binding affinity varies broadly.

Effect of single amino acid mutation on GLC binding

In both expression systems, some mutants were found to bind worse than the remaining ones, including M3 (G22P), M5 (Y5W) and M8 (T23A). To quantitatively compare the binding behavior, the adsorption data for GLCs was modeled by Langmuir isotherm (Payne, Resch et al. 2013). The Langmuir constant maximum binding capacity (B_{max}) for all variants are listed in **Table 12** in ascending order.

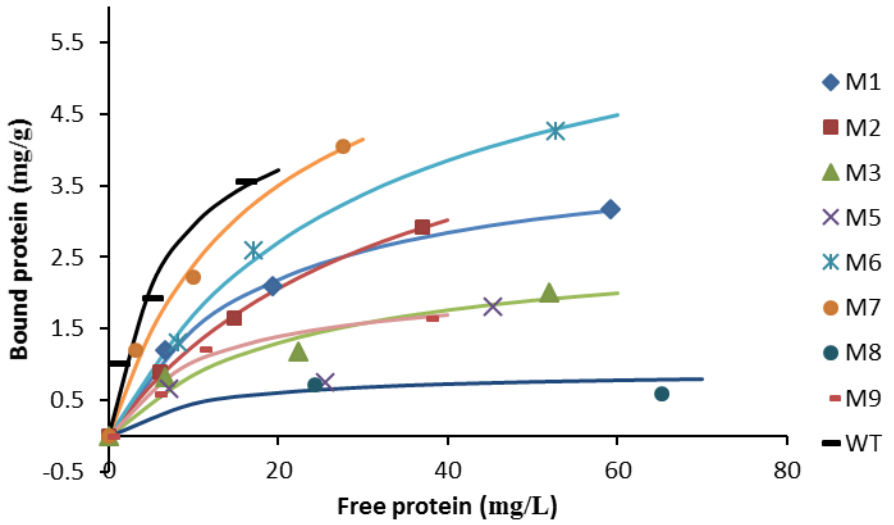


Figure 24. Adsorption isotherms and non-linear Langmuir isotherm fitting for GLC variants expressed from *E. coli*.

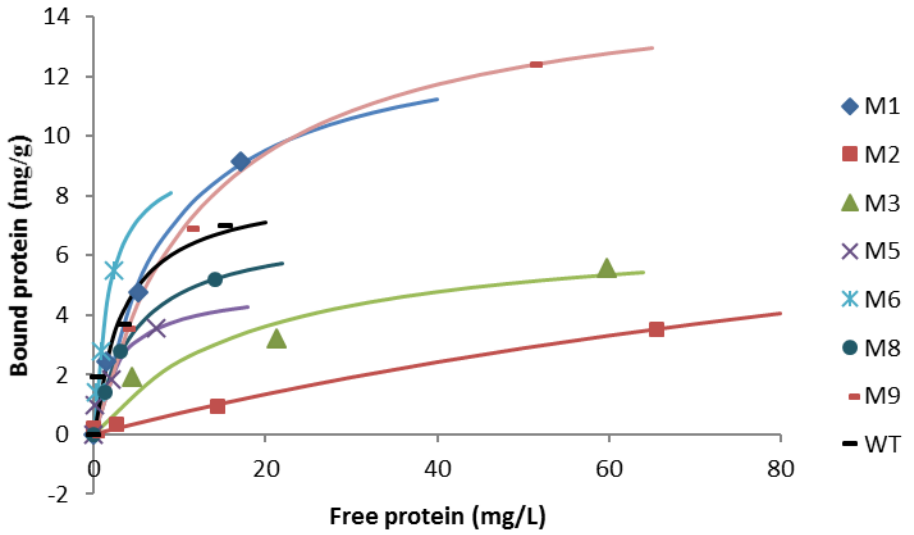


Figure 25. Adsorption isotherms and non-linear Langmuir isotherm fitting for GLC variants expressed from *P. pastoris*.

For both cases, the M5 (Y5W) displayed impaired binding capacity comparing to WT CBD. Replacing a binding face tyrosine residue with tryptophan is a routinely constructed CBD mutation that usually utilized for illustration of the significance of aromatic side chain on cellulose binding. Compared to tyrosine, the tryptophan residue with larger aromatic side chain could enhance the van der Waals force of π-π stacking and eliminate the phenolic hydroxyl group which contributes to hydrogen bonding (Reinikainen, Ruohonen et al. 1992). Contradictory results were reported in the previously published work on this particular mutation. Linder et al. found Y5W enhanced the binding affinity by 33% (Linder, Lindeberg et al. 1995), which could possibly be attributed to the larger binding surface provided by tryptophan. However, a ΔΔG (defined as ΔG_{W5} - ΔG_{Y5}) value -0.24 ± 0.55 kcal/mol was observed from computational simulation and it seems to indicate that both hydrophobic surface area and electrostatics are important for binding (Li, Yan et al. 2012). In some cases, the electrostatic interaction

was suggested to dominate the energy changes in CBD-mediated processivity and substrate binding (Beckham, Matthews et al. 2010, Li, Yan et al. 2012).

In this work, the reduction in binding affinity induced by Y5W could possibly provide experimental support for the simulation finding that H bond dominates van der Waals interaction in the CBD-cellulose binding. Another factor that is also likely to contribute to the loss in binding is the potential conformational change. The Y5 region was known to be sensitive to amino acid substitution as it exists in a type II turn at a position usually occupied by glycine (Linder, Lindeberg et al. 1995). The discrepancy between the experimental measurements could be resulted from the different means of peptide production, i.e., peptide synthesis, different expression host and expression conditions.

Regarding M3 (G22P), it was also found to display impaired binding affinity in both expression hosts. Similar to Y5, this can still be ascribed to the secondary structure of the region where G22 locates, a tight turn which was originally occupied by a glycine residue (**Figure 26**). Glycine imparts high flexibility to the backbone; proline, to the contrast, brings considerable conformational restrict in the backbone and it is not able to act as a H-bond donor to stabilize the turn structure (Gouverneur and Müller 2012). Therefore, the Gly to Pro mutation could potentially introduce conformational changes in the backbone structure and further impact the binding behavior of CBD.

For M8 (T23A), it is adjacent to the G22 residue and the T23 likewise lies within the abovementioned turn structure. The mutation from a polar amino acid (Thr) to a nonpolar one (Ala) usually interferes with the H-bond formation, which possibly contribute to the reduction in binding affinity. Out of the nine suggested mutations, there

is another one involving the position T23, i.e. M7 (T23Y). Unlike the T23A mutation, M7 retained the H-bond formation potential by introducing a Tyr with aromatic side chain and hydroxyl group. To the contrary, the binding affinity of M7 is dramatically enhanced, implicating the crucial role of position 23 within the tight turn.

The three residues Y5, G22 and T23 are all located in turn structures that are sensitive to mutation; the binding behavior was impacted by the selected mutations at these positions. The adsorption results present here provide evidence that tight turns connecting adjacent segments of secondary structures are capable of altering binding, although it is not involved in the direct interaction of peptides with polysaccharides.

Table 12. Ranking of B_{max} of GLC variants expressed from *E. coli* and *P. pastoris*

Rank	*E. coli*		*P. pastoris*	
	Mut #	**B_{max} (mg/g)**	**Mut #**	**B_{max} (mg/g)**
1	8	0.91	5	4.96
2	9	2.17	8	6.98
3	3	2.72	3	7.04
4	1	4.06	WT	8.37
5	WT	5.06	6	9.69
6	2	5.62	2	12.44
7	7	6.56	1	13.72
8	6	6.68	9	15.54

Besides the mutations with negative effects, M2 (V18T) and M6 (Q26K) expressed from both expression systems displayed improved binding capacity on

microcrystalline cellulose (**Table 12**). V18 is located at the rough face of CBD. V18T introduced a reactive hydroxyl group with the H-bond formation potential that could help stabilize the coil structure and even facilitate the binding. In fact, although the cellulose binding face was postulated to be the hydrophobic flat face comprised of aromatic residues, there are still some controversial reports suggesting the rough face might also be involved in polysaccharide binding. The mutation P16R located on the rough face was introduced and found to greatly reduce both binding and activity due to the break in the surface, which provides evidence to its direct involvement in peptide-carbohydrate recognition (Reinikainen, Ruohonen et al. 1992). The discrepancies found in publications might result from the differences in the peptide preparation methods. **Table 13** shows the comparison of B_{max} for Cel7A and its domains from microorganisms and chemical synthesis. The same enzyme or domain expressed from different microorganisms can be quite different in terms of B_{max}. In general, the binding capacity of the heterologously expressed peptide tends to be compromised, possibly due to the overglycosylation problem. The synthetic CBD shows 100-fold higher B_{max}, the dramatic difference could be explained by the different preparation method and the presence of GFP which might induce a steric effect and block the surface of cellulose. Besides, it cannot be ruled out that the fitting method and extent of saturation for binding isotherms impact the resulting fitting parameters.

Table 13. Comparison of B_{max} for Cel7A and its components

Protein	**Expression host**	**B_{max} (nmol/g)**	**Reference**
Cel7A	*T. reesei*	1200	(Reinikainen,

Table 13 continued

	S. ceravisiae	820	Ruohonen et al. 1992)
Cel7A CD	*T. reesei*	580	
	S. ceravisiae	320	
Cel7A CBD	Synthetic	25000	(Linder, Lindeberg et al. 1995)
	P. pastoris	251	This work
	E. coli	151	This work

The same position V18 appeared twice in the top nine list from PCA: M2 (V18T) and M9 (V18A). The M9 binding behavior in terms of B_{max} varied depending on the status of glycosylation. The glycosylated M9 had the highest binding capacity among all the variants from *P. pastoris*, once again showing the potential involvement of CBD rough face in cellulose binding. Similarly, M1 (S14T) located at the rough face also imposed a significant positive impact in B_{max} with the presence of glycosylation.

M6 (Q26K) is located within the β sheet closest to the rough face. Upon mutation to a charged Lys residue, the binding was improved for both systems compared to the WT. This could imply the significance of electrostatic interactions in the cellulose recognition process. Since cellulose is negatively charged (Sood, Tyagi et al. 2010), the positively charged Lys residue might create additional favorable interaction toward the cellulose surface.

In this CBD mutational work, these variants, except for M5, have not been studied previously, thus there is no adsorption data for comparison. Although the potential functional roles for certain positions have been discussed, the vast majority was derived from molecular simulation findings (Hoffrén, Teeri et al. 1995).

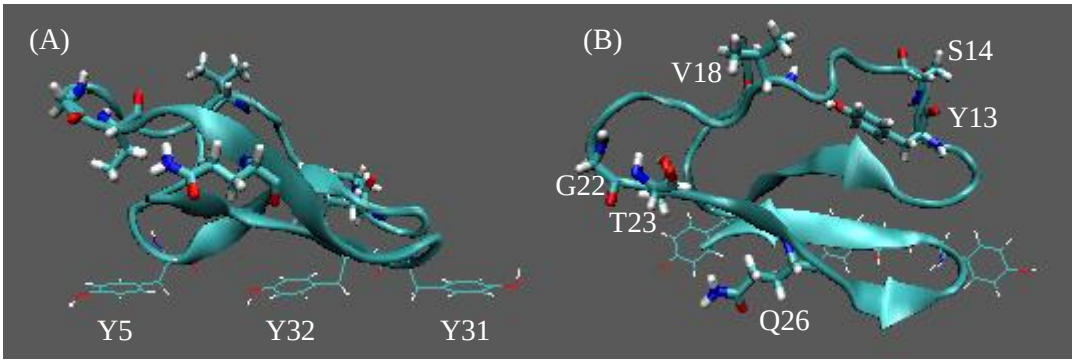


Figure 26. Side view (A) and top view (B) of Cel7A CBD. Three tyrosine residues on the flat binding face are indicated in (A), six target mutation positions located away from flat binding face are indicated in (B).

Effect of glycosylation on GLC binding

Apart from the effect of single amino acid substitutions in CBD sequence, the glycosylation was demonstrated to significantly impact the adsorption results as well. For all CBD variants, eight from *E. coli* and seven from *P. pastoris*, the most crucial difference between the proteins from two expression systems was the presence of glycosylation in *P. pastoris*. The binding results for the expressed GLCs are summarized in **Table 14**. Clearly, for all GLC variants, the binding capacity of glycosylated form exhibit 1.45-fold to 7.67-fold enhancement over the non-glycosylated counterparts, meaning peptide-bond glycans on linker and CBD might be directly involved in the interaction with cellulose. The mechanism for the sugar-cellulose interaction is similar to peptide-cellulose, which is based on π-stacking and H-bonding (Chen, Drake et al. 2014).

Table 14. Summary of binding results for GLCs expressed from *E. coli* and *P. pastoris*.

Mut #	Mutation	B_{max} (mg/g)	K_{ad} (L/mg)	Regional structure	Binding results*

Table 14 continued

		E. coli	*P. pastoris*	*E. coli*	*P. pastoris*		B_{max}	K_{ad}
WT	-	5.06	8.37	0.14	0.286	-	-	-
1	**S14T**	4.06	13.72	0.058	0.114	Coil	+	
2	**V18T**	5.62	12.44	0.029	0.006	Coil	++	
3	**G22P**	2.72	7.04	0.046	0.053	Turn		
5	**Y5W**	N/D	4.96	N/D	0.331	Turn		+
6	**Q26K**	6.68	9.69	0.034	0.556	β sheet	++	+
7	**T23Y**	6.56	N/A	0.057	N/A	Turn	+	
8	**T23A**	0.91	6.98	0.1	0.208	Turn		
9	**V18A**	2.17	15.54	0.09	0.077	Coil	+	

*Binding results: + indicates the binding parameter is improved in either E. coli or P. pastoris expression system; ++ indicates improvement in both systems.

To date, the majority of the research effort has been directed towards understanding the structure-function relationship for the functional domains of carbohydrate-active enzymes, especially the core domain CD. Recently, additional function of the linker region and CBD was postulated by molecular dynamics (MD) simulation. The O-glycosylated linker could not only impart protease resistance but also dynamically bind onto cellulose surface (Payne, Resch et al. 2013). Similarly, various native and artificial glycosylation patterns of CBD were also studied and a 140-fold increase in binding affinity relative to non-glycosylated counterpart reported (Taylor, Talib et al. 2012). In addition to binding affinity enhancement, the glycosylated CBDs were experimentally found to confer 50-fold increased thermolysin resistant and improve thermostability by up to 16 ℃ (Chen, Drake et al. 2014). Therefore, it has been proved

both computationally and experimentally that glycosylation could render additional functions to linker and CBD.

However, there was only one experimental validation of the computational results and the maximum increase in glycosylated CBD binding affinity was found to be 7.4-fold which is significantly lower than the 140-fold revealed by MD simulation (Chen, Drake et al. 2014). Our binding results shown the glycosylated fusion protein GFP-linker (GL) expressed from *P. pastoris* did not adsorb on microcrystalline cellulose by itself; a CBD domain was required to exhibit binding affinity (**Figure 20**). Although the simulation results do not completely agree with experimental findings, the potential functional role of covalently bond glycans on CBD and synergistic effect of glycosylated linker and CBD on binding was observed. To draw a conclusion on the potential novel functions of glycosylation, more experimental validation on various families of CBDs from different origins is needed.

3.3.5 Utilize GLCs as biological pretreatment agents

Since the linker-CBD was found to act as a biological pretreatment agent in a previous publication from our group, the GLCs obtained from *P. pastoris* were utilized in pretreatment step prior to addition of cellulolytic enzymes. *P. pastoris* expressed GLCs are glycosylated and structurally closer to the natively expressed counterparts, therefore, they are preferred for the pretreatment process. However, the 1-h hydrolysis results after 15 hour pretreatment with 1 to 5 mg/g GLC loadings do not seem promising (**Figure 27**). All variants exhibit similar hydrolysis yield at three GLC loadings; as GLC dosage increases, the subsequent hydrolysis rate reduces. This phenomena could be potentially

ascribed to the binding competition between GLC and cellulase, especially the cellobiohydrolase Cel7A. Both proteins consist of the same CBD domain competing for the same binding sites; at higher dosages of GLC, the available sites are largely blocked, which in turn cause the drop in hydrolysis yield. Due to the blocking effect of GLC, the non-hydrolytic function facilitating cellulose hydrolysis is likely masked.

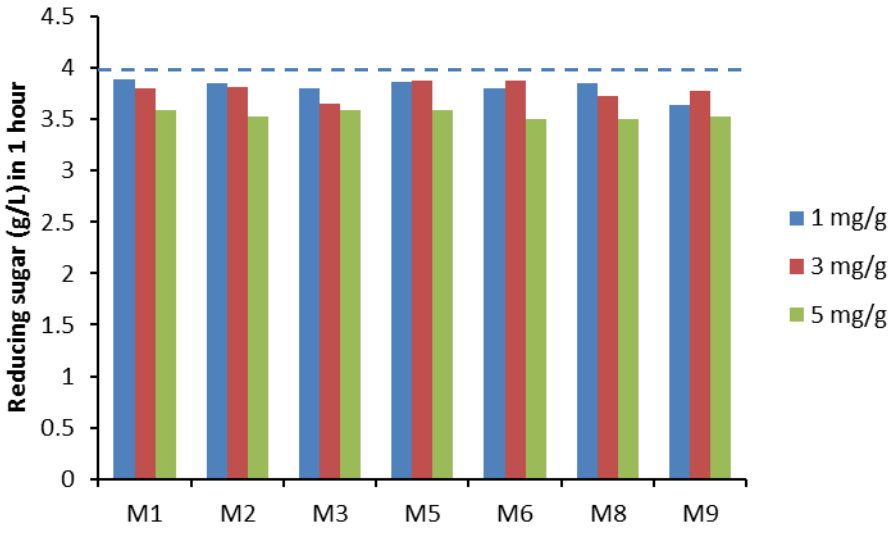


Figure 27. Initial 1 hour hydrolysis rate of Avicel after GLC variants pretreatment at 45 ℃ for 15 hours. Dotted line shows 1-h hydrolysis rate for non-pretreated control.

To alleviate the binding competition, a desorption step was added prior to the hydrolysis step to dilute the system and re-establish the equilibrium. An initial concentration of 40 g/L cellulose loading was utilized for 15 h pretreatment step, followed by dilution with NaOAc buffer (pH 5.0, 50 mM) to 20 g/L and 30 min was allowed for the system to re-reach equilibrium.

Although only 10% bound GLC was desorbed from cellulose surface (results not shown), the conversion yield for the select two mutants M3 and M9 were clearly

improved than the control throughout the conversion range studied (**Figure 28**). To further enhance the pretreatment outcome, complete elimination of competitive binding was desired.

The crystallinity of Avicel after M3 and M9 pretreatment was calculated to be 50.0% and 47.1% respectively based on XRD spectra (results not shown). The increase in decrystalization effect corresponds to increased hydrolysis yield, meaning the enhancement in hydrolysis could possibly due to the reduction in crystallinity.

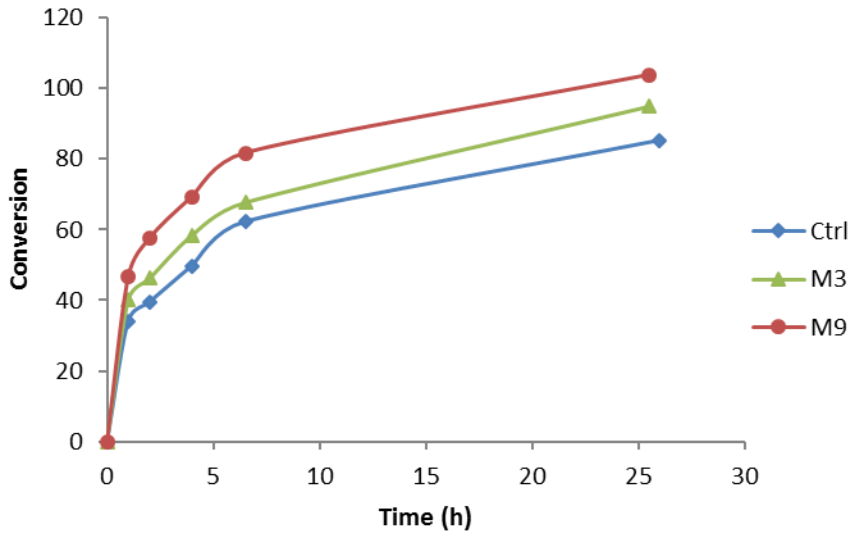


Figure 28. Avicel conversion curve after GLC variants pretreatment and dilution. Pretreatment with GLC variants was performed at 42°C for 15 hours with 40 g/L Avicel and 5 mg/g GLC loading.

3.4 Conclusions

In this work, nine CBD variants were identified by Principal Component Analysis (PCA), a data-driven protein engineering method, with the concept of consensus. Wild-type and variant CBD were successfully expressed as GFPuv fusion proteins in active

form from bacterial (*E. coli*) and yeast (*P. pastoris*) expression systems. Both single amino acid mutation and existence of glycosylation were found to significantly impact the binding behavior of GLC variants. Thus, the direct interaction between peptide and cellulose involving both surface amino acid residues and peptide-bond glycans. The mutants with improved cellulose binding affinities were identified, opening the chance to combine the CBDs with the catalytic domain to yield variant cellulases with improved enzymatic activity on cellulose substrate. In addition to binding, the GLC variants were also investigated for their potential to be utilized as biological pretreatment agents. An effective biological pretreatment in time-efficient manner without the need to remove inhibitors and exchange buffer could be derived from the GLC variants.

CHAPTER 4

A NOVEL CHEMICAL PRETREATMENT METHOD FOR CELLULOSE AND LIGNOCELLULOSICS

4.1 Introduction

Lignocellulosic biomass is an important and renewable alternative resource for liquid transportation fuels and for chemicals. Compared to the food-based raw materials for first generation biofuels, non-edible plant-based biomass is generally agreed to be highly recalcitrant due to the complexity of plant cell wall structure. To fully explore the potential of biomass utilization in biofuel industries, a certain pretreatment step is crucial to break down the complex and compact three-dimensional structure of cell wall polymers to fractionate into its three major components, which are cellulose, hemicelluloses and lignin (Rinaldi and Schuth 2009). The current pretreatment techniques can be grouped into four major categories: physical, physiochemical, chemical and biological. Although enormous efforts have been put into developing different methods, the cost of pretreatment for vast majority of the current techniques limits their commercial viability and researchers around the world are focusing on finding new cost-effective methods and/or optimizing the present techniques.

The existence of lignin in the plant accounts for a great portion of the added difficulties for second-generation biofuels production. Lignin can dramatically impair the enzymatic hydrolysis efficiency by both chemically bonding to hemicelluloses and cellulose and physically blocking cellulose and cause enzymes to non-productively bind

to the lignin (Galbe and Zacchi 2012). Lignin is a complex three-dimensional natural compound with irregular structure. It can be considered as the polymerization product of three fundamental phenylpropane units: p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Pinkert, Goeke et al. 2011). These monolignols are incorporated into lignin macromolecules in the forms of phenylpropanoid, i.e. p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S). Lignin content and composition vary depending on the sources of biomass. In general, softwood lignin from coniferous trees has a higher lignin content compared to hardwood and is primarily comprised of G-lignin, while hardwood is rich in both S and G lignin. H lignin on the other hand is less prevalent in woody biomass, however, it is commonly found in grasses and agricultural residues. Among the different biomass, softwood is the so-called "worst case senario," which renders greatest challenge to the biofuel production processes such as pretreatment and fractionation (Kang, Bansal et al. 2013).

For the past decade, ionic liquids (ILs) have been intensively studied as a biomass solvent and fractionation agent for a variety of biomass species (A. da Costa Lopes 2013). Joint BioEnergy Institute (JBEI), one of the three DOE funded Bioenergy Research Centers (BRCs), has focused on ionic liquid pretreatment to understand the technology at the cellular level to develop a more effective, affordable and scalable pretreatment technology (Singh, Simmons et al. 2009, Sun, Li et al. 2013). As a molten salt with melting temperature lower than 100 ℃, both the cation and anion have been studied either separately, or in combination, to understand their respective effect on biomass swelling, dissolution and structural modification (George, Tran et al. 2011). The role of cation and anion is still under debate. Both cations and anions of ILs have been

demonstrated to pose profound impact on the plant cell wall swelling, dissolution and physical properties of ILs (Zhang, Sun et al. 2006, Brandt, Hallett et al. 2010, Lu, Xu et al. 2014). The imidazolium based ILs are the most intensively studied and earlier publications reported that imidazolium cation has significant π-stacking and hydrogen bonding interaction with lignin through both experimental and modeling results (Kilpelainen, Xie et al. 2007, Janesko 2011). Although encouraging results have been achieved with ILs, the often high cost of IL synthesis and purification and its commonly high viscosity add to the difficulties of handling and retard its commercialization with current technology. Recently, the inexpensive $[HSO_4]^-$ based ILs show promise in fractionation and deconstruction of lignocellulosic biomass, which could potentially lead to a cost-effective IL pretreatment (Verdía, Brandt et al. 2014).

Inspired by the earlier work, a novel biomass pretreatment/fractionation agent 1-methylimidazole (MI) has been studied for its impact on cellulose and lignocelluloses structure, composition and subsequent enzymatic hydrolysis step. MI is a precursor to some imidazolium-based ILs. MI's first reported use in biomass dissolution was as part of a bi-solvent system with dimethylsulfoxide (DMSO) to fully dissolve finely ball-milled biomass at room temperature (Lu and Ralph 2003). The reaction mechanism for lignin esterification with acid anhydrides using MI as catalyst was also revealed (Connors and Pandit 1978, Fox and McDonald 2010). In this work, we will focus on the potential of utilizing MI as a pretreatment agent for both cellulose and lignocelluloses and understanding its impact on different substrates. To our knowledge, this is the first investigation of MI as a means of pretreatment to facilitate the subsequent processes in biofuel production.

4.2 Materials and methods

Materials

Avicel PH-101 (Cat. No. 11365) and Fibrous cellulose (Cat. No. C6288, from cotton linters) purchased from Sigma-Aldrich (St Louis, MO, USA) were employed as model cellulose substrates in this study. The steam-exploded loblolly pine (SELP) samples were provided by Dr. John Muzzy at Georgia Institute of Technology, Atlanta, GA. Steam-exploded bagasse (SEB) and steam-exploded wheat straw (SEWS) were provided by Dr. Guido Zacchi at the University of Lund, in Lund, Sweden. 1-Methylimidazole (99%), Cellulase from Trichoderma reesei (Celluclast®, 159 FPU/mL) and β-glucosidase from almonds (5 U/mg) were purchased from Sigma-Aldrich. The BCA protein assay kit was purchased from Thermo Fischer Scientific (Rockford, IL, USA).

MI pretreatment

The MI pretreatment was typically conducted with 100% MI at 40 g/L substrate loading at 25℃ for 5 min with thorough mixing unless otherwise stated. Following the pretreatment step, the resulting substrate was washed with 2 vol. of DI water for four times (total 8:1 v/v DI water/MI). The washed sample was subjected to enzymatic degradation and the loss of cellulose during pretreatment was quantified by mass balance and dinitrosalicylic acid (DNS) assay (Kang, Bansal et al. 2013). The loss of lignin for lignin-containing sample was quantified by mass balance. All these experiments were duplicated.

Enzymatic hydrolysis

Batch hydrolysis reactions were carried out in 1.5 mL eppendorf tubes with 2% substrate consistency. Enzymatic hydrolysis was carried out at 0.63 FPU/mL cellulase loading, 45°C with thorough mixing for desired time. β-Glucosidase was added at 5 U/mL to minimize product inhibition(Bommarius, Katona et al. 2008). The total reaction volume was 1 mL unless otherwise stated. At the desired time point, the tubes containing reaction mixture were immediately chilled on ice and centrifuged at 14,000 rpm for 3 min. The reducing sugar concentration in the supernatant was measured by DNS assay with glucose standards. All the hydrolysis experiments were carried out in duplicate.

Measurement of degree of polymerization (DP)

The number-average DP was measured and calculated as the ratio of total glycosyl monomer and reducing end concentration according to the published protocol(Zhang and Lynd 2005). Phenol-sulfuric acid method and BCA assay were utilized to determine the two concentrations respectively. The experiments for DP measurement were carried out in triplicate.

Adsorption studies

Substrates (2% consistency) were pre-incubated in 1 mL NaOAc buffer (50 mM, pH 5.0) at 45 ℃, 1000 rpm for 1 h followed by cooling on ice. A broad range of cellulase loading ranging from 0.0048 to 3.8 FPU/mL were added and further agitated for 30 min at 4 ℃ to minimize the influence from hydrolysis. Following 30 min adsorption, all samples were immediately centrifuged for 3 min at 4 ℃, 14,000 rpm and supernatant was analyzed for free protein concentration by BCA assay kit. To eliminate the interference from reducing sugar, the free protein in supernatant was precipitated out and re-dissolved for BCA assay(Bansal P., Vowell B. J. et al. 2012, Kang, Bansal et al. 2013). The bound

protein (mg/g) is calculated as the difference between total and free protein concentration (mg/L) divided by the substrate concentration (g/L). All the adsorption experiments were carried out in duplicate.

X-ray diffraction (XRD)

XRD patterns of lyophilized samples were recorded with a X'pert PRO X-ray diffractometer (PANanalytical BV, Almelo, the Netherlands) using Cu/Kα1 irradiation (1.54 Å) at 45 kV and 40 mA. The scattering angle (2θ) ranged from 10° to 40° with a scan speed 0.021425 s^{-1} and step size 0.0167 °.

SDS-PAGE analysis of enzymatic hydrolysate

20 µl enzymatic hydrolysate after 96 h conversion of SELP was separated on 12% SDS-PAGE along with protein standards with known molecular weight. After electrophoresis, the gel was stained with 0.1% w/v Coomassie Brilliant Blue R250 to visualize the protein bands.

4.3 Results and discussion

4.3.1 Effect of MI on pure cellulose

The effect of MI was first investigated on two pure model cellulose substrates Avicel and fibrous cellulose (FC) with experimental conditions of 37℃ and 1 hour. Based on mass balance and reducing sugar assay results of the MI supernatant, there is negligible cellulose dissolution induced by the pretreatment (< 5 wt%). However, under the mild pretreatment conditions, the conversion yield of the two MI-pretreated substrates (**Figure 29**) is substantially higher compared to the control which underwent the pretreatment process at the same conditions with 50 mM, pH 5.0 sodium acetate

(NaOAc) buffer instead of MI. To obtain 60% conversion, MI-pretreated Avicel requires less than 3 hours while 24 hours is needed for the control sample. The conversion yield for FC is slower than Avicel partly due to the difference in crystallinity index (CrI) (72% and 60% for FC and Avicel, respectively) (Hall, Bansal et al. 2011).

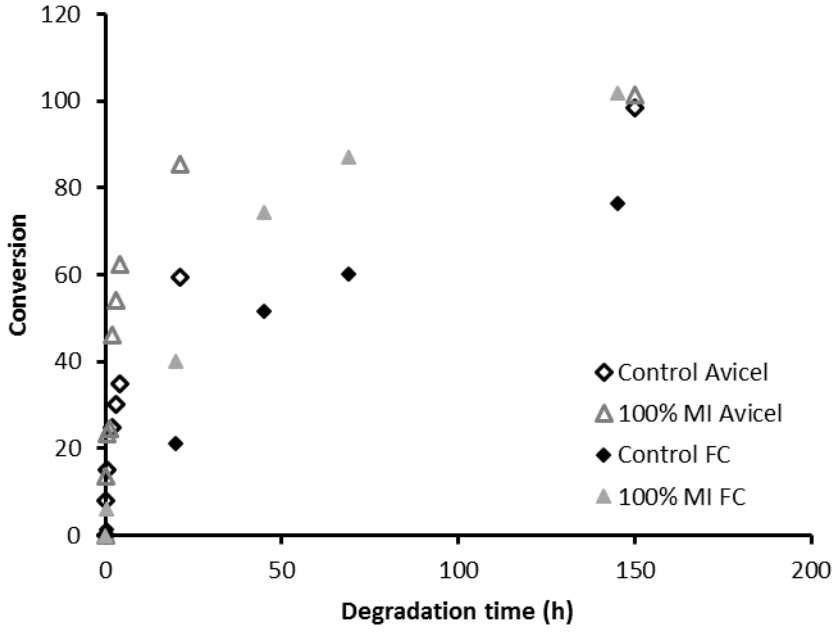


Figure 29. Conversion curves for control and MI-pretreated (a) Avicel and (b) FC. Cellulase loading: 0.63 FPU/ml, β-glucosidase loading: 5 U/mL at 45 ℃.

The conversion curves of pretreated cellulose shown in **Figure 29** are obtained after extensive washing prior to enzyme addition. The effect of MI in the enzymatic system was investigated with varied MI loadings ranging from 0 to 5%. According to the results in **Figure 30**, dosage of MI as low as 1% can be detrimental to the cellulolytic enzymes; beyond 1%, there is no detectable reducing sugar in the hydrolysate after 1 hour of hydrolysis. Therefore, elimination of residual MI in the pretreatment system is

crucial prior to enzymatic degradation due to the strong inhibitory effect of MI on the following step. The initial pretreatment conditions worked well for both cellulose substrates, showing MI pretreatment is a promising process operated at mild conditions.

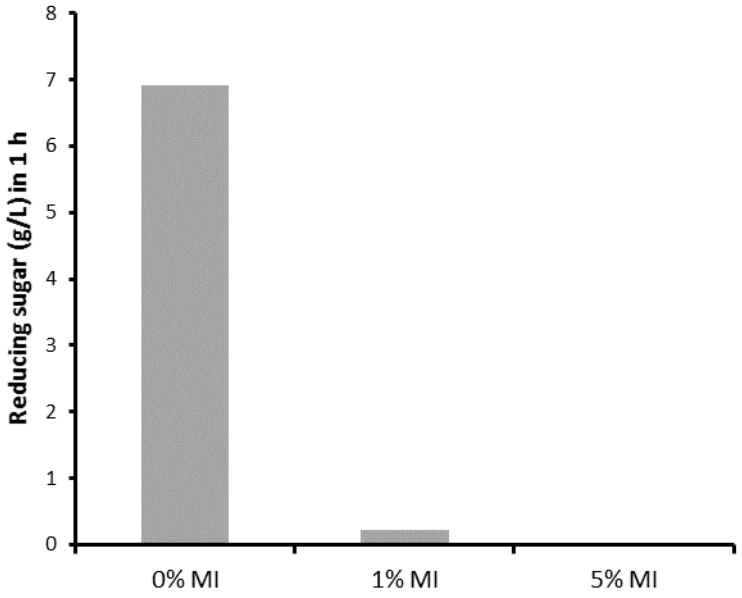


Figure 30. Inhibitory effect of MI on enzymatic hydrolysis of MI-pretreated Avicel. To study the reducing sugar release rate in the presence of MI, 0, 1 and 5 vol% MI was added into the enzymatic system for extensively washed MI-pretreated Avicel. Reaction conditions: 20 g/L Avicel in pH 5.0 NaOAc buffer at 45 ℃ with 0.63 FPU/ml cellulases and 5 U/ml β-glucosidase.

Next, the optimization of MI pretreatment conditions was performed. The three most relevant and independent factors determining pretreatment severity were chosen: MI concentration, pretreatment time and temperature. A two-level full factorial experimental design was carried out. **Table 15** shows the chosen low (-) and high (+) level for the three factors and the average of all observed 1-h hydrolysis rate at each level. All eight runs were duplicated to provide enough degrees of freedom for estimating residual error for

Analysis of Variance (ANOVA), which was computed with Minitab. The planning matrix and 1-h rate for the duplicates in each run are presented in **Table 16**. Based on the ANOVA results (not shown), at the significance level of 0.05, MI concentration and temperature are statistically significant (both p-value = 0.000) on the hydrolysis rate. However, pretreatment time is not a significant factor at the 0.05 level (p-value = 0.060). The fact that pretreatment time does not play a role (i.e. there is a negligible difference between 1 and 15 hours) is unexpected; we later compared 1 hour with 5 min, and observed very similar initial hydrolysis rate (results not shown), which indicates to us that the effect of MI on cellulose substrate is almost instantaneous. The hydrolysis outcome is dependent on MI concentration; the higher level resulted in higher average 1-h hydrolysis rate as expected. The MI pretreatment does not tolerate dilution by water: the presence of 20 vol% water in the mixture reduces the rate improvement effect from 55% to 24%; further addition of water dropped the effect down to ~10% (**Figure 31**). For several reasons, ambient temperature (25 ℃) is preferred compared to elevated temperatures (50 ℃). The decrease in improved hydrolysis rate can be possibly attributed to the increase of molecular motion at higher temperature, which lowers the potential of hydrogen bond formation. Thus, the optimal pretreatment condition within the studied range is 100% MI at 25 ℃ for 5 min and the following pretreatment experiments are conducted at these conditions.

Table 15. 3-Factor full factorial design at two levels

Factor	**Level**		**Average 1-h hydrolysis rate (g/L/h)**	
	-	+	-	+

Table 15 continued

MI conc. (vol %)	80	100	4.14	5.12
Time (h)	1	15	4.60	4.67
Temperature (℃)	25	50	4.75	4.51

Table 16. Planning matrix and observed 1-h hydrolysis rate

Run #	**MI conc.**	**Time**	**Temperature**	**Hydrolysis rate (g/L/h)**	
1	-	-	-	4.15	4.33
2	-	-	+	4.11	4.01
3	-	+	-	4.16	4.23
4	-	+	+	4.02	4.10
5	+	-	-	5.19	5.11
6	+	-	+	4.90	4.96
7	+	+	-	5.44	5.39
8	+	+	+	4.96	5.02

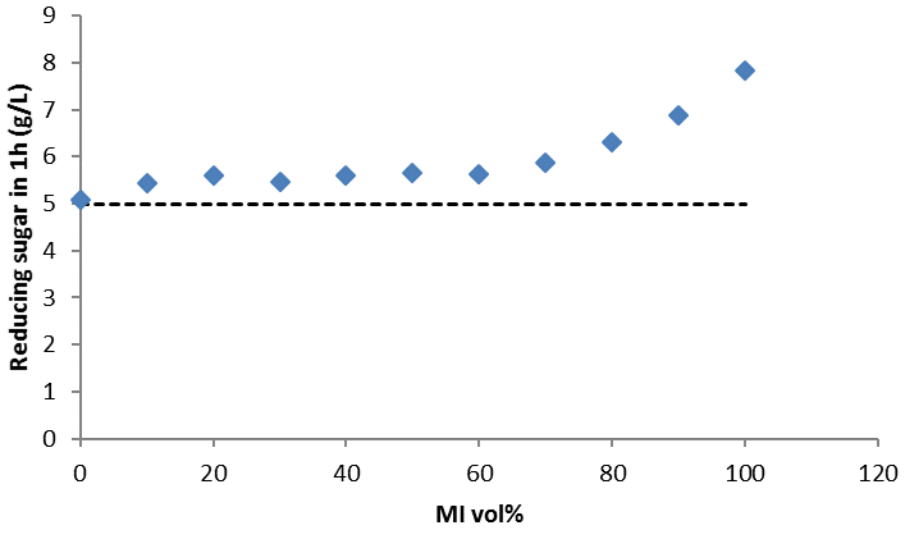


Figure 31. 1-h enzymatic hydrolysis rate for Avicel pretreated with various concentrations of MI

4.3.2 Understanding the effect of MI pretreatment on cellulose

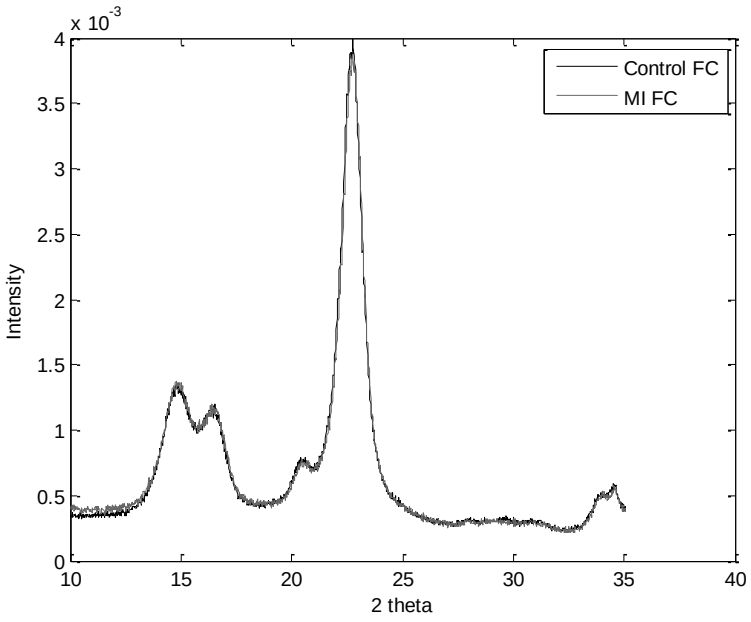
After establishing the accelerating effect of MI on cellulose hydrolysis rate, we sought to find the cause for such acceleration. We first hypothesized that MI can partially break down the crystal structure of cellulose to enable a higher hydrolysis rate. Crystallinity, characterized by the crystallinity index CrI is known to be a key factor governing the enzymatic rate (Hall, Bansal et al. 2010, Kang, Bansal et al. 2013). The XRD results, however, do not support this hypothesis as Avicel pretreated with NaOAc or MI exhibits similar diffraction patterns (results not shown). With the curving-fitting method we developed and published earlier (Kang, Bansal et al. 2013), the CrI of NaOAc- and MI-pretreated Avicel were determined to be 57.9% and 57.7%, respectively, and within the margin of error. XRD spectra of differently pretreated FC samples also support the conclusion (**Figure 32** (a)). Furthermore, no peak shift has been detected for

either Avicel or FC, indicating constant interlayer distances for the main cellulose crystalline phases identified from XRD spectra.

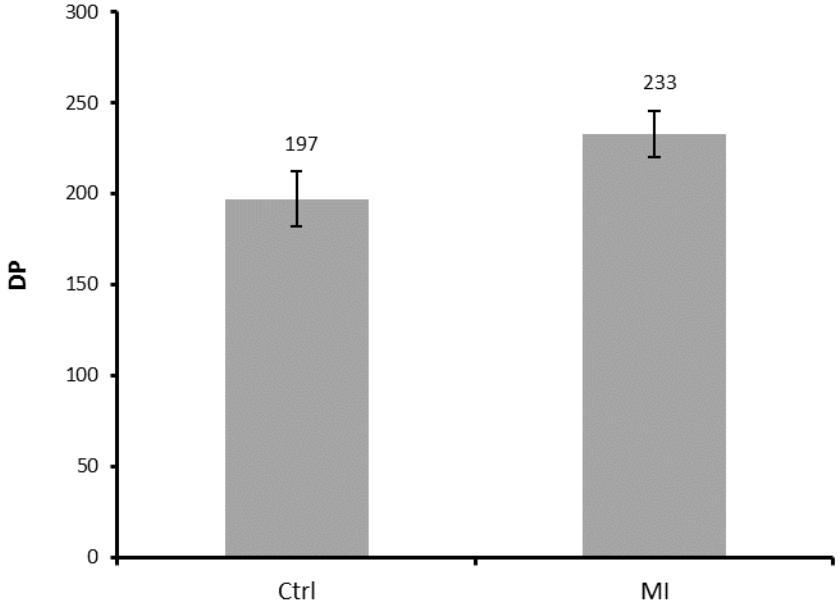
Degree of polymerization (DP) is another physical property of cellulose, which frequently found to impact cellulose degradation rate (Liu and Chen 2006). A reduced DP is generally related to hydrolysis rate improvement as a result of increased cellulose chain end concentration. Liu et al. reported that by combining steam explosion with IL pretreatment on switchgrass, cellulose DP was reduced, an observation that partly explained the increase in cellulose degradation rate (Liu and Chen 2006). Model simulation results also demonstrated an inverse relationship between substrate DP and hydrolysis rate (Kumar and Murthy 2013). However, the measured number-average DP with a published method (Zhang and Lynd 2005) revealed a higher DP from MI pretreatment compared to the control (**Figure 32** (b)). These results indicate the preferential dissolution of short cellobiose oligomers by MI, corresponding to a reduction in chain end concentration. Therefore, the rate enhancement observed during hydrolysis is not attributable to reduction in DP.

The surface area of cellulose substrate is generally agreed to be a crucial contributor to enzymatic hydrolysis rate as the cellulases' catalytic reaction is essentially a surface reaction (Arantes and Saddler 2010, Yang 2011). However, the total surface area does not usually reflect enzyme-accessible surface, as the bulky size of cellulase enzymes hinders their accessibility to smaller pores on the substrate surface.

(a)


x 10^-3
Control FC
MI FC
Intensity
2 theta
0
0.5
1
1.5
2
2.5
3
3.5
4
10
15
20
25
30
35
40

(b)


DP
197
233
Ctrl
MI
0
50
100
150
200
250
300

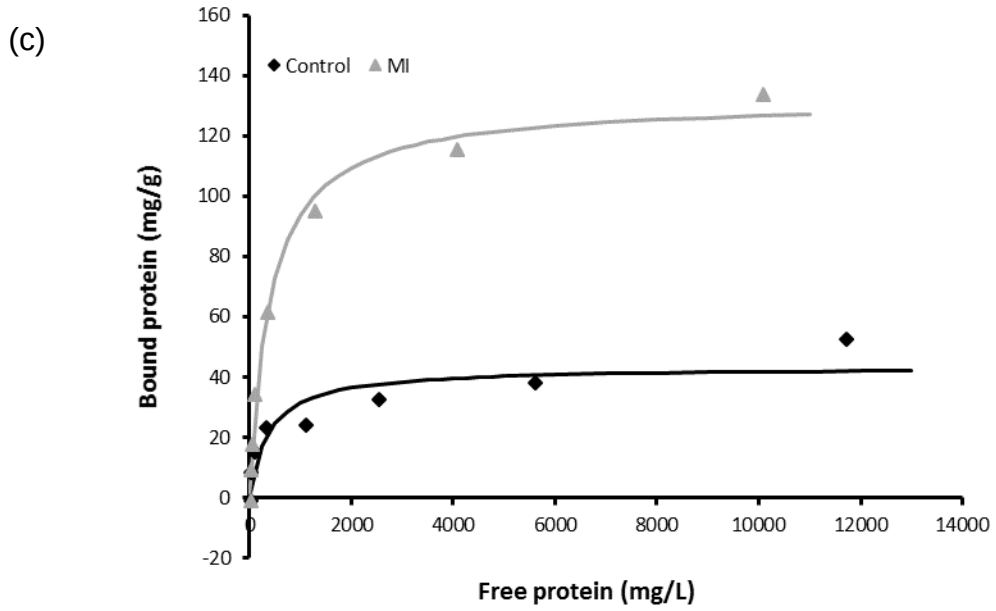


Figure 32. (a) XRD spectra for pretreated-FC. Black: control, Grey: MI. (b) Number-average DP values measured for control and MI-pretreated Avicel. (c) Adsorption isotherms of control and MI-pretreated Avicel. The respective curve shows a good non-linear fit with the Langmuir model of adsorption.

There are many assays available to estimate surface area/accessibility such as nitrogen gas adsorption (Lee, Teramoto et al. 2009), simon's staining (Arantes and Saddler 2010), fluorescent fusion protein (Hong, Ye et al. 2007, Rollin, Zhu et al. 2011) and carbohydrate-binding modules (Gourlay, Arantes et al. 2012). The most convenient and relevant way to assess the cellulose accessibility is the cellulase adsorption assay performed at low temperature (Bansal P., Vowell B. J. et al. 2012). Although the underlying assumptions of a Langmuir isotherm do not necessarily hold true in cellulase adsorption, the model still gives a reasonable description of the protein binding behavior in heterogeneous systems (Carta and Jungbauer 2010). The adsorption isotherms obtained for Avicel (**Figure 32** (c)) were non-linearly fitted with the Langmuir equation and the fitted parameters are listed in **Table 17**. B_{max} is the maximum binding capacity, which is

a measure of available surface area to cellulases and K_{ad} is the adsorption equilibrium constant reflecting the binding affinity of cellulases to cellulose. Comparing the fitted B_{max} and K_{ad} values for NaOAc- and MI-pretreated Avicel, the most significant difference lies in B_{max} while the binding affinity was effectively the same. MI pretreated Avicel can take up to 130 mg/g substrate; however, the control sample (NaOAc-pretreated) only attained 43 mg/g which is about 3-fold lower.

Table 17. Langmuir isotherm parameters determined by non-linear regression.

	B_{max} (mg/g)	K_{ad} (L/g)	R^2 (%)
Control	43.4	2.63E-3	85.2
MI	131.7	2.46E-3	98.8

The significant improvement in B_{max} is also indirectly reflected by the volume of the sediment after washing and gravity settling (**Figure 33**). Furthermore, scanning electron microscope (SEM) images were utilized to obtain microscopic morphological information (**Figure 34**). The MI-pretreated Avicel appeared slightly rougher at the edge of the particles. However, the difference observed under SEM is not as dramatic as the never-dried samples, possibly attributable to the drying process which causes collapse of porous structure.

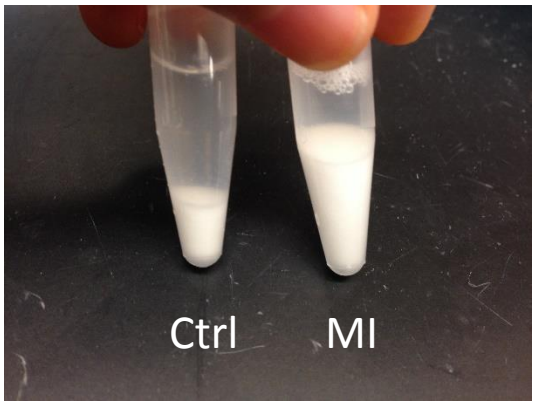


Figure 33. Control and MI pretreated-Avicel (20 g/L) after washing.

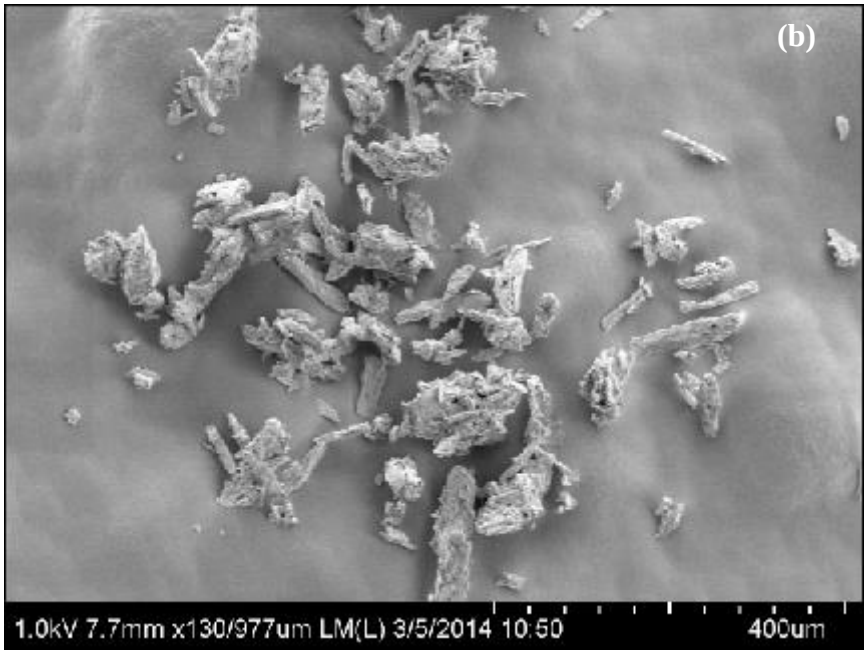
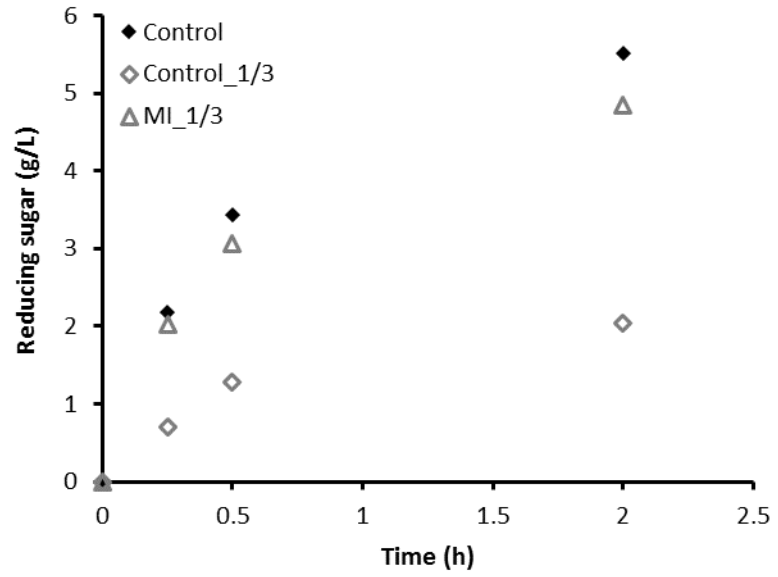


Figure 34. SEM images of control (a) and MI-pretreated (b) Avicel.

Upon decreasing the MI-pretreated Avicel concentration to 1/3 that of NaOAc, both reducing sugar-releasing curves superimpose on top of each other (**Figure 35** (a)), which is consistent with the 3-fold difference in accessible surface area identified from the binding isotherms. The control sample, however, showed a 3-fold reduction in sugar concentration (**Figure 35** (a)) but retained the same conversion yield (**Figure 35** (b)) as a result of the substrate concentration change. It can be inferred from these results that available surface area limits the hydrolysis rate; the released sugar concentration is proportional to the accessibility throughout the initial 2 hours hydrolysis and the conversion yield for MI-pretreated Avicel exhibits about 200% improvement under the studied conditions. Based on these findings, the observed rate enhancement on MI-pretreated cellulose can be attributed to the threefold higher accessible surface area.

(a)



(b)

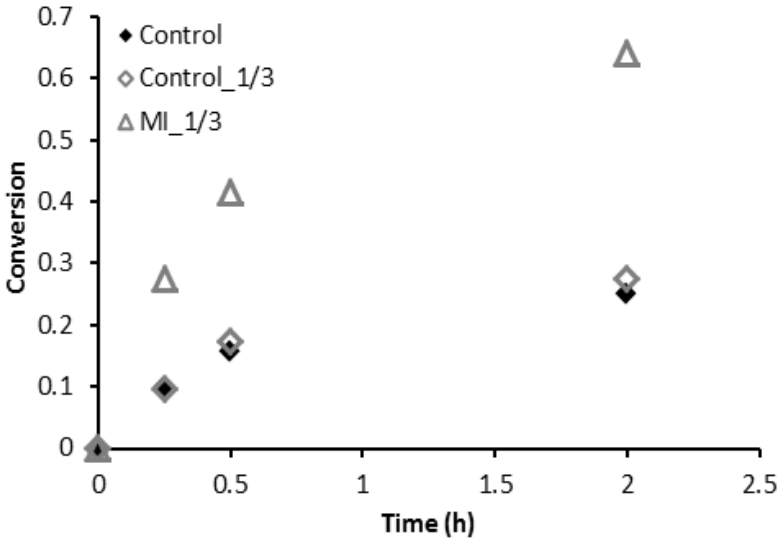


Figure 35. The effect of varying pretreated substrate concentration on (a) time course of reducing sugar release (b) conversion curve. Reaction conditions: 20 g/L (control) or 6.7 g/L (control_1/3 and MI_1/3) control or MI-pretreated Avicel in pH 5.0 NaOAc buffer at 45°C with 0.63 FPU/mL cellulases and 5 U/mL β-glucosidase.

To further corroborate the hypothesis that available surface area is a key parameter controlling the reaction rate, MI pretreatment was conducted on phosphoric

acid swollen cellulose (PASC) to investigate the effect on amorphous cellulose. In **Figure 36**, there is no difference in the conversion yield for the PASC with or without MI pretreatment. This provides further supporting evidence for the belief that MI is functioning as a "surface expander" on the cellulose substrate. Once a substrate is already made fully accessible to enzymes, pretreatment with MI will not render any additional benefit to the ensuing hydrolysis, nor suffer any detrimental effect, provided MI is completely removed before enzyme addition.

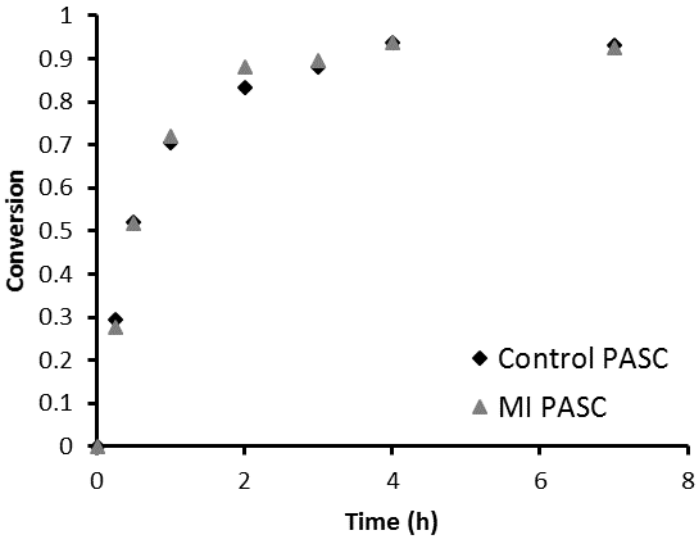


Figure 36. Effect of MI on PASC. Pretreatment conditions: 20 g/L PASC in pH 5.0 NaOAc buffer or 100% MI at 37 ℃ for 5 min. Pretreated PASC was extensively washed with DI water, the washed samples were subjected to enzymatic degradation with 0.63 FPU/mL cellulases and 5 U/mL β-glucosidase loadings at 47 ℃.

4.3.3 Further optimization of MI pretreatment

As demonstrated earlier, the pretreatment effect of MI is maximized in pure (100%) MI. A slight reduction in MI concentration can significantly impact pretreatment

outcome. Therefore, to reduce MI consumption per unit of biomass, a higher solid-to-liquid ratio during the process is desired.

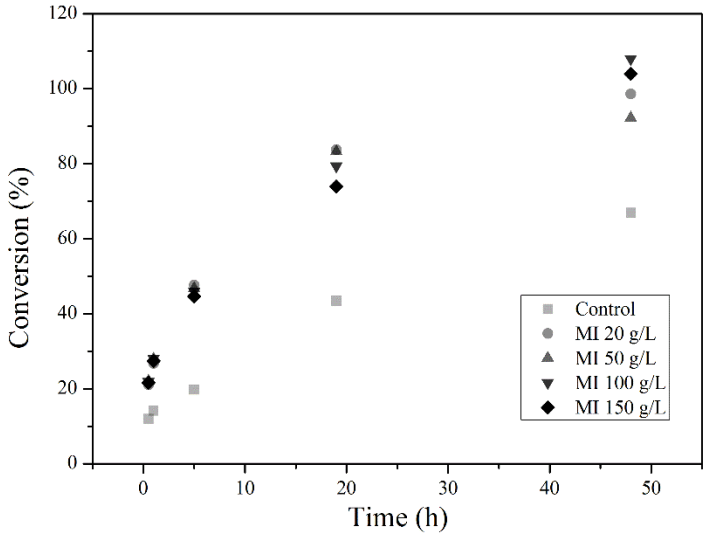


Figure 37. The enzymatic conversion curves for Avicel pretreated at different solid loadings in 100% MI. Pretreatment conditions: 20 – 150 g/L Avicel in 100% MI at 25 ℃ for 5 min; control was 20 g/L Avicel pretreated with pH 5.0 NaOAc buffer under the same conditions. The pretreated Avicel was extensively washed and subjected to enzymatic degradation with 0.63 FPU/ml cellulases and 5 U/ml β-glucosidase at 20 g/L substrate concentration at 45 ℃.

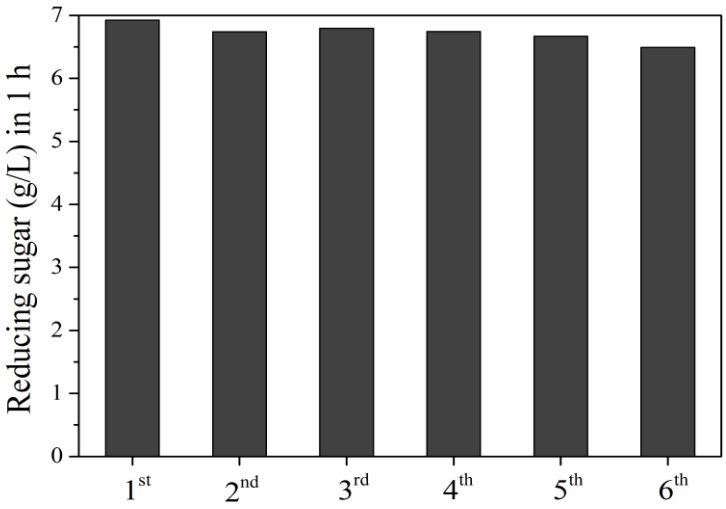


Figure 38. The reducing sugar released after 1 hour hydrolysis from six successive batches of Avicel pretreated with the same batch of MI. Pretreatment conditions: 40 g/L Avicel in fresh (1^{st}) or recycled (2^{nd} – 6^{th}) 100% MI at 25 ℃ for 5 min.

However, higher solid-to-liquid ratios may create additional problems such as inadequate mixing. Still, we found that by varying substrate concentration from 20 to 150 g/L, the MI efficiency on cellulose is not affected according to the subsequent conversion results (**Figure 37**). Based on this observation, it can be inferred that the action of MI on cellulose is controlled by the rate of penetration into cellulose particles (internal mass transfer) rather than the mixing of particles (external mass transfer).

To further improve MI pretreatment, we also studied the reusability of spent MI. As MI dissolves negligible amounts (<< 5 wt%) of cellulose under the mild pretreatment conditions, the supernatant separated from pretreated cellulose contained mainly MI and there is nearly no visual difference between spent and fresh MI. **Figure 38** shows the initial 1 hour rate for six consecutive batches of Avicel after re-utilizing the same batch of MI for pretreatment. There is no evidence that reusing MI for six successive cycles

diminished the effectiveness of MI on Avicel, especially for the initial five cycles which shown less than 4% reduction in hydrolysis (**Figure 38**). Therefore, the pretreatment method stands an improved chance to be economical through re-utilization of spent MI on cellulose.

4.3.4 Effect of MI on lignocellulosic biomass

To study the effect of MI pretreatment on realistic biomass, raw loblolly pine (LP) was employed as the model lignocellulosic substrate. The same condition at 25℃ for 5 min was applied and no significant changes can be discerned visually. The enzymatic conversion results showed 92% to 230% rate enhancement compared to control for the entire conversion process for up to 102 hours (**Figure 39**). Furthermore, the model LP sample has an average particle size of ~300 μm which is substantially larger than Avicel paritcles (~50 μm). Further improvement of LP digestibility can be expected with particle size reduction. Although remarkable rate enhancement was achieved with MI pretreatment, the conversion yield for pretreated LP is still low compare to the leading pretreatment methods typically performed under severe conditions. To further investigate and understand the MI effect under ambient conditions, which is likely masked by the presence of complex plant polymers, lignin-carbohydrate complexes (LCCs) and poor accessibility, we extended the study to steam-exploded biomass with an open structure.

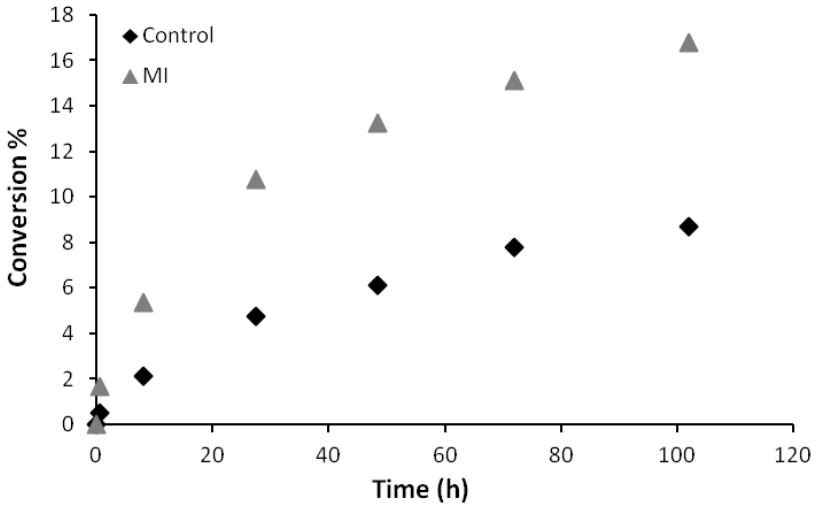


Figure 39. The enzymatic conversion curves for 100% MI and NaOAc buffer pretreated raw loblolly pine. Pretreatment conditions: 20 g/L LP in pH 5.0 NaOAc buffer or 100% MI at 25 ℃ for 5 min. Pretreated LP was extensively washed with DI water, the washed samples were subjected to enzymatic degradation with 0.63 FPU/mL cellulases and 5 U/mL β-glucosidase.

Before moving on to steam-exploded lignocelluloses, a mechanical mixture of Avicel and commercial sugarcane bagasse lignin (SBL) was prepared and utilized for 100% MI pretreatment and subsequent enzymatic degradation to investigate the effect of lignin on the process. After 5 min pretreatment at 25 ℃, a significant amount of lignin dissolution was observed (**Figure 40**). In the MI-pretreated Avicel/SBL mixture, the majority of SBL was preferentially and instantaneously dissolved by MI while Avicel in the system was conserved. Negligible dissolution occurred during the NaOAc buffer pretreatment. The action of MI on lignin and cellulose is similar to monoethanolamine (MEA), which is known as a selective delignifying agent. To extract lignin from raw woody biomass with MEA, much more severe conditions should be employed (Claus, Kordsachia et al. 2004).

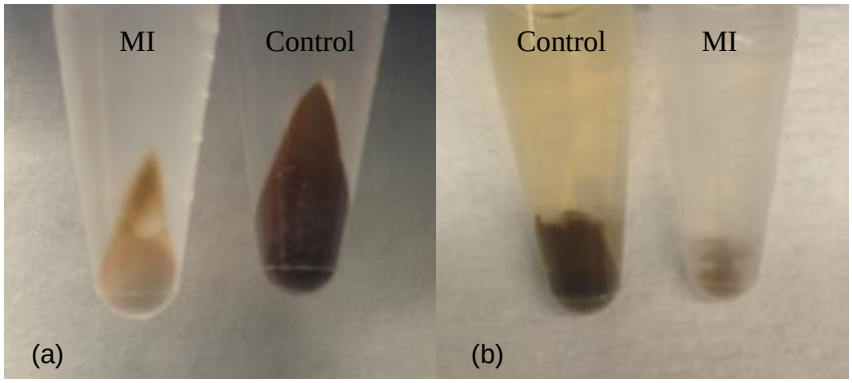


Figure 40. Mechanically mixed 1:1 mass ratio of Avicel and SBL after (a) control or MI pretreatment and extensive wash step (b) pretreatment, wash and extended enzymatic hydrolysis step. The remaining solid was un-dissolved lignin during pretreatment.

The behavior of lignin and cellulose dissolution in MI differs vastly from each other, thus the MI concentration effect was revisited and results are shown in **Figure 41**. A comparison study of steam-exploded sugarcane bagasse (SEB) and SBL dissolution in varying MI concentrations ranging from 0 to 100% was performed. As expected, the higher concentration renders higher amounts of dissolved lignin. For the case of SBL, a steep increase in lignin dissolution capacity was observed from 0% to 10% MI. More than 95% SBL was dissolved in 10% MI after a short incubation time (i.e. 5 min) at room temperature. The saturation point was reached at around 10%, beyond which no additional dissolution effect was observed with rising MI concentration. For SEB, there seems to exist a threshold MI concentration (i.e. 5%), lower than which negligible lignin dissolution was achieved compare to 0% control. In addition, the maximum dissolvable lignin fraction in SEB under those incubation conditions is around 60%, which is significantly lower than SBL. These results suggest lignin dissolution is more water-

tolerant in SEB and lignin embedded in SEB was not as easily dissolvable as in SBL, probably due to the limited accessibility and structural/compositional difference of lignin.

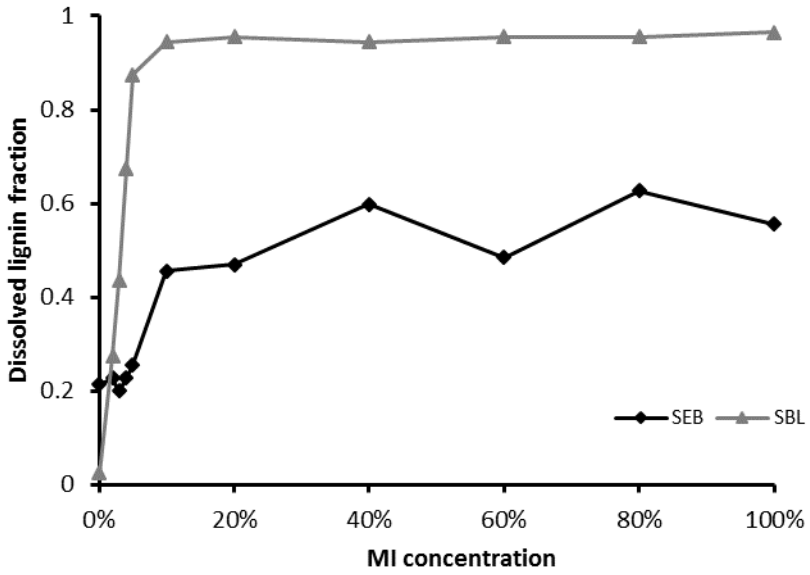


Figure 41. Dissolved lignin faction (based on total lignin content in respective sample) from SEB and SBL during pretreatment with varied MI concentration. Pretreatment conditions: 40 g/L substrate in MI at 25 ℃ for 5 min. The lignin dissolution was calculated based on mass balance.

Since the lignin dissolution was probably controlled by MI concentration, lignin composition and degree of condensation, the MI concentration was fixed at 100% for steam-exploded lignocellulose pretreatment to study the different effect of MI on different substrates. **Figure 42** shows the lignin removal efficiency on SBL and three differently steam-exploded substrates at different solid loadings. Three solid loadings ranging from 20 to 100 g/L were studied in this work; the higher solid loading leads to a slight decrease in the percentage of dissolved lignin. Among the four substrates, SBL is

the most readily dissolvable at all loadings: more than 94% SBL was quickly dissolved under the mild condition (25℃ for 5 min), even at 100 g/L. Besides, solid loading does not play a role in lignin dissolution for SBL. The dissolved amount is very similar for all three loadings, which is indicative of the absence of external mass transfer limitation and good accessibility of solute. In contrast, the percentage of dissolved lignin from steam-pretreated agricultural residues SEB and steam-exploded wheat straw (SEWS) is slightly lower than SBL and a downward trend was observed as solid loading increased. The results indicate that at a given MI concentration lignin dissolution is controlled by both its accessibility to solvent as well as mass transfer. For the woody biomass steam-exploded loblolly pine (SELP), the lignin embedded is much more resistant and less accessible to the solvent MI and lignin composition probably also plays a role in its dissolution in solvents.

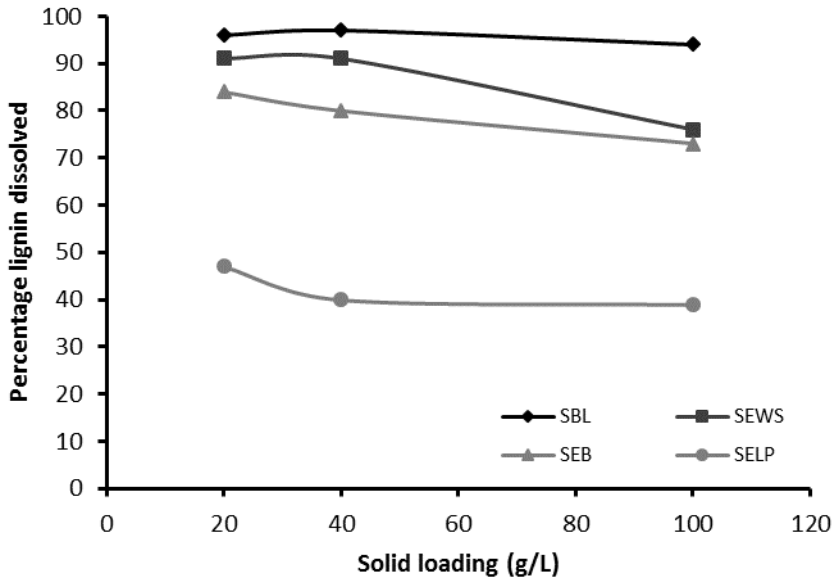


Figure 42. Percentage of lignin removed in dissolved form during MI pretreatment for different lignin-containing substrates.

The crystalline structure of the pretreated-biomass was analyzed by X-ray diffractometry (XRD) before and after MI pretreatment (**Figure 43**). The most apparent change in the XRD spectra is the increase in diffraction intensities around 22.5 ° which is the characteristic position for the 200-phase of cellulose crystal. The 200 peak intensity change is usually a sign of crystal structure modification of cellulose during the pretreatment and/or change in relative content of crystalline and amorphous components in the sample.

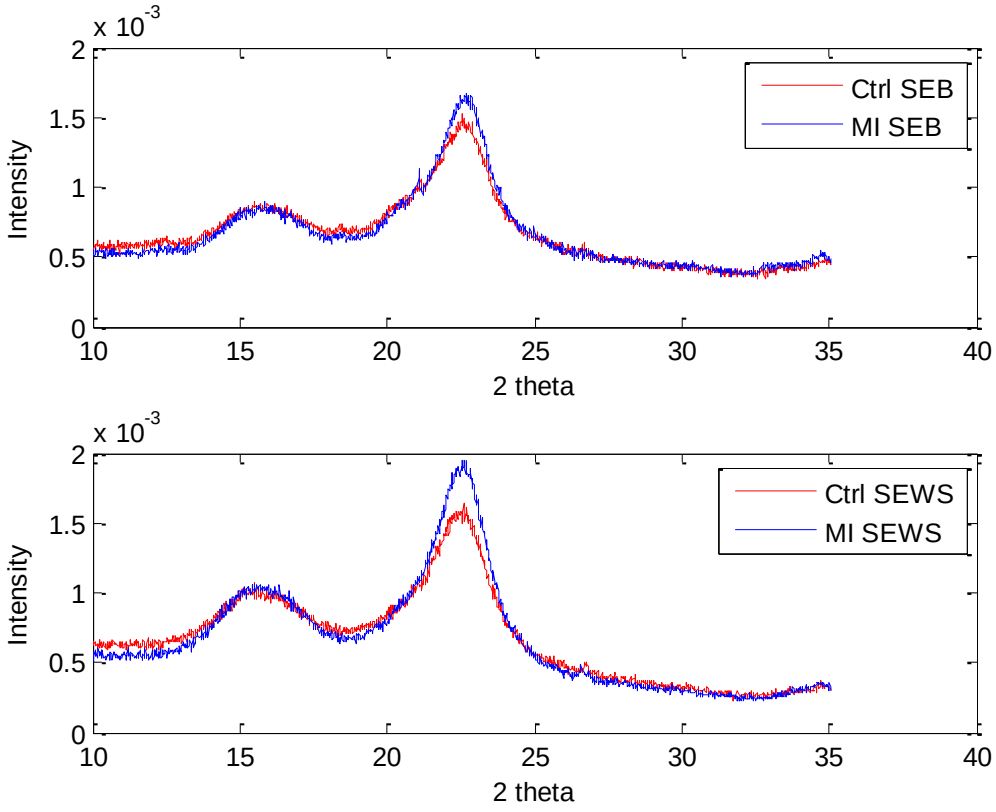


Figure 43. XRD spectra for SEB and SEWS samples pretreated with MI or control.

To quantify the CrI for each sample, a previously established curve-fitting method was utilized and the calculated results are tabulated in **Table 18** (Kang, Bansal et al. 2013). The overall CrI of MI-pretreated biomass is higher than that of the control due to the dissolution of amorphous lignin, as a result of which the relative cellulose content is increased. Comparing to overall CrI, $CrI_{cellulose}$ is a more relevant parameter that is calculated as overall CrI divided by cellulose content for individual sample. The calculated $CrI_{cellulose}$ for both SEB and SEWS are slightly lowered (both decreased by less than 5%) after MI pretreatment, however, the decrystallization effect is not significant judging from XRD results. To further understand the potential changes in biomass

structure induced by MI, other analytical techniques such as solid-state NMR and FTIR should be utilized.

Table 18. Calculated overall crystallinity (CrI) and cellulose crystallinity ($CrI_{cellulose}$) from XRD spectra.

	Ctrl SEB	**MI SEB**	**Ctrl SEWS**	**MI SEWS**
CrI %	21.9	27.5	27.4	37.6
Cellulose %	55.4	78.5	56.5	80.8
$CrI_{cellulose}$ %	39.5	35.0	48.5	46.6

Lignin can form both a chemical and a physical barrier to the enzymatic degradation through non-productive binding and physical blockage to impair the accessibility of cellulase to cellulose. The most desired outcome in pretreatment is to completely remove lignin but conserve polysaccharides, expose and expand the cellulose surface, and decrystallize cellulose structure. Since the majority of the lignin can be efficiently extracted from the steam-pretreated agricultural biomass, the next step is to study the enzymatic degradation efficiency of the MI-pretreated samples. **Figure 44** shows the conversion profiles. For the woody biomass SELP, the MI pretreatment does not have a substantial positive effect. Within 30 min, the rate is 20% higher compare to the control. However, after 30 min, the reducing sugar releasing yield for MI-pretreated SELP was surpassed by the control. For the case of agricultural residues (i.e. SEB and SEWS), MI showed promising rate-enhancing effect for the entire time frame investigated. These results suggest MI is not as effective on woody biomass as on

agricultural residues (**Figure 44**). However, within the 30 min hydrolysis time mentioned earlier, the 20% rate improvement for MI-pretreated SELP is reproducible and consistent, which led us to believe there are other interfering factors that complicate the interpretation of hydrolysis results.

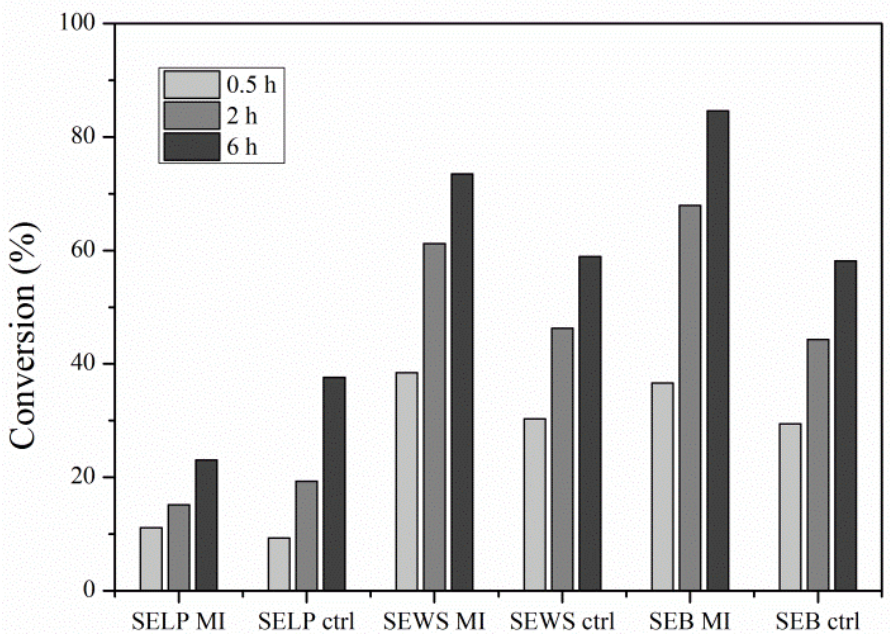


Figure 44. Cellulose conversion from MI-pretreated lignocellulosic biomass.

During the pretreatment step, the pores in the cell wall and thus the overall surface area expand, thus enhanced accessibility for both lignin and cellulose can be achieved. The effect of MI on lignin-containing biomass can be both beneficial and detrimental as the increased non-productive binding is balanced by improved productive binding and lignin removal. The ultimate pretreatment outcome depends on the relative extent of lignin removal and cellulose expansion. **Table 19** summarizes the lignin content

for the control and after MI pretreatment. As expected, SELP contains more lignin for both cases; after MI pretreatment, the residual lignin content is almost double that of both SEWS and SEB. It is apparent that lignin removal by MI from woody biomass is far less effective than from non-woody biomass; therefore, the enhanced cellulose accessibility cannot compensate for the increased non-productive binding onto residual lignin during hydrolysis. The reason we observed a rate enhancement at 30 min is probably due to the slower lignin adsorption kinetics compared to cellulose (Zheng, Zhang et al. 2013). However, at equilibrium, the conversion yield is exceeded by the control due to the high adsorption capacity of MI-pretreated lignin.

Table 19. Lignin content (based on total mass of respective sample) in lignocellulosic biomass after control and MI pretreatment.

	SELP	**SEWS**	**SEB**
Control (%)	52.8	33.0	35.1
MI (%)	40.2	19.2	21.5

The dramatic increase in cellulase non-productive binding to lignin for SELP is also supported by the SDS-PAGE total protein analysis of hydrolysate (**Figure 45**). After 96 hours of enzymatic degradation on MI-pretreated and control SELP (20 g/L), the residual solid in the system is predominantly lignin. Although the residual lignin content in the control sample is about 30% higher than in the case of MI pretreatment (**Table 19**), the amount of bound cellulases on lignin is much lower. 12-fold more cellulases remained in free form compared to MI (**Figure 45**, data not shown), indicating a

substantial difference in cellulase affinity to the lignin from differently-pretreated SELP. In **Figure 45**, the EGI and EGII bands are not discernable in the lane for MI-pretreated SELP; CBHI, accounting for the thickest band in the control lane with more than 60% of total protein secreted from *T. reesei*, almost completely disappeared in the case of MI, indicating a high binding capacity of CBHI onto residual lignin.

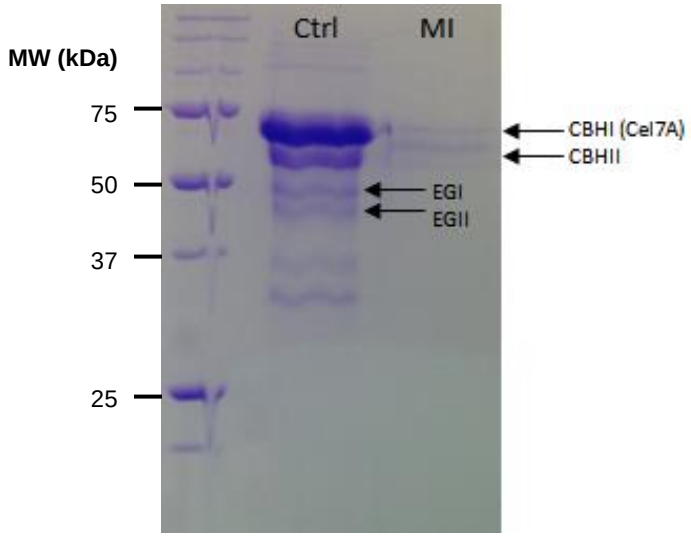


Figure 45. Analysis of 96 hour hydrolysates from MI and control pretreated SELP.

To study the reusability of MI on lignin-containing materials, six consecutive batches of SEB were tested. Upon direct reuse of MI, the rate improvement effect continuously dropped to below 10% in comparison to the control (**Table 20**). This necessitates an additional recovery step in order to separate MI from the dissolved lignin for reuse.

Table 20. Effect of reused MI on six consecutive batches of SEB.

# of reuse	0	1	2	3	4	5
1-h rate enhancement %	22.7	12.1	12.7	11.7	9.5	7.0

4.3.5 Recovery of MI from spent liquor

To reduce chemical consumption and pretreatment cost, recycling of MI is required and lignin needs to be precipitated from the spent liquor. Since lignin is extremely hydrophobic, acid was chosen to protonate MI to increase its hydrophilicity and thus reduce affinity toward dissolved lignin. A comparison study of utilizing hydrochloric acid and sulfuric acid as anti-solvents was performed (**Table 21**). Both acids precipitate lignin equally well for SEB; however, a dramatic difference was noted for SEWS and SBL. Sulfuric acid is capable of precipitating more than 53% dissolved lignin from the SBL/MI supernatant while a very small amount of lignin precipitation, which could not be pelletized, was observed for the case of SEWS. With hydrochloric acid, the amount of precipitated lignin from MI supernatant is comparable for SEB and SEWS; however, HCl is not as effective an anti-solvent while dealing with SBL, where only 15% lignin was precipitated and pelletized. Therefore, the effectiveness of different acids on dissolved lignin precipitation depends strongly on the biomass species. HCl is good for both agricultural residues we investigated, while H_2SO_4 is preferred for pure lignin SBL. All three lignin-rich MI can be precipitated with acids to extract around 50% of the dissolved lignin. Further work on separation and purification is still needed to generate relatively pure recycled MI. **Figure 46** shows the schematic representation of a MI-based pretreatment process.

Table 21. Comparison of precipitated lignin fraction (based on the total lignin dissolved) out of spent liquors from MI pretreatment systems with two acids.

	SEB	SEWS	SBL
H_2SO_4 (%)	47.2	NA*	53.1
HCl (%)	47.2	45.5	14.9

* Precipitated lignin fraction was not available as a small amount of lignin precipitation which cannot be pelletized was observed

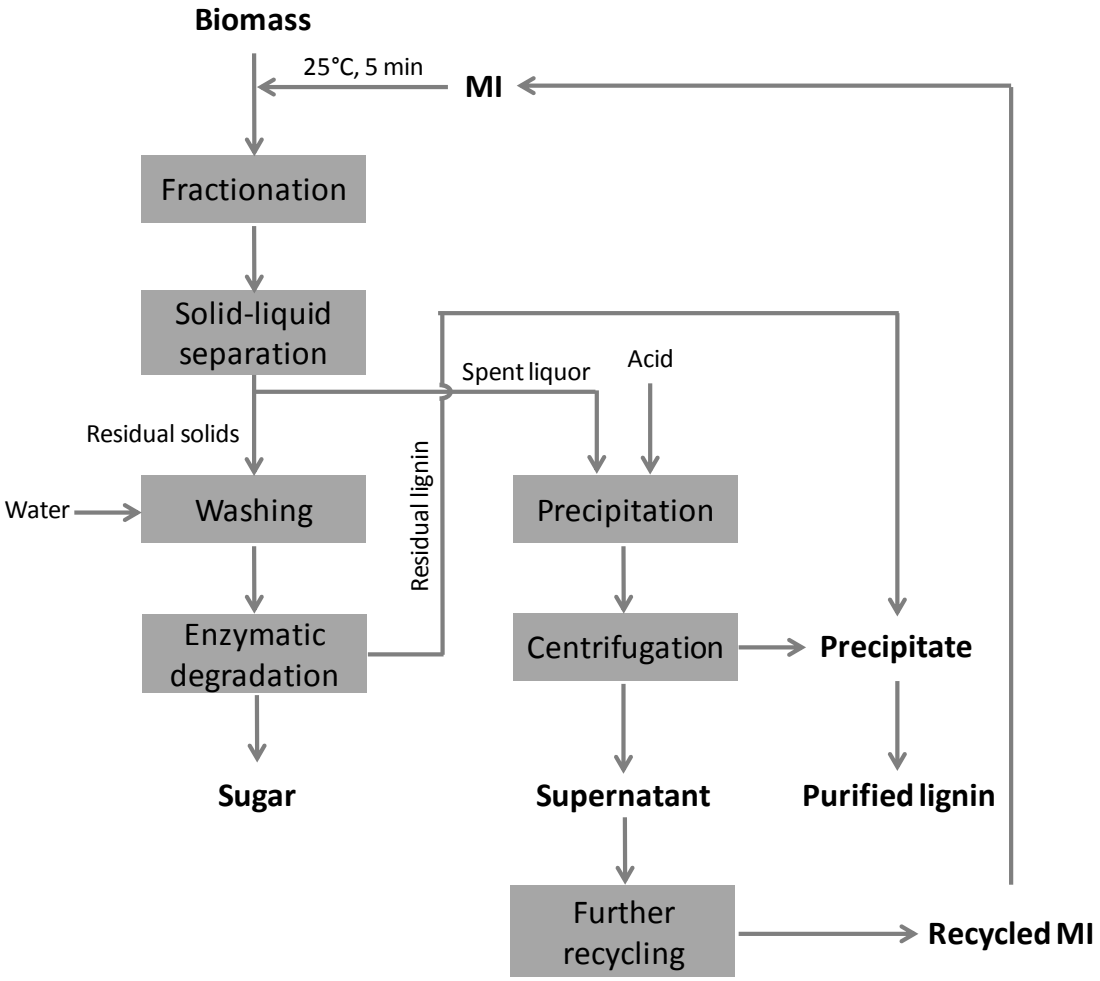


Figure 46. Schematic representation of a MI-based pretreatment process for biomass.

4.4 Conclusions

In the present work, MI is shown to be an efficient pretreatment agent for both cellulose and lignocelluloses substrates. Significant hydrolysis rate enhancement was observed for a variety of both lignin-containing and lignin-free substrates at ambient pretreatment conditions. MI is found to conserve cellulose as well as its crystal structure. The effectiveness of MI on model cellulose substrates was hypothesized and demonstrated to be due to the penetration into and expansion of the structure at molecular level to dramatically improve available surface area, which was found to be proportional to the rate of sugar release.

On the other hand, the effect of MI on lignocelluloses is a combined result of efficient delignification and cellulose expansion. Lignin can be effectively extracted from model lignin-containing substrates at mild conditions without discernable dissolution, degradation or decrystallization of cellulose. The lignin removal increases accessibility and reduces physical barriers to cellulose, which in turn promotes the enzymatic hydrolysis rate.

To further improve the pretreatment process and explore its potential industrial applications, chemical recycling and process-related factors need to be optimized and tailored for different substrates. A more detailed study of process development is needed and ongoing.

CHAPTER 5

DIRECT DYES AS MOLECULAR PROBES FOR THE MEASUREMENT OF CELLULOSE ACCESSIBILITY AND PORE SIZE DISTRIBUTION

5.1 Introduction

To date, there is no agreement on the rate-limiting parameters that control the enzymatic hydrolysis rate of cellulose-containing substrates to produce monomeric sugar (Kang, Bansal et al. 2013). In fact, cellulose properties such as crystallinity (Hall, Bansal et al. 2010, Kang, Bansal et al. 2013), accessibility (Jeoh, Ishizawa et al. 2007, Rollin, Zhu et al. 2011), degree of polymerization (DP) (Puri 1984, Zhang and Lynd 2004) have all been considered to be the key parameters affecting enzymatic kinetics. However, many cellulose properties are not stand-alone concepts and they could be related to, or depend on, one another. Besides, the determining factors for enzymatic step are likely to change with respect to the pretreatment method, origin of biomass, compositional differences, etc. Although limited knowledge is available and no conclusive statement can be made (Mansfield, Mooney et al. 1999), the accessibility of cellulose is valuable information to obtain due to the heterogeneous nature of the reaction.

Numerous studies have examined the accessibility of cellulose-containing substrates by different means. All the techniques quantify accessibility by direct contact of probe molecules with substrates. A wide selection of molecular probes have been tested including water (i.e. water retention value), dextran solutes (i.e. solute exclusion),

iodine, nitrogen, direct dyes, cellulase, cellulose-binding domain-containing fusion proteins (Inglesby and Zeronian 1996, Mansfield, Mooney et al. 1999, Hong, Ye et al. 2007). Among the different probe molecules, direct dye staining is a promising technique that could eliminate the possible problems associated with the remaining methods, such as non-specific binding and procedural complexity and length (Rahikainen, Mikander et al. 2011).

The employment of direct dyes to characterize cellulose surface properties has received considerable attention (Morton 1935). A variety of direct dyes, such as Direct Red 2, Direct Red 28, Direct Blue 1 and Direct Yellow 4, have long been recognized to bind selectively and strongly on cellulose substrates (Inglesby and Zeronian 2002). Due to the effectiveness and ease of use, various direct dyes were frequently employed as molecular sensors to characterize cellulose for textile and pulp and paper studies (Chandra, Ewanick et al. 2008).

Single direct dye with low molecular weight and volume is effective at evaluating the internal surface area of cellulose, however, the probed area is not necessarily accessible to cellulolytic enzymes with much larger sizes. Thus, it is frequently found the total surface area and enzymatic activity is not directly correlated.

The Simons' staining is a two color differential staining assay which appeared in 1950 for the investigation of cellulose accessibility to cellulases. It was first developed for the pulp and paper industry as a microscopic assessment of the degree of fibrillation for beaten fibers (Simons 1950). The original Simons' staining assay utilizes a dye mixture consisting of a blue and a polymeric orange dye to determine the extent of fiber damage. The blue dye Direct Blue 1 (DB1) has a lower molecular weight and binding

affinity compared to the orange dye Direct Orange 15 (DO15) with molecular size similar to cellulase enzymes. After dyeing, the smaller cellulose pores will be populated with DB1 which has an estimated molecular volume of 697 $Å^3$ (Inglesby, Zeronian et al. 2002); while DO15, a condensation product of 5-nitro-O-toluenesulfonic acid under alkali condition (Yu, Minor et al. 1995), occupies larger pores and cellulose surface. As a result of their different binding affinities, the fibers appear blue in color for mechanical and non-beaten pulp; for well exposed fibers, they stain orange (Chandra, Ewanick et al. 2008).

With the additional orange dye, the Simons' staining measurement were reported to directly correlate with enzymatic activity (Esteghlalian, Bilodeau et al. 2001). The molecular structure of polymeric DO15 is ill-defined, it is known to be comprised of low molecular weight (LMW) and high molecular weight (HMW) fractions. The latter was found to exhibit stronger adsorption due to the existence of abundant available binding site per molecule (Yu, Minor et al. 1995). Although only HMW DO15 is utilized in the modified Simons' stain protocol, two sub-fractions with hydrodynamic sizes of 5-7 nm and 14-36 nm were identified. The polydispersion of the orange dye molecules is likely to render them different binding behaviors with some closer to cellulase. Besides, the DO15 is not commercially available from the US vendors anymore, therefore it is time to seek an alternative technique to estimate the cellulose accessibility and pore size distribution, and finally to predict the response to enzymatic hydrolysis. In this chapter, we used a small direct dye probe in combination with cellulase to investigate the microscopic cellulose structure.

5.2 Materials and methods

Materials

Avicel PH-101 (Fluka 11365) and fibrous cellulose (FC) from cotton linters (Sigma C6288, medium) were purchased from Sigma-Aldrich (St Louis, MO, USA) and employed as model cellulose substrates in this study. DB1 and DV9 dyes were provided by Pylam Products (Garden City, NY). Three differently-pretreated steam-exploded loblolly pine (SELP) samples were provided by Dr. John Muzzy from the Georgia Institute of Technology, Atlanta, GA, USA. Cellulase from *Trichoderma reesei* (Celluclast®, 159 FPU/ml) and β-glucosidase from almonds (5 U/mg) were purchased from Sigma-Aldrich and phosphoric acid (85%) from EMD (Gibbstown, NJ, USA). The BCA protein assay kit was purchased from Thermo Fischer Scientific (Rockford, IL, USA).

Preparation of phosphoric acid swollen Avicel and FC (PASC and PAFC)

PA treatment of Avicel and FC were carried out following the procedure described by M. Hall et al. (Hall, Bansal et al. 2010). Thirty mL of ice-cold phosphoric acid with desired concentration was used to pretreat 1 gram slightly hydrated Avicel or FC for 40 min with occational stirring. After the treatment, 30 ml of ice-cold acetone was added with stirring. The resulting regenerated slurry was filtered with a fritted filter funnel and washed with 20 ml ice-cold acetone twice. The acetone-washed slurry was further washed with 100 ml DI water four times and freeze-dried for X-ray diffraction analysis and staining experiments.

Preparation of partially converted cellulose

Partially converted cellulose was prepared following the procedure described by P. Bansal et al. (Bansal P., Vowell B. J. et al. 2012). Twenty g/L Avicel was pre-incubated in sodium acetate buffer (NaOAc, pH 5.0, 50 mM) for 1 h at 50 ℃. The enzymatic hydrolysis was initiated by adding Celluclast® cocktail 150 FPU/g and β-glucosidase 15 kU/L. After desired amount of time, the hydrolysis was quenched and Avicel with desired conversion level was collected by centrifugation. The bound cellulase enzymes were removed from cellulose surface with a procedure modified from a published one (Hong, Ye et al. 2007, Bansal P., Vowell B. J. et al. 2012). In brief, the partially converted Avicel was resuspended in 1.1% SDS and incubated for 15 min at 80 ℃, followed by washing with 75% ethanol three times and four times with a DI water wash. The washed cellulose slurry free of enzymes was freeze-dried for X-ray diffraction analysis and Simons' staining experiments.

Staining procedure

The staining assay was carried out with Direct Violet 9 (DV9) or Direct Blue 1 (DB1) dye. Stock solutions of each dye with known concentration (around 10 g/L) were prepared by dissolving in DI water. The staining step was performed according to a modified protocol (Chandra, Ewanick et al. 2008). To each tube containing cellulose substrate and 100 µl phosphate-buffered saline (PBS, pH 6, 0.3 M PO_4, 1.4 M NaCl), increasing amount of stock solution (0 to 200 µl) was added and DI water was used to bring up the final volume to 1 ml. The final dye concentration ranges from 0 to 2 g/L. The mixture was incubated at 70 ℃ and 1000 rpm for 6 hours. After the incubation period, the supernatant was separated and absorbance was read at 624 nm. The free dye

concentration was calculated with the respective extinction coefficient derived from calibration curves at 624 nm (**Figure 60**).

Protein adsorption

Substrates at 2% consistency were preincubated in 1 ml NaOAc buffer (50 mM, pH 5.0) at 45 ℃, 1000 rpm for 1 h followed by cooling on ice. Various amounts (0.0048 – 7.6 FPU/ml) of cellulase were added and further agitated for 30 min at 4 ℃ to minimize the influence from hydrolysis. Following 30 min adsorption, all samples were immediately centrifuged for 3 min at 4 ℃ and 18,625 g. The control group went through the same process without substrate. The supernatant was subjected to BCA analysis to determine the protein concentration based on BSA standards.

X-ray diffraction (XRD)

XRD patterns of lyophilized samples were recorded with a X'pert PRO X-ray diffractometer (PANanalytical BV, Almelo, the Netherlands) using Cu/Kα_1 irradiation (1.54 Å) at 45 kV and 40 mA. The scattering angle (2θ) ranged from 10 ° to 40 ° with a scan speed 0.021425 s^{-1} and step size 0.0167 °.

5.3 Results and discussion

5.3.1 Binding specificity of direct dyes

In this work, two direct dyes DB1 and DV9 were tested for their suitability to be utilized as molecular probes for estimation of the internal surface area of the cellulose substrates. **Table 22** shows their color index numbers and molecular weights. The DB1 dye is traditionally employed in Simons' staining, in combination with the polymeric DO15, to act as the smaller molecular probe that is known to be capable of penetrating to

the capillary internal cellulose structure and permitting a rapid estimation of total accessibility. The DV9 dye has not been studied before for this purpose, so it was first characterized for its binding specificity with pure commercial lignin and three steam-exploded loblolly pine (SELP) samples (See **Table 4** for pretreatment conditions).

Table 22. Direct dyes studied in this work

Dye	CI*	Molecular weight (g/mol)
Direct Blue 1 (DB1)	24410	993
Direct Violet 9 (DV9)	27885	692

*CI: color index number

The binding isotherms for all four substrates are shown in **Figure 47**. Similar to DB1, the DV9 dye selectively binds onto cellulose but not lignin. The overall B_{max} for DV9 on lignin is almost 10-time lower than the SELPs. If we take into account the glucan content for all SELPs and assume no dye binds on lignin, the specific binding on cellulose ($B_{max,cellulose}$) can be calculated by dividing the overall B_{max} with the actual glucan concentration (**Table 23**). Less than 5 mg/g of DV9 adsorbed on lignin, while between 110 and 150 mg/g was observed on cellulose fraction in SELPs. Based on the binding results shown in **Table 23**, the cellulase prefers to bind on more severely pretreated SELP, which agrees with the hydrolysis results shown in Chapter 2.

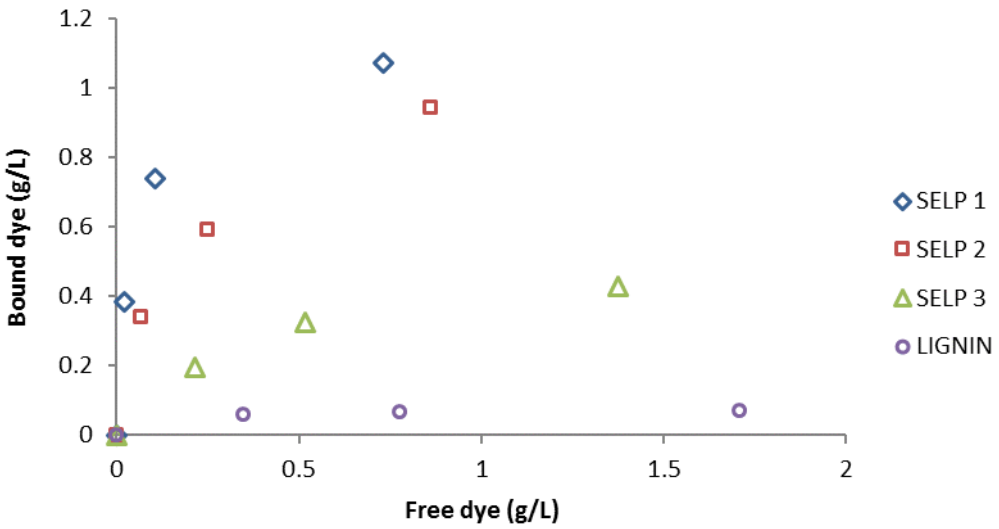


Figure 47. Binding specificity of DV9 on differently pretreated SELP and lignin.

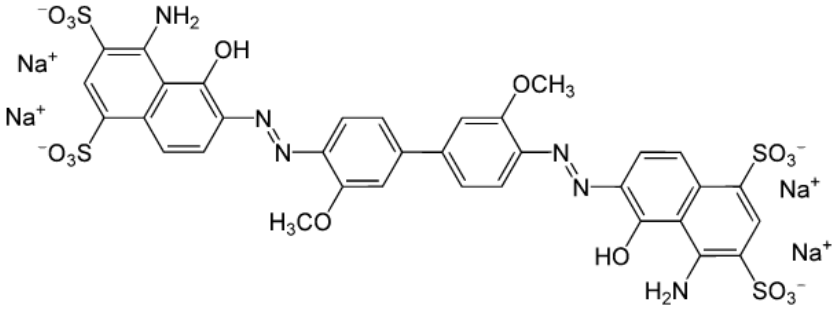
Table 23. Estimated binding parameters of DV9 on three SELPs

	SELP 1	**SELP 2**	**SELP 3**	**Lignin**
Glucan wt%	47.62	42.21	18.64	-
Overall B_{max} (g/L)	1.126	1.126	0.552	0.074
$B_{max,cellulose}$ (mg/g)	118.2	133.4	148.1	3.7 (Lignin)
K_{ad} (L/g)	21.03	5.26	2.63	11.32
R^2 (%)	99.6	98.9	99.9	100.0

The binding capacity of DV9 is significantly higher than that of DB1. The reasons for this discrepancy in binding could possibly be induced from the molecular structures (**Figure 48**). The more branches/substitutions on the backbone of the direct dyes tends to

decrease the binding affinity; this phenomena was also observed in a previous study which compared direct red 2 and direct blue 1 (Inglesby and Zeronian 2002). It has been demonstrated the dye molecules prefer to diffuse into the porous cellulose structure with its longitudinal axis parallel to the pore walls (Hori and Zollinger 1986, Inglesby and Zeronian 2002). The existence of fewer side chains presents less steric hindrance during diffusion thus results in higher binding affinity.

(A)



(B)

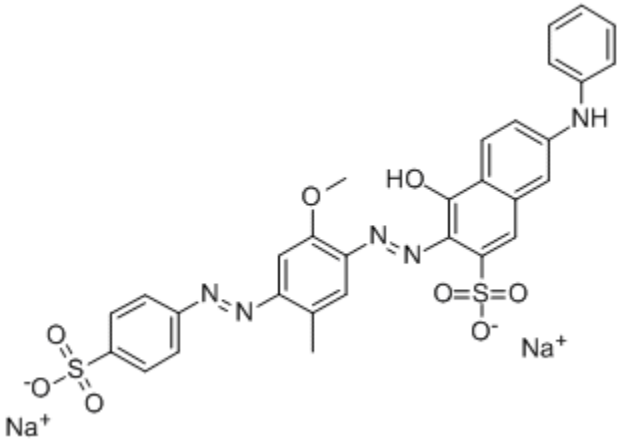


Figure 48. Molecular structure of the two direct dyes (A) DB1 and (B) DV9.

5.3.2 Characterization of pretreated Avicel with the modified Simons' staining technique

1-Methylimidazole (MI) was utilized as chemical pretreatment agent for Avicel with the optimal conditions (i.e. 25℃ for 5 min) identified in Chapter 4. The accessibility of the MI-pretreated and control Avicel were explored with the two molecule probes DV9 and cellulases. The pretreated cellulose was found to yield higher maximum binding capacity B_{max} based on the non-linear fitting with Langmuir model for both adsorption agents. The B_{max} increased 37% from 115.9 to 159 mg/g for DV9, while for the larger molecule cellulases, B_{max} increased more than 3-fold (**Table 24**).

Table 24. Summary of Avicel binding parameters for DV9 and cellulase.

Adsorption agent	Substrate	B_{max} (mg/g)	R^2 (%)	C/V ratio
DV9	Avicel	115.9	98.4	0.374
	MI_Avicel	159.2	99.5	0.827
Cellulase	Avicel	43.4	95.2	-
	MI_Avicel	131.7	98.8	-

Based on this observation, it can be inferred the MI pretreatment primarily creates larger pores which are accessible to the enzymes. Meanwhile, a fraction of tiny pores used to be non-accessible to both probes are also enlarged to be capable of accommodating DV9 but not cellulases. Since no change was observed in the crystalline index of cellulose after MI pretreatment, the crystalline to amorphous region ratio

remains constant. The observed improvement on cellulase accessibility is assumed to be the outcome of expansion and swelling of the amorphous region alone. Although the amorphous cellulose region is reported to be readily accessible to small molecules such as water (Nakamura, Hatakeyama et al. 1981, Vittadini, Dickinson et al. 2001), its accessibility to cellulase enzymes can be hindered by the bulky size and steric effect. Even if the MI molecule cannot penetrate into the highly-ordered crystalline cellulose to alter the crystal structure, we speculate it promotes cellulase accessibility by creating larger void volume in the amorphous region. **Figure 49** schematically illustrate the MI pretreatment impact on cellulose structure. This findings agree with earlier work which demonstrated the adsorption properties of cellulose fibers predominantly dependent on the void diameter and volume (Kreze, Jeler et al. 2002).

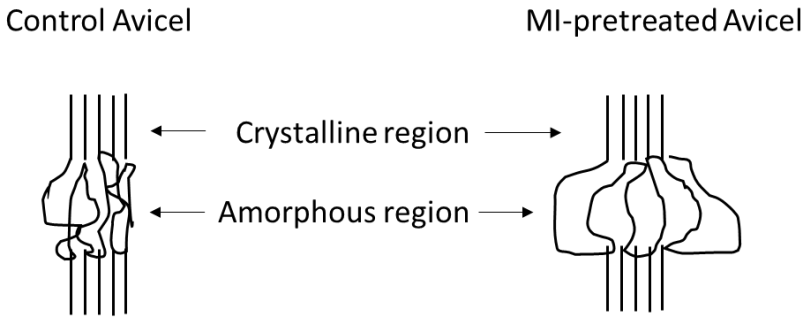


Figure 49. Schematic representation of MI pretreatment effect on Avicel.

For the conventional Simons' stain, the ratio of adsorbed DO15 and DB1, designated as O/B ratio, is frequently employed to evaluate the pretreatment effectiveness. O/B ratio is commonly used for the comparison of relative porosity of pretreated lignocellulosic biomass (Yu and Atalla 1998, Esteghlalian, Bilodeau et al.

2001). Various studies have provided evidence for the linear dependence of cellulose conversion on O/B ratio (Chandra, Ewanick et al. 2008). In this work, the ratio of adsorbed cellulase to DV9 is defined as C/V ratio which is equivalent to O/B ratio. The larger value of C/V ratio is indicative of more effective pretreatment technique which enhance the void volume for cellulase attack.

For differently pretreated Avicel, the C/V ratio is calculated in addition to the estimated binding parameters according to Equation (8) (**Table 24**). After MI pretreatment, the C/V ratio increased by 2.2-fold. However, the concept of C/V ratio can only be utilized for the assessment of pretreated pure cellulose substrate. When lignin is present in the system, a lignin blocking agent is required to prevent the non-productive cellulase binding that often leads to overestimation of accessibility.

$$C/V\ ratio = \frac{B_{max,cellulase}}{B_{max,DV9}} \quad (8)$$

5.3.3 The correlation between B_{max} and crystallinity

The two cellulose characteristics often identified as the rate-limiting parameters are accessibility and crystallinity (Jeoh, Ishizawa et al. 2007, Hall, Bansal et al. 2011). Correlation between the two factors has been studied with a variety of molecules including water, methanol, ethanol and formic acid (Ioelovich 2009). The water molecules were reported to have full access to the non-crystalline region of cellulose; as expected, accessibility reduces with increasing molecular size (Ioelovich 2009). However, the quantification of adsorption extent can be tedious with these solvent molecules. To explore the relationship between accessibility and crystallinity and provide

a convenient tool to estimate the substrate reactivity, the two direct dyes DB1 and DV9 were employed.

The small direct dyes can act as molecular probes for estimating the cellulose surface area of the amorphous region. **Figure 50** shows the partition of DV9 on four cellulose substrates from different origins or pretreatment conditions after 6 hour adsorption at 70 ℃. 77% and 81% PASC indicates the phosphoric acid concentrations utilized for sample preparation. The order of crystalline content in these samples are: FC > Avicel > 77% PASC > 81% PASC. A clear increasing trend in bound dye was observed with increasing amorphous cellulose fraction.

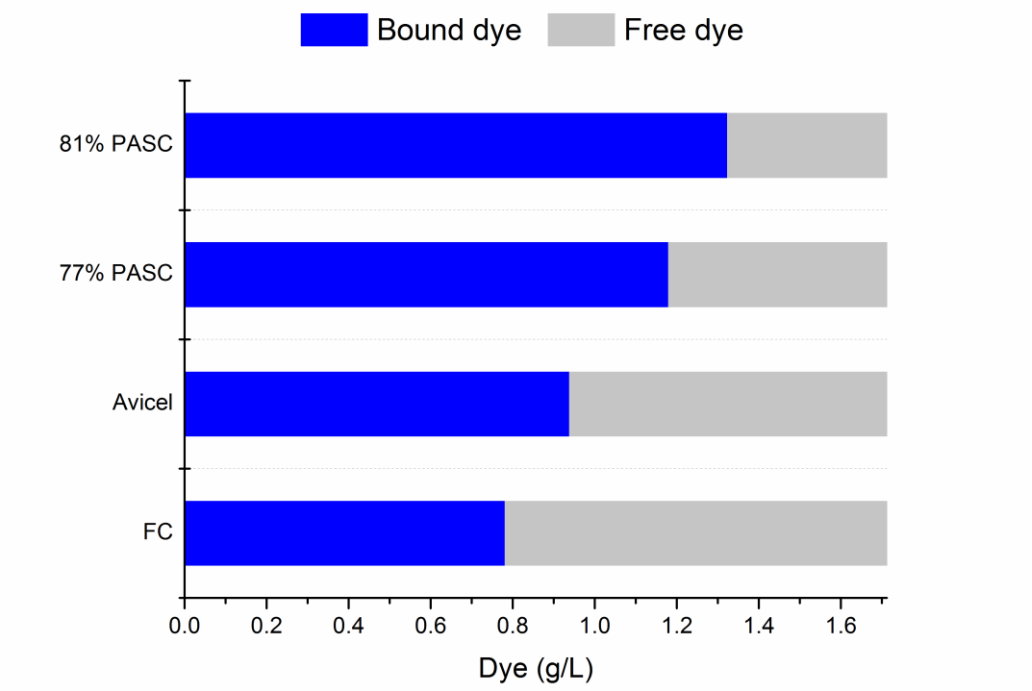


Figure 50. DV9 adsorption on various cellulose samples

To study the direct dye binding behavior on cellulose with various crystallinity, we generated intermediate crystallinity samples from Avicel and FC with different

phosphoric acid concentrations. The binding isotherms of DV9 and DB1 on the series of phosphoric acid swollen Avicel (PASC) and phosphoric acid swollen FC (PAFC) were obtained (**Figure 62**). The ten sets of adsorption data were fitted to Langmuir equation and the estimated B_{max} are listed in **Table 25**. In general, stronger binding affinity is achieved with higher acid concentration, which decrystallize cellulose crystal structure to a greater extent. B_{max} were plotted against calculated CrI with the developed spectrum fitting method introduced in Chapter 2 (**Figure 51**). For both DB1 and DV9, cellulose from different origins and pretreatment conditions shows strong inverse linear relationship between CrI and B_{max}. The B_{max} measured with small-sized molecular probes closely related to surface area can be conveniently used to predict the crystallinity of cellulose.

Table 25. Estimated B_{max} from Langmuir isotherm

Sample	**B_{max} (mg/g)**	**Sample**	**B_{max} (mg/g)**
Avicel	25.1	FC	23.2
76% PASC	52.6	78.09% PAFC	38.7
77% PASC	47.1	79.94% PAFC	60.3
80% PASC	57.8	81.06% PAFC	53.5
81% PASC	73.4	82.03% PAFC	79.0

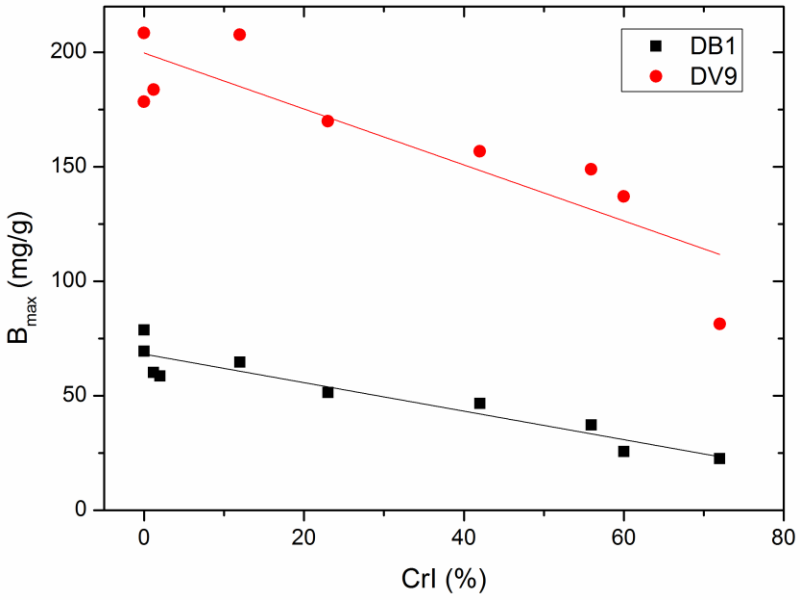


Figure 51. Correlation between B_{max} and CrI for cellulose with intermediate crystallinities derived from Avicel and FC.

To validate the relationship between B_{max} and CrI, partially converted Avicel was obtained. Enzymatic hydrolysis of Avicel was terminated at six desired time points to harvest residual solids with various conversion levels. Based on our earlier findings, the cellulose crystallinity index maintained at a constant (~ 61%) until approximately 90% conversion (Hall, Bansal et al. 2010). If the inverse linear dependency of B_{max} on CrI holds, the accessibility of DV9 on Avicel with different conversion will be a constant. The B_{max} was almost held at a constant for the initial 48% conversion; at 66% conversion, the reduction in B_{max} is less than 10% (**Figure 52**). These results agree with the cellulase accessibility data published for Avicel samples with the same conversion levels. However, more significant decrease (i.e. around 20%) in B_{max} at 66% was observed for cellulase (Bansal P., Vowell B. J. et al. 2012). Taking both cellulase and DV9 binding

results into account, negligible change in the overall porous structure occurred at initial hydrolysis stage; toward the end of hydrolysis, some larger pores are consumed by the enzymes while the inaccessible pores are largely retained.

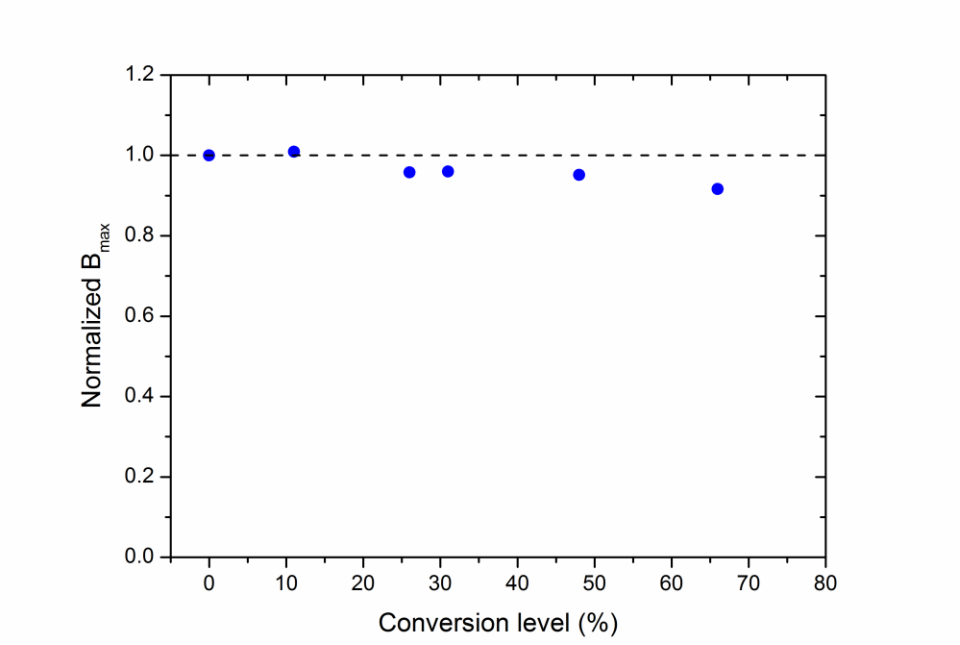


Figure 52. Normalized DV9 B_{max} values as a function of Avicel conversion level

5.4 Conclusions

In this chapter, the direct dye probes were found to be suitable to serve as convenient tool for the estimation of crystallinity for original and regenerated cellulose from different origins. When combining with larger probes such as cellulases, the adsorption results can be utilized to semi-quantitatively analyze the porous structure and pore size distribution to understand the effect of different pretreatment techniques and enzymatic hydrolysis.

With the small-sized direct dyes, the dependency of maximum binding capacity on crystallinity index was explored and a strong inverse linear relationship was revealed. Therefore, the direct dyes could also be used to estimate the relative crystallinity of cellulose samples.

CHAPTER 6

RECOMMENDATIONS AND CONCLUSIONS

The major limitation for biofuel production from lignocellulosic biomass through biological routes is the high recalcitrance to enzymatic saccharification. Low biodegradation efficiency necessitates the introduction of a biomass pretreatment process. However, the vast majority of the current pretreatment techniques are not economically attractive. Optimization of the existing routes and further exploration of new techniques are highly desired.

6.1 Recommendations

6.1.1 SO_2-catalyzed steam explosion

Steam explosion is one of the most successfully applied biomass pretreatment techniques. Compared to non-catalyzed steam explosion, acid catalyst SO_2 enables the recovery of hemicellulose in its monomeric sugar form to prevent further degradation (Clark and Mackie 1987). The pretreatment severity should be carefully controlled to avoid cellulose degradation, extensive lignin redistribution and re-condensation (Shevchenko, Beatson et al. 1999, Ibrahim, Dufresne et al. 2010). The trade-off between hemicellulose hydrolysis and cellulose retention is frequently observed. The ultimate goal is the improvement on overall sugar recovery and enzymatic digestibility of pretreated solids.

The cellulose crystalline fraction was identified to be one of the key rate-limiting parameters for steam-exploded lignocellulosic biomass. Other factors such as cellulose degree of polymerization, accessibility, lignin composition and molecular weight may also control the rate of hydrolysis. A more comprehensive study of a wide severity range with various substrate characteristics is required to gain better understanding of the rate-limiting parameters and steam explosion effectiveness to facilitate its optimization.

6.1.2 CBD-aided biological pretreatment of biomass

In Chapter 3, nine CBD mutations were identified based on PCA analysis, these mutations can be introduced to the WT CBD sequence to investigate their combined effect. The second round protein engineering introducing two of the top nine mutations of CBD in the GLC fusion protein were studied. The M4 (Y13W) and M7 (T23Y) mutations, which failed to be expressed in the *P. pastoris* strain as single mutations, were successfully introduced and expressed in the double variants. The preliminary binding results are shown in **Table 26**. Further studies combining the known beneficial variants could be conducted to look at the existence of synergic/additive effect on cellulose binding.

Table 26. Langmuir parameters estimated from the GLC variants binding isotherms.

	M9 (V18A)	**M4 + M9 (Y13W, V18A)**	**M7 + M9 (T23Y, V18A)**
B_{max} (mg/g)	15.54	22.87	15.34
K_{ad} (L/mg)	0.077	0.011	0.026

Besides the binding effect, the CBD was found to be involved in a variety of non-hydrolytic phenomena of cellulolytic enzymes (Bommarius, Sohn et al. 2014). A particularly attractive target for protein engineering is the improvement of thermostability. The CBD has been identified as a thermostabilizing domain for *T. reesei* Cel7A (Hall, Rubin et al. 2011), further improvement on the melting temperature of CBD could be desired for the development of a more thermostable Cel7A. Meanwhile, the effects of CBD binding affinity and thermostability could be studied simultaneously on the hydrolytic activity of the intact Cel7A enzyme with linker and catalytic domain.

The linker domain is often overlooked in the protein engineering of hydrolytic enzymes. The effect of linker length has been studied in our group, varying the Cel7A linker length by three amino acids (PPG, see **Figure 53**) causes both the cellulose binding affinity and reversibility of WT GLC to change dramatically (M. Sohn, unpublished results). The shorter linker with 21 amino acids exhibits almost 10-fold higher binding capacity compared to the longer linker with 24 amino acids, this phenomenon is possibly explained by the existence of a kink (formed by PPG) which changes the relative orientation of the N-terminal GFP domain to cellulose substrate. Moreover, the adsorption of the longer linker GLC is completely reversible (M. Sohn, unpublished results).

21 aa: -GNRGTTTTRRPATTTGSSPGP-
24 aa: -PPGGNRGTTTTRRPATTTGSSPGP-

Figure 53. The amino acid sequences of two Cel7A linkers in the GLCs studied (M. Sohn, unpublished results).

All GLC variants studied in Chapter 3 are comprised of 21 amino acids shown in **Figure 53**. The significant difference in binding behavior (i.e. decreased binding capacity and increased binding reversibility) induced by linker length can also be investigated in conjunction with CBD mutations to yield improved biological pretreatment agents. This could have an important implication that carefully designed linker might contribute to more dynamic binding and eventually interfere with the catalytic activity of enzymes.

For the identification of CBD mutation, the PCA analysis can be further improved by limiting the family 1 CBD sequence pool to only those with better binding affinity compared to Cel7A CBD. In this way, the resulted PCA suggested mutations are likely to produce improved binding as the mutated sequence aligns more closely with the underlying sequence pattern for stronger-binding family 1 CBDs. However, only limited comparative adsorption data was published in previous reports. A comparative study on a wide range of family 1 CBDs is desired to determine which sequences to be included in the sequence pool for PCA analysis.

6.1.3 Chemical pretreatment of lignocellulosic biomass

In Chapter 4, the discovery of a substituted imidazole was identified as an efficient lignin solvent and cellulose surface expander. The mechanism for this phenomenon is not clear. In depth understanding of the effect of organic solvent

pretreatment agents on biomass can be achieved by first gathering more biomass-dissolution information on different solvents. The effect of 1,2-dimethylimidazole (DMI) and 4(5)-hydroxymethylimidazole (HMI) (chemical structure shown in **Figure 54**) was also studied, both show similar pretreatment effect on microcrystalline cellulose. It seems the difference in ring substitution mainly affects the physical properties of the chemicals, such as melting and boiling temperature. The cellulose pretreatment effect remains for all three chemicals tested, indicating the potential role of the heterocycle in the pretreatment process. More analytical tools should be utilized to further investigate how the heterocycle expand the cellulose surface and void volume. Some factors that might contribute to the effect are reducing end concentration, H-bonding and cellulose derivatization. A detailed mechanistic study will reinforce our knowledge on this efficient pretreatment solvent system.

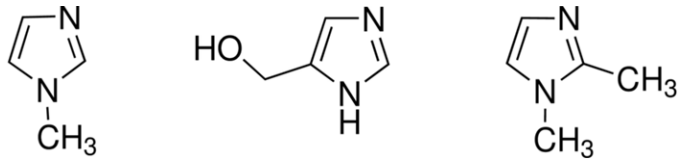


Figure 54. Chemical structures of the MI, HMI and DMI studied in this work.

Although MI-based pretreatment displays high pretreatment efficiency, the substituted imidazoles cannot completely solve the problems associated with biomass pretreatment. One biggest concern is the solvent cost. Large scale production cost of 1-methylimidazole (MI) was estimated to be around €8.00 per kg (Chen, Sharifzadeh et al.

2014), which is significantly higher than regular organic solvents. Although this cost is already reduced dramatically compared to commonly used imidazolium-based ionic liquids, further process development and optimization is strongly desired.

In addition to the cost issue, the current MI-based pretreatment at low severity is not effective enough for the raw biomass compared to other leading pretreatment techniques. MI is not capable of decrystallizing of cellulose. Combined with the limited accessibility of non-treated biomass to solvent, it is highly recommended the MI pretreatment to be utilized with another physical or thermal pretreatment method to synergistically deconstruct the biomass and markedly cut down the energy input for raw lignocellulosics. Meanwhile, hunting for an even more cost-efficient cellulose and/or biomass solvent can be an alternative pathway to achieve satisfactory pretreatment outcome with lower cost. This part of work will be continued in our group by Thomas Kwok.

6.1.4 Development of new biomass characterization assay

Chapter 5 discussed a new assay utilizing low molecular weight direct dyes with cellulase enzymes to semi-quantitatively determine the accessibility and pore size distribution. It could be used as an efficient method to quickly estimate and understand the effectiveness of pretreatment techniques. Conceptually, it is similar to Simons' staining. However, the availability of the polymeric direct orange 15 limits the applicability of this assay.

Currently, the assay has been demonstrated for pretreated pure cellulose. The presence of lignin in the system will complicate the assay and generate meaningless adsorption data, thus an effective lignin-blocking agent must be employed to block the

non-productive cellulase binding sites. A range of proteins/surfactants have been tested for this purpose, bovine serum albumin, Tween 20, Tween 80, polyethylene glycol are the most frequently utilized ones which show good blocking effect in earlier studies (Kumar and Wyman 2009, Qing, Yang et al. 2010). With the addition of appropriate lignin-blocking agent, the direct dye/cellulase method could be extended to lignin-containing substrates.

To estimate the cellulase-accessible region of the substrates, the affinity adsorption assay conducted with cellulase could be replaced with GLC. With the lignin binding sites blocked, the green fluorescent GLC offers a convenient and potentially high throughput assay to quickly assess the relative differences between differently pretreated substrates. Besides the employment of biological affinity probes, the rarely-available direct orange 15 dye could also be replaced by another direct dye after screening. Polymeric dyes with hydrodynamic size of around 5-7 nm and selective cellulose binding are possible alternatives.

6.2 Conclusions

6.2.1 Physicochemical pretreatment – SO_2-catalyzed steam explosion

Lignocellulosic biomass is the most abundant naturally renewable organic resource for biofuel production. Because of its high recalcitrance to enzymatic degradation, pretreatment is a crucial step before hydrolysis of the feedstock. A variety of pretreatment methods have been developed and intensively studied to achieve optimal yield without imposing significant adverse impact on the environment. The SO_2-catalyzed steam explosion is one of the most commonly utilized pretreatment methods.

The effect of steam explosion severity on the digestibility, accessibility, and crystallinity of loblolly pine were investigated in Chapter 2. A X-ray diffraction spectra-fitting method was developed and demonstrated to yield reliable crystallinity index (CrI) measurement for both raw and pretreated cellulose and lignocellulosics. Three steam-exploded loblolly pine (SELP) samples with different pretreatment severity were obtained, the CrIs are higher than raw biomass meaning the amorphous region is liable to degradation. High severity shows higher preference to hydrolyze crystalline region to yield pretreated solid with lower CrI. Strong inverse linear relationship between initial hydrolysis rates and CrI of pretreated loblolly pine was found. Similar to the case of pure cellulose (Hall, Bansal et al. 2010), the findings demonstrate the significance of CrI in enzymatic hydrolysis of pretreated lignocellulosic biomass.

6.2.2 Biological pretreatment – cellulose-binding domain as a biological pretreatment agent

Conventional biological pretreatment suffers from the extremely slow biomass degradation rate and low selectivity. A more realistic biological pretreatment method utilizing the cellulose-binding domain (CBD) was studied and demonstrated to render cellulose substrate better digestibility after 15 hour pretreatment period at 42 – 45 ℃. The current CBD-based biological pretreatment technique could be performed without chemical addition at mild conditions in a time-efficient manner and eliminate the need to wash and change buffer prior to the initiation of enzymatic hydrolysis. Although the sugar yield under current conditions was improved by less than 30%, development of a CBD with fully reversible cellulose binding could potentially realize further enhancement in the pretreatment outcome.

Protein engineering is a valuable tool to gain better understanding of the CBD structure-function relationship. It could not only generate CBDs with improved functionality for biological pretreatment, but also open up the chances to enhance the cellulolytic activity of intact Cel7A enzymes by combining with mutated CBDs.

6.2.3 Chemical pretreatment – a new technique with substituted imidazoles

A novel chemical pretreatment method using substituted heterocycles with low temperature and short residence time requirements was developed. 1-Methylimidazole (MI) is a precursor to some imidazolium-based ionic liquids. In Chapter 4, its potential utilization as a biomass pretreatment agent is being investigated for the first time. At mild conditions, such as 25℃ for 5 min at ambient pressure, a substantial increase in the hydrolysis rate throughout the entire course of conversion for cellulose substrate was obtained. Furthermore, the pretreatment effectiveness of MI on both untreated and steam-exploded lignocellulosic biomass including loblolly pine, switchgrass, and sugarcane bagasse has been studied, and MI was found to be an efficient delignifier. Remarkable rate enhancement was also observed for the non-woody lignocellulosic substrates after a short period of MI pretreatment at ambient conditions. The effect of MI on cellulose and lignocellulosics was studied through analysis of cellulose physical properties including crystallinity index, degree of polymerization and accessibility. The mechanism of MI pretreatment is yet to be determined.

6.2.4 Molecular probes – a convenient tool to assess pretreatment effectiveness

Two low molecular-weight direct dyes, Direct blue 1 and Direct violet 9, were used in combination with cellulases as convenient tools to semi-quantitatively evaluate

the effectiveness of pretreatment technique on pure cellulose. In addition, cellulose crystallinity was found to depend inverse linearly on the accessibility of both direct dyes, opening the opportunity to developing a fast and convenient new crystallinity measurement technique.

Appendix A

SUPPLEMENTARY MATERIALS FOR CHAPTER 3

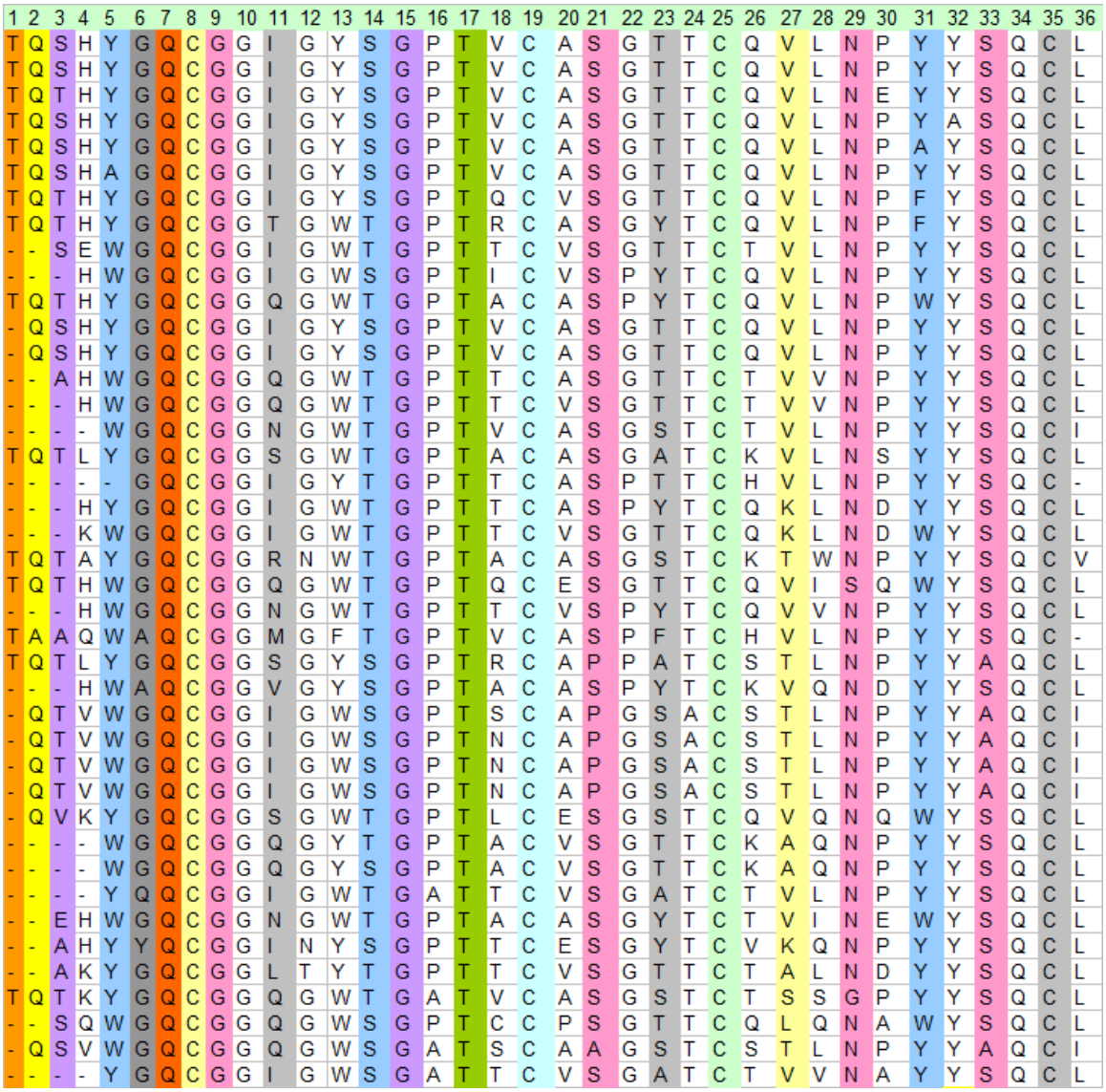


Figure 55. Sequence alignment of the selected 41 family 1 CBDs.

cellobiohydrolase I [Trichoderma reesei]
cellobiohydrolase I [Trichoderma viride]
cellobiohydrolase I [Trichoderma sp. XST1]
cellobiohydrolase I mutant [Trichoderma reesei]
cellobiohydrolase I mutant [Trichoderma reesei]
cellobiohydrolase I mutant [Trichoderma reesei]
cellobiohydrolase [Hypocrea virens]
cellobiohydrolase [Hypocrea lixii]
cellobiohydrolase [Penicillium Chrysoporium]
endo-1,4-xylanase D [Penicillium funiculosum]
endo-1,4-xylanase D Penicillium funiculosum]
endoglucanase [Aspergillus fumigatus Af293]
unnamed protein product [Sordaria macrospora]
xylanase/cellobiohydrolase [Penicillium funiculosum]
cellobiohydrolase I [Penicillium occitanis]
unnamed protein product [Podospora anserina]
glycosyl hydrolase family 45 protein Neosartorya fischeri]
cellobiohydrolase I [Volvariella volvacea]
cellobiohydrolyase [Aspergillus fumigatus]
endo-1,4-beta-xylanase [Aspergillus fumigatus]
Glycosyl hydrolase family 61 [Neosartorya fischeri]
acetyl xylan esterase [Hypocrea jecorina]
endo-1,4-beta-xylanase, putative [Talaromyces stipitatus]
cellulase [Irpex lacteus]
endoglucanase IV [Hypocrea jecorina]
endoglucanase 1 [Penicillium echinulatum]
endoglucanase II [Trichoderma viride]
endoglucanase III [Trichoderma viride]
endoglucanase III [Hypocrea jecorina]
endoglucanase II [Hypocrea jecorina]
glycoside hydrolase family 5 [Nectria haematococca]
endoglucanase I [Penicillium oxalicum]
endoglucanase I [Penicillium decumbens]
cellulose-binding beta-glucosidase [Phanerochaete chrysosporium]
acetyl xylan esterase [Aspergillus fumigatus]
endoglucanase/cellulase, putative [Aspergillus flavus]
family 61 endoglucanase [Phanerochaete chrysosporium]
Cel7A [Lentinula edodes]
exoglucanase [Verticillium albo-atrum]
cellobiohydrolase II [Acremonium cellulolyticus]
Cel6B [Lentinula edodes]

Figure 56. The corresponding origins of the 41 family 1 CBDs.

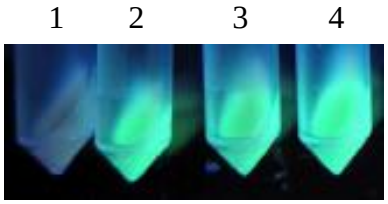


Figure 57. Cell pellets from SHuffle strain expression for 6 hours. 1: negative control with empty pET28a-sumo; 2: pET28a-sumo-GFPuv; 3: pET28a-sumo-GL; 4: pET28a-sumo-GLC.

ATGCATCATCATCATCATCAT*AGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCC*

CAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAG

AGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTG

GAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAAT

GCTTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGC

CCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGAACTACAAG

ACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAA

GGTATTGATTTTAAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTAT

AACTCACACAATGTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAA

CTTCAAAATTCGCCACAACATTGAAGATGGATCCGTTCAACTAGCAGACCATTATCA

ACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTC

GACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTC

*TTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAGCTCTACAAA*GGA

AATCGTGGCACCACCACCACCAGAAGACCAGCTACTACCACTGGAAGCTCTC

CCGGACCT**ACCCAATCTCATTACGGGCAATGCGGCGGTATTGGTTACAGC**

GGACCTACTGTCTGTGCTAGTGGCACAACTTGTCAAGTACTGAATCCTTA

TTACAGTCAGTGCCTGTAA

Figure 58. WT GLC gene sequence. 6X his tag is double underlined, followed by GFP sequence in Italic, underlined linker region and bold CBD sequence.

M H H H H H H *S K G E E L F T G V V P I L V E L D G D V N G H K F S V S G E G E G D A T Y G K L T L K F I C T T G K L P V P W P T L V T T F S Y G V Q C F S R Y P D H M K R H D F F K S A M P E G Y V Q E R T I S F K D D G N Y K T R A E V K F E G D T L V N R I E L K G I D F K E D G N I L G H K L E Y N Y N S H N V Y I T A D K Q K N G I K A N F K I R H N I E D G S V Q L A D H Y Q Q N T P I G D G P V L L P D N H Y L S T Q S A L S K D P N E K R D H M V L L E F V T A A G I T H G M D E L Y K* G N R G T T T T R R P A T T T G S S P G P **T Q S H Y G Q C G G I G Y S G P T V C A S G T T C Q V L N P Y Y S Q C L**

Figure 59. WT GLC amino acid sequence. 6X his tag is double underlined, followed by GFP sequence in Italic, underlined linker region and bold CBD sequence.

Table 27. Expression level of WT and variant GLCs in *E. coli*

GLC variant	Yield (mg/L)
WT	136
M1	150
M2	103
M3	147
M5	127
M6	112
M7	122
M8	157
M9	150

Table 28. Expression level of WT and variant GLCs in *P. pastoris*

GLC variant	Yield (mg/L)
M1	86.8
M2	102.3
M3	71.6
M5	73.6
M6	68.1
M8	70.4
M9	92.8
WT	48.2

Appendix B

SUPPLEMENTARY MATERIALS FOR CHAPTER 5

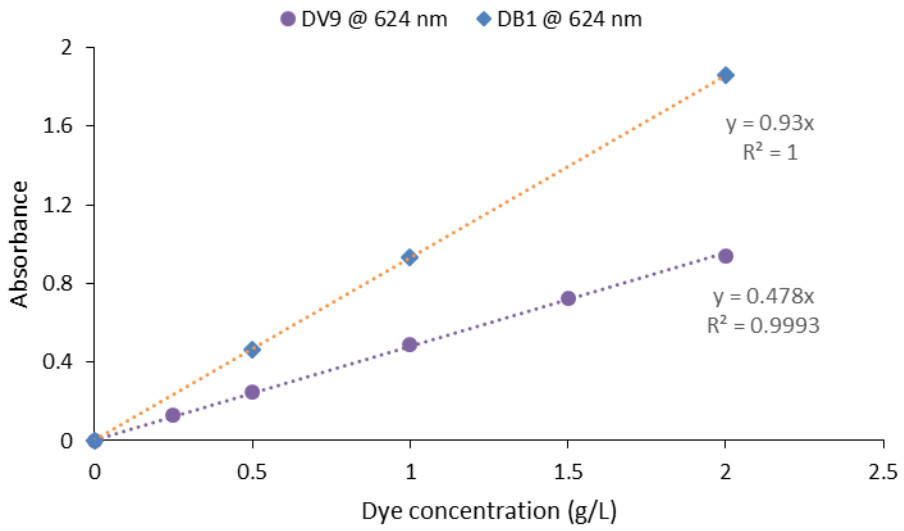


Figure 60. Calibration curves for DB1 and DV9 at 624 nm

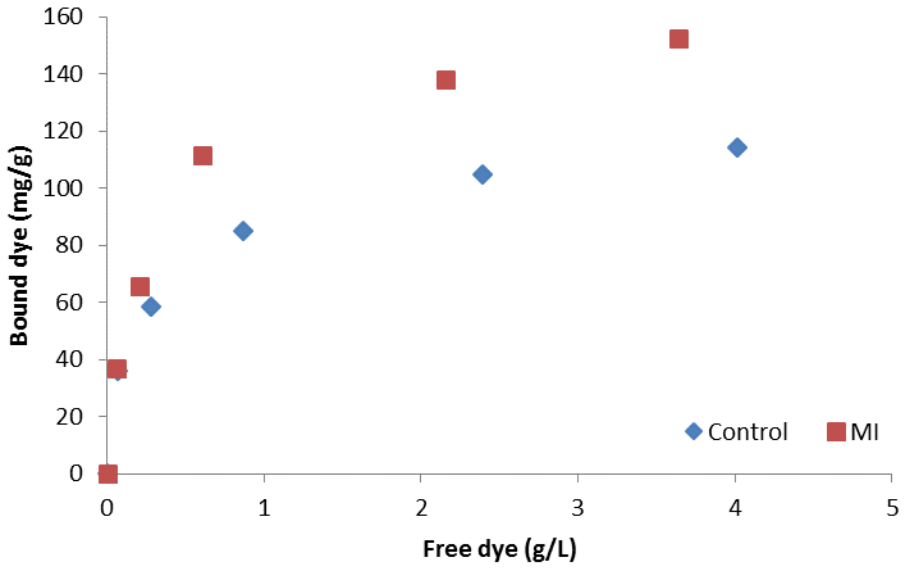


Figure 61. Binding isotherm of DV9 on MI-pretreated and control Avicel

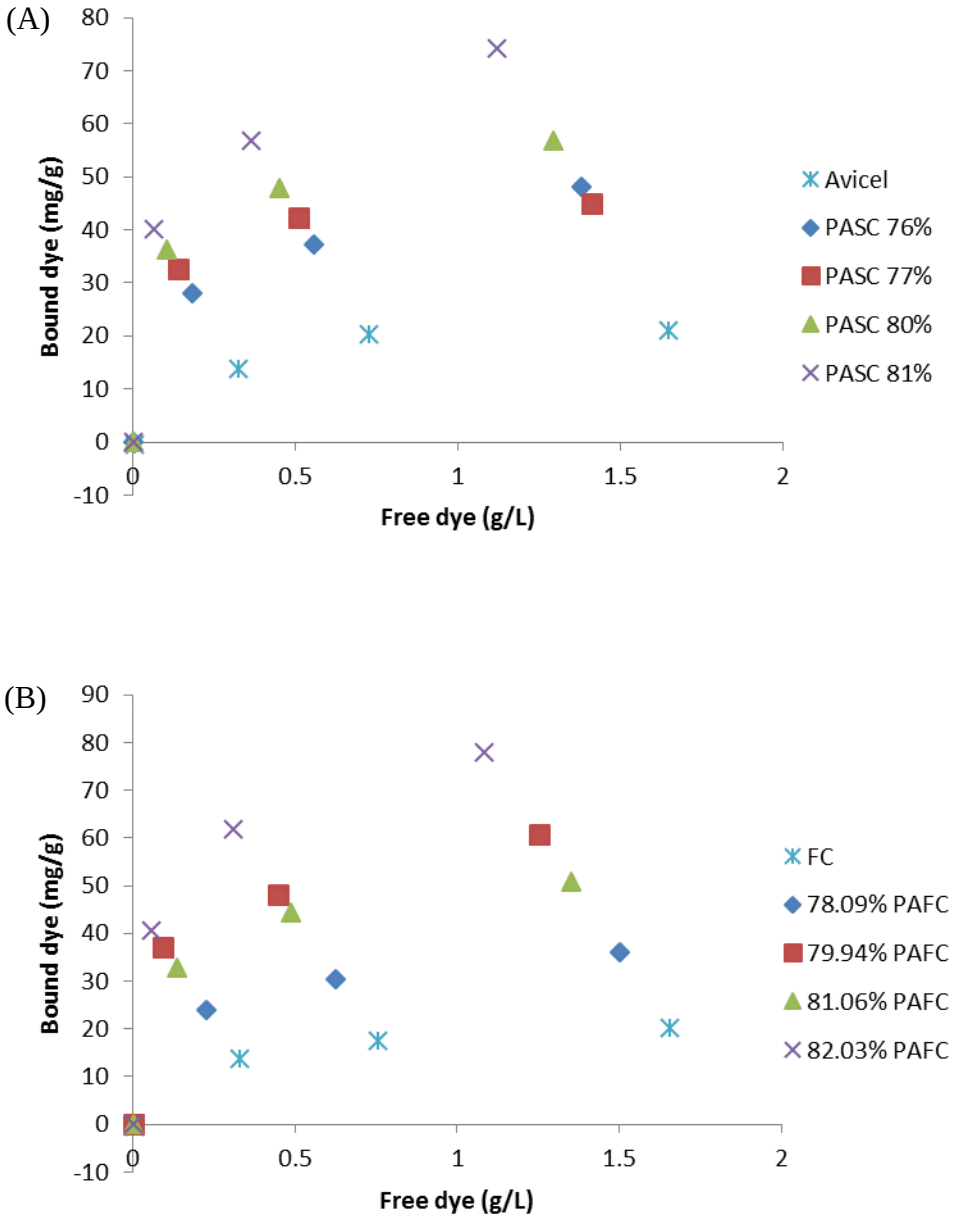


Figure 62. Adsorption isotherms for (A) PASCs and (B) PAFCs

Table 29. Phosphoric acid concentrations used for PAFC preparation

Sample	PAFC1	PAFC2	PAFC3	PAFC4	PAFC5
PA conc %	0	78.09	79.94	81.06	82.03

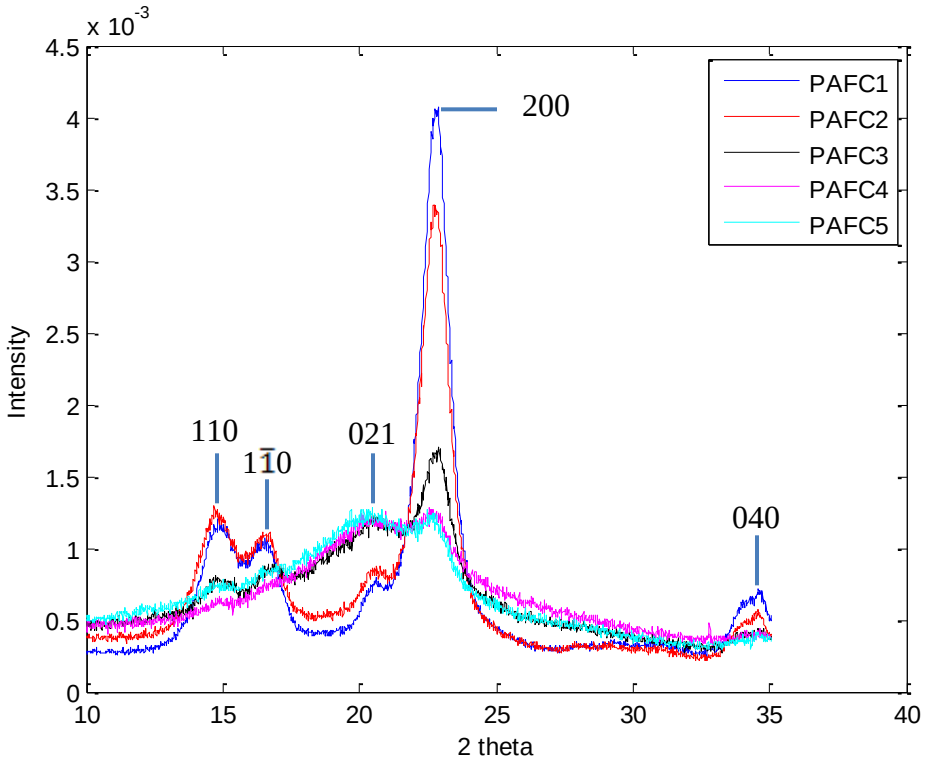


Figure 63. XRD spectra of the differently pretreated FC as indicated in **Table 29**.

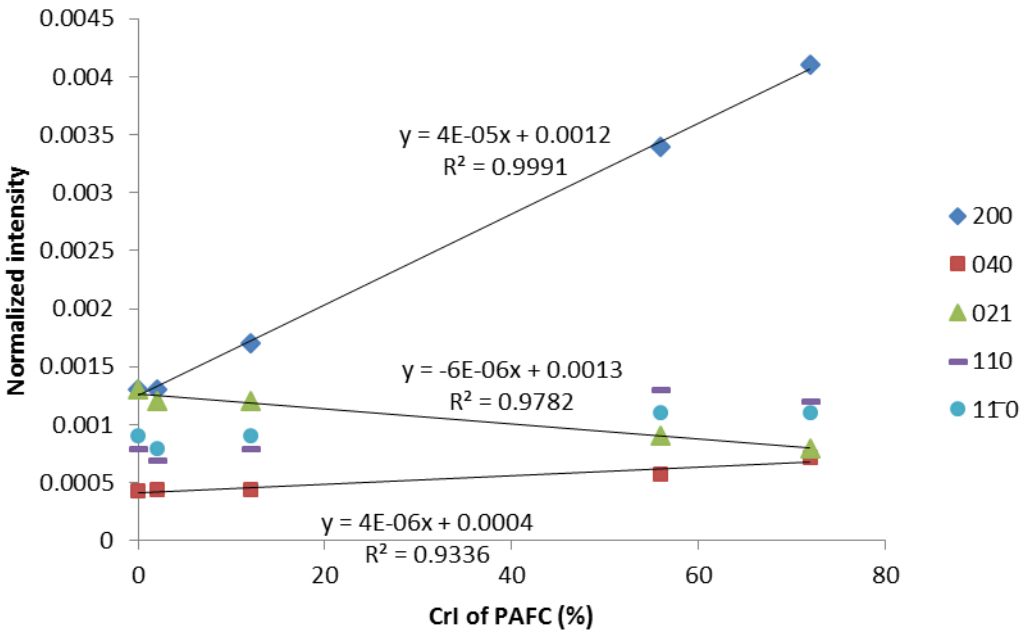


Figure 64. Correlation between normalized intensity and CrI of PAFCs at characteristic peaks corresponding to different crystal planes

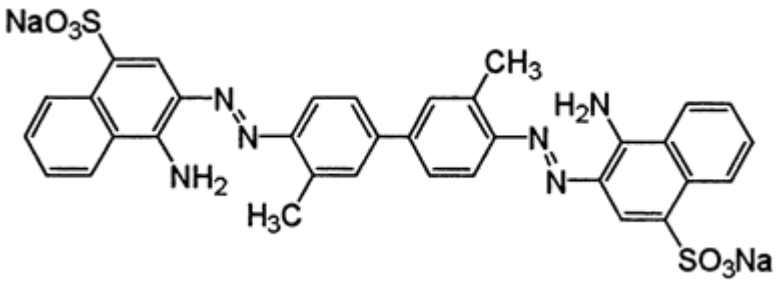


Figure 65. Chemical structure of Direct Red 2 (DR2)

REFERENCES

A. da Costa Lopes, K. J., R. Morais, E. Bogel-Lukasik and R. Bogel-Lukasik (2013). "Ionic liquids as a tool for lignocellulosic biomass fractionation." Sustain. Chem. Process. **1:3**: 1-31.

Adorjan, I., et al. (2004). "Kinetic and chemical studies on the isomerization of monosaccharides in< i> N</i>-methylmorpholine-< i> N</i>-oxide (NMMO) under Lyocell conditions." Carbohydr Res **339**(11): 1899-1906.

Alizadeh, H., et al. (2005). "Pretreatment of switchgrass by ammonia fiber explosion (AFEX)." Applied biochemistry and biotechnology **124**(1-3): 1133-1141.

Alzate, C. A. C. and O. J. S. Toro (2006). "Energy consumption analysis of integrated flowsheets for production of fuel ethanol from lignocellulosic biomass." Energy **31**(13): 2447-2459.

Andersson, S., et al. (2003). "Crystallinity of wood and the size of cellulose crystallites in Norway spruce (Picea abies)." Journal of Wood Science **49**(6): 531-537.

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

Ballesteros, I., et al. (2006). Ethanol production from steam-explosion pretreated wheat straw. Twenty-Seventh Symposium on Biotechnology for Fuels and Chemicals, Springer.

Bansal, P. (2011). "Computational and experimental investigation of the enzymatic hydrolysis of cellulose."

Bansal, P., et al. (2010). "Multivariate statistical analysis of X-ray data from cellulose: a new method to determine degree of crystallinity and predict hydrolysis rates." Bioresour Technol **101**(12): 4461-4471.

Bansal, P., et al. (2012). "Elucidation of cellulose accessibility, hydrolysability and reactivity as the major limitations in the enzymatic hydrolysis of cellulose." Bioresour Technol **107**: 243-250.

Beckham, G. T., et al. (2012). "Harnessing glycosylation to improve cellulase activity." Current opinion in biotechnology **23**(3): 338-345.

Beckham, G. T., et al. (2010). "Identification of amino acids responsible for processivity in a Family 1 carbohydrate-binding module from a fungal cellulase." The Journal of Physical Chemistry B **114**(3): 1447-1453.

Blanchette, R., et al. (1985). "Changes in structural and chemical components of wood delignified by fungi." Wood science and technology **19**(1): 35-46.

Blanchette, R. A. (1995). "Degradation of the lignocellulose complex in wood." Canadian Journal of Botany **73**(S1): 999-1010.

Boer, H., et al. (2000). "Characterization of Trichoderma reesei cellobiohydrolase Cel7A secreted from Pichia pastoris using two different promoters." Biotechnology and bioengineering **69**(5): 486-494.

Bommarius, A. S., et al. (2011). "Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst." Current opinion in chemical biology **15**(2): 194-200.

Bommarius, A. S., et al. (2008). "Cellulase kinetics as a function of cellulose pretreatment." Metabolic Engineering **10**(6): 370-381.

Bommarius, A. S., et al. (2014). "Protein engineering of cellulases." Current opinion in biotechnology **29**: 139-145.

Bommarius, B., et al. (2010). "Cost-effective expression and purification of antimicrobial and host defense peptides in< i> Escherichia coli</i>." Peptides **31**(11): 1957-1965.

Boraston, A. B., et al. (2004). "Carbohydrate-binding modules: fine-tuning polysaccharide recognition." Biochem J **382**(Pt 3): 769-781.

Boussaid, A., et al. (1999). "Fermentability of the hemicellulose-derived sugars from steam-exploded softwood (douglas fir)." Biotechnol Bioeng **64**(3): 284-289.

Brandt, A., et al. (2010). "The effect of the ionic liquid anion in the pretreatment of pine wood chips." Green Chemistry **12**(4): 672-679.

Brodeur, G., et al. (2011). "Chemical and physicochemical pretreatment of lignocellulosic biomass: a review." Enzyme research **2011**.

Brown, R. E., et al. (1989). "Protein measurement using bicinchoninic acid: elimination of interfering substances." Anal Biochem **180**(1): 136-139.

Brownell, H., et al. (1986). "Steam-explosion pretreatment of wood: Effect of chip size, acid, moisture content and pressure drop." Biotechnology and bioengineering **28**(6): 792-801.

Bura, R., et al. (2003). "Optimization of SO2-catalyzed steam pretreatment of corn fiber for ethanol production." Appl Biochem Biotechnol **105 -108**: 319-335.

Cai, C. M., et al. (2013). "THF co-solvent enhances hydrocarbon fuel precursor yields from lignocellulosic biomass." Green Chemistry **15**(11): 3140-3145.

Cantarella, M., et al. (2004). "Effect of Inhibitors Released during Steam-Explosion Treatment of Poplar Wood on Subsequent Enzymatic Hydrolysis and SSF." Biotechnology progress **20**(1): 200-206.

Cara, C., et al. (2006). "Enhanced enzymatic hydrolysis of olive tree wood by steam explosion and alkaline peroxide delignification." Process Biochemistry **41**(2): 423-429.

Cara, C., et al. (2008). "Production of fuel ethanol from steam-explosion pretreated olive tree pruning." Fuel **87**(6): 692-700.

Carrasco, C., et al. (2011). "Steam pretreatment and fermentation of the straw material "Paja Brava" using simultaneous saccharification and co-fermentation." J Biosci Bioeng **111**(2): 167-174.

Carta, G. and A. Jungbauer (2010). "Protein Chromatography: Process Development and Scale-up." Wiley-VCH, Weinheim.

Cereghino, J. L. and J. M. Cregg (2000). "Heterologous protein expression in the methylotrophic yeast Pichia pastoris." FEMS microbiology reviews **24**(1): 45-66.

Chandra, R., et al. (2008). "The characterization of pretreated lignocellulosic substrates prior to enzymatic hydrolysis, part 1: a modified Simons' staining technique." Biotechnology progress **24**(5): 1178-1185.

Chang, V. S., et al. (1997). Lime pretreatment of switchgrass. Biotechnology for Fuels and Chemicals, Springer**:** 3-19.

Chang, V. S., et al. (1998). "Lime pretreatment of crop residues bagasse and wheat straw." Applied biochemistry and biotechnology **74**(3): 135-159.

Chang, V. S., et al. (2001). "Oxidative lime pretreatment of high-lignin biomass." Applied biochemistry and biotechnology **94**(1): 1-28.

Chen, L., et al. (2014). "Specificity of O-glycosylation in enhancing the stability and cellulose binding affinity of Family 1 carbohydrate-binding modules." Proceedings of the National Academy of Sciences **111**(21): 7612-7617.

Chen, L., et al. (2014). "Inexpensive ionic liquids:[HSO 4]−-based solvent production at bulk scale." Green Chemistry **16**(6): 3098-3106.

Cherian, B. M., et al. (2010). "Isolation of nanocellulose from pineapple leaf fibres by steam explosion." Carbohydrate Polymers **81**(3): 720-725.

Chiriac, A. I., et al. (2010). "Engineering a family 9 processive endoglucanase from Paenibacillus barcinonensis displaying a novel architecture." Appl Microbiol Biotechnol **86**(4): 1125-1134.

Chundawat, S. P., et al. (2010). "Multifaceted characterization of cell wall decomposition products formed during ammonia fiber expansion (AFEX) and dilute acid based pretreatments." Bioresource technology **101**(21): 8429-8438.

Ciolacu, D., et al. (2010). "The effect of the cellulose-binding domain from Clostridium cellulovorans on the supramolecular structure of cellulose fibers." Carbohydr Res **345**(5): 621-630.

Clark, T. and K. Mackie (1987). "Steam explosion of the softwood Pinus radiata with sulphur dioxide addition. I. Process optimisation." Journal of Wood Chemistry and Technology **7**(3): 373-403.

Claus, I., et al. (2004). "Monoethanolamine (MEA) pulping of beech and spruce wood for production of dissolving pulp." Holzforschung **58**(6): 573-580.

Connors, K. A. and N. K. Pandit (1978). "N-Methylimidazole as a Catalyst for Analytical Acetylations of Hydroxy Compounds." Analytical Chemistry **50**(11): 1542-1545.

da Costa Sousa, L., et al. (2009). "'Cradle-to-grave'assessment of existing lignocellulose pretreatment technologies." Current opinion in biotechnology **20**(3): 339-347.

Dagel, D. J., et al. (2010). "In situ imaging of single carbohydrate-binding modules on cellulose microfibrils." The Journal of Physical Chemistry B **115**(4): 635-641.

Davin, L. B. and N. G. Lewis (2005). "Lignin primary structures and dirigent sites." Current opinion in biotechnology **16**(4): 407-415.

Dekker, R. and A. Wallis (1983). "Enzymic saccharification of sugarcane bagasse pretreated by autohydrolysis–steam explosion." Biotechnology and bioengineering **25**(12): 3027-3048.

Den Haan, R., et al. (2007). "Hydrolysis and fermentation of amorphous cellulose by recombinant Saccharomyces cerevisiae." Metab Eng **9**(1): 87-94.

Dijkerman, R., et al. (1996). "Adsorption characteristics of cellulolytic enzymes from the anaerobic fungus Piromyces sp. strain E2 on microcrystalline cellulose." Appl Environ Microbiol **62**(1): 20-25.

Doherty, W. O., et al. (2011). "Value-adding to cellulosic ethanol: Lignin polymers." Industrial Crops and products **33**(2): 259-276.

Donohoe, B. S., et al. (2008). "Visualizing Lignin Coalescence and Migration Through Maize Cell Walls Following Thermochemical Pretreatment." Biotechnology and Bioengineering **101**(5): 913-925.

Escovar-Kousen, J. M., et al. (2004). "Integration of computer modeling and initial studies of site-directed mutagenesis to improve cellulase activity on Cel9A from Thermobifida fusca." Applied biochemistry and biotechnology **113**(1-3): 287-297.

Esteghlalian, A. R., et al. (2001). "Do enzymatic hydrolyzability and Simons' stain reflect the changes in the accessibility of lignocellulosic substrates to cellulase enzymes?" Biotechnology progress **17**(6): 1049-1054.

Excoffier, G., et al. (1991). "Saccharification of steam-exploded poplar wood." Biotechnol Bioeng **38**(11): 1308-1317.

Fan, L. T., et al. (1980). "Mechanism of the Enzymatic-Hydrolysis of Cellulose - Effects of Major Structural Features of Cellulose on Enzymatic-Hydrolysis." Biotechnology and Bioengineering **22**(1): 177-199.

Foston, M., et al. (2011). "Chemical, ultrastructural and supramolecular analysis of tension wood in Populus tremula x alba as a model substrate for reduced recalcitrance. ." Energy Environ. Sci.(4): 4962.

Fox, S. C. and A. G. McDonald (2010). "Chemical and Thermal Characterization of Three Industrial Lignins and Their Corresponding Lignin Esters." BioResources **5**(2): 990-1009.

Galbe, M. and G. Zacchi (2012). "Pretreatment: The key to efficient utilization of lignocellulosic materials." Biomass and Bioenergy **46**: 70-78.

Gasparatos, A., et al. (2013). "Sustainability impacts of first-generation biofuels." Animal Frontiers **3**(2): 12-26.

George, A., et al. (2011). "The effect of ionic liquid cation and anion combinations on the macromolecular structure of lignins." Green Chemistry **13**(12): 3375-3385.

Gericke, M., et al. (2012). "Ionic liquids—promising but challenging solvents for homogeneous derivatization of cellulose." Molecules **17**(6): 7458-7502.

Gladden, J. M., et al. (2014). "Discovery and characterization of ionic liquid-tolerant thermophilic cellulases from a switchgrass-adapted microbial community." Biotechnology for Biofuels **7**(1): 15.

Gollapalli, L. E., et al. (2002). Predicting digestibility of ammonia fiber explosion (AFEX)-treated rice straw, Springer.

Goodell, B., et al. (2003). Brown-rot fungal degradation of wood: our evolving view. Current knowledge of wood deterioration mechanisms and its impact on biotechnology and wood preservation. Symposium at the 221st National Meeting of the American Chemical Society, San Diego, California, USA, 1-5 April 2001., American Chemical Society.

Gourlay, K., et al. (2012). "Use of substructure-specific carbohydrate binding modules to track changes in cellulose accessibility and surface morphology during the amorphogenesis step of enzymatic hydrolysis." Biotechnology for Biofuels **5**.

Gouverneur, V. and K. Müller (2012). Fluorine in Pharmaceutical and Medicinal Chemistry: From Biophysical Aspects to Clinical Applications, World Scientific.

Hall, M., et al. (2010). "Cellulose crystallinity--a key predictor of the enzymatic hydrolysis rate." FEBS J **277**(6): 1571-1582.

Hall, M., et al. (2011). "Biological pretreatment of cellulose: Enhancing enzymatic hydrolysis rate using cellulose-binding domains from cellulases." Bioresource technology **102**(3): 2910-2915.

Hall, M., et al. (2011). "Biological pretreatment of cellulose: enhancing enzymatic hydrolysis rate using cellulose-binding domains from cellulases." Bioresour Technol **102**(3): 2910-2915.

Hall, M., et al. (2011). "The cellulose-binding domain of cellobiohydrolase Cel7A from Trichoderma reesei is also a thermostabilizing domain." J Biotechnol **155**(4): 370-376.

Himmel, M. E., et al. (2007). "Biomass recalcitrance: engineering plants and enzymes for biofuels production." Science **315**(5813): 804-807.

Hoffrén, A.-M., et al. (1995). "Molecular dynamics simulation of fungal cellulose-binding domains: differences in molecular rigidity but a preserved cellulose binding surface." Protein engineering **8**(5): 443-450.

Holtzapple, M. T., et al. (1991). "The ammonia freeze explosion (AFEX) process." Applied biochemistry and biotechnology **28**(1): 59-74.

Hong, J., et al. (2007). "Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications." Langmuir **23**(25): 12535-12540.

Hori, T. and H. Zollinger (1986). "Role of water in the dyeing process." Textile chemist & colorist **18**(10): 19-25.

Huber, G. W., et al. (2006). "Synthesis of transportation fuels from biomass: chemistry, catalysts, and engineering." Chemical reviews **106**(9): 4044-4098.

Hui, J. P., et al. (2001). "Characterization of cellobiohydrolase I (Cel7A) glycoforms from extracts of< i> Trichoderma reesei</i> using capillary isoelectric focusing and electrospray mass spectrometry." Journal of Chromatography B: Biomedical Sciences and Applications **752**(2): 349-368.

Ibrahim, M., et al. (2010). "Isolation and characterization of cellulose and lignin from steam-exploded lignocellulosic biomass." Bioresources **5**(1): 397-418.

Ibrahim, M. M., et al. (2010). "Banana fibers and microfibrils as lignocellulosic reinforcements in polymer composites." Carbohydrate Polymers **81**(4): 811-819.

Igarashi, K., et al. (2009). "High speed atomic force microscopy visualizes processive movement of Trichoderma reesei cellobiohydrolase I on crystalline cellulose." J Biol Chem **284**(52): 36186-36190.

Inglesby, M. and S. Zeronian (1996). "The accessibility of cellulose as determined by dye adsorption." Cellulose **3**(1): 165-181.

Inglesby, M. and S. Zeronian (2002). "Direct dyes as molecular sensors to characterize cellulose substrates." Cellulose **9**(1): 19-29.

Inglesby, M., et al. (2002). "Aggregation of direct dyes investigated by molecular modeling." Textile research journal **72**(3): 231-239.

Inoue, H., et al. (2008). "Combining hot-compressed water and ball milling pretreatments to improve the efficiency of the enzymatic hydrolysis of eucalyptus." Biotechnology for Biofuels **1**(2): 1-9.

Ioelovich, M. (2009). "Accessibility and crystallinity of cellulose." BioResources **4**(3): 1168-1177.

Irwin, D., et al. (1998). "Roles of the catalytic domain and two cellulose binding domains of Thermomonospora fusca E4 in cellulose hydrolysis." J Bacteriol **180**(7): 1709-1714.

Janesko, B. G. (2011). "Modeling interactions between lignocellulose and ionic liquids using DFT-D." Physical Chemistry Chemical Physics **13**(23): 11393-11401.

Jeoh, T., et al. (2007). "Cellulase digestibility of pretreated biomass is limited by cellulose accessibility." Biotechnology and Bioengineering **98**(1): 112-122.

Jin, M., et al. (2013). "Phenotypic selection of a wild Saccharomyces cerevisiae strain for simultaneous saccharification and co-fermentation of AFEX™ pretreated corn stover." Biotechnol. Biofuels **6**(1): 108.

Kaar, W., et al. (1998). "Steam explosion of sugarcane bagasse as a pretreatment for conversion to ethanol." Biomass and Bioenergy **14**(3): 277-287.

Kaar, W. E. and M. T. Holtzapple (2000). "Using lime pretreatment to facilitate the enzymic hydrolysis of corn stover." Biomass and Bioenergy **18**(3): 189-199.

Kang, Y. Z., et al. (2013). "SO2-catalyzed steam explosion: The effects of different severity on digestibility, accessibility, and crystallinity of lignocellulosic biomass." Biotechnology progress **29**(4): 909-916.

Kilpelainen, I., et al. (2007). "Dissolution of wood in ionic liquids." Journal of Agricultural and Food Chemistry **55**(22): 9142-9148.

Kim, Y. S. and R. H. Newman (1995). "Solid state 13C NMR study of wood degraded by the brown rot fungus Gloeophyllum trabeum." Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood **49**(2): 109-114.

Kraulis, P. J., et al. (1989). "Determination of the 3-Dimensional Solution Structure of the C-Terminal Domain of Cellobiohydrolase-I from Trichoderma-Reesei - a Study Using Nuclear Magnetic-Resonance and Hybrid Distance Geometry Dynamical Simulated Annealing." Biochemistry **28**(18): 7241-7257.

Kreze, T., et al. (2002). "Correlation between structure characteristics and adsorption properties of regenerated cellulose fibers." Materials Research Innovations **5**(6): 277-283.

Kumar, D. and G. S. Murthy (2013). "Stochastic molecular model of enzymatic hydrolysis of cellulose for ethanol production." Biotechnology for Biofuels **6**.

Kumar, L., et al. (2011). "Influence of steam pretreatment severity on post-treatments used to enhance the enzymatic hydrolysis of pretreated softwoods at low enzyme loadings." Biotechnol Bioeng.

Kumar, P., et al. (2009). "Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production." Industrial & Engineering Chemistry Research **48**(8): 3713-3729.

Kumar, R. and C. E. Wyman (2009). "Effect of additives on the digestibility of corn stover solids following pretreatment by leading technologies." Biotechnology and bioengineering **102**(6): 1544-1557.

Kuo, C.-H. and C.-K. Lee (2009). "Enhanced enzymatic hydrolysis of sugarcane bagasse by< i> N</i>-methylmorpholine-< i> N</i>-oxide pretreatment." Bioresource technology **100**(2): 866-871.

Lau, M. W., et al. (2009). "The impacts of pretreatment on the fermentability of pretreated lignocellulosic biomass: a comparative evaluation between ammonia fiber expansion and dilute acid pretreatment." Biotechnology for Biofuels **2**(1): 30.

Le Costaouec, T., et al. (2013). "The role of carbohydrate binding module (CBM) at high substrate consistency: comparison of Trichoderma reesei and Thermoascus aurantiacus Cel7A (CBHI) and Cel5A (EGII)." Bioresour Technol **143**: 196-203.

Lee, J., et al. (2007). "Biological pretreatment of softwood Pinus densiflora by three white rot fungi." JOURNAL OF MICROBIOLOGY-SEOUL- **45**(6): 485.

Lee, S. H., et al. (2009). "Enzymatic saccharification of woody biomass micro/nanofibrillated by continuous extrusion process I - Effect of additives with cellulose affinity." Bioresource technology **100**(1): 275-279.

Li, C., et al. (2010). "Comparison of dilute acid and ionic liquid pretreatment of switchgrass: biomass recalcitrance, delignification and enzymatic saccharification." Bioresource technology **101**(13): 4900-4906.

Li, T., et al. (2012). "The impact of Trichoderma reesei Cel7A carbohydrate binding domain mutations on its binding to a cellulose surface: a molecular dynamics free energy study." Journal of molecular modeling **18**(4): 1355-1364.

Li, Y., et al. (2007). "Processivity, substrate binding, and mechanism of cellulose hydrolysis by Thermobifida fusca Cel9A." Appl Environ Microbiol **73**(10): 3165-3172.

Li, Y., et al. (2010). "Increased crystalline cellulose activity via combinations of amino acid changes in the family 9 catalytic domain and family 3c cellulose binding module of Thermobifida fusca Cel9A." Appl Environ Microbiol **76**(8): 2582-2588.

Liebert, T. (2010). Cellulose solvents-remarkable history, bright future. ACS symposium series, Oxford University Press.

Lin, L., et al. (2009). "Improved catalytic efficiency of Endo-β-1, 4-glucanase from Bacillus subtilis BME-15 by directed evolution." Applied microbiology and biotechnology **82**(4): 671-679.

Lin, Z., et al. (2010). "Ball milling pretreatment of corn stover for enhancing the efficiency of enzymatic hydrolysis." Applied biochemistry and biotechnology **162**(7): 1872-1880.

Linder, M., et al. (1995). "The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution." FEBS letters **372**(1): 96-98.

Linder, M., et al. (1995). "Identification of functionally important amino acids in the cellulose-binding domain of Trichoderma reesei cellobiohydrolase I." Protein Science **4**(6): 1056-1064.

Liu, L. Y. and H. Z. Chen (2006). "Enzymatic hydrolysis of cellulose materials treated with ionic liquid [BMIM]Cl." Chinese Science Bulletin **51**(20): 2432-2436.

Liu, Z.-H., et al. (2013). "Effects of biomass particle size on steam explosion pretreatment performance for improving the enzyme digestibility of corn stover." Industrial Crops and products **44**: 176-184.

Lobstein, J., et al. (2012). "SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm." Microb Cell Fact **11**(1): 56-56.

Lu, B., et al. (2014). "Cation does matter: how cationic structure affects the dissolution of cellulose in ionic liquids." Green Chemistry.

Lu, F. C. and J. Ralph (2003). "Non-degradative dissolution and acetylation of ball-milled plant cell walls: high-resolution solution-state NMR." Plant Journal **35**(4): 535-544.

Lu, Y., et al. (2002). "Cellulase adsorption and an evaluation of enzyme recycle during hydrolysis of steam-exploded softwood residues." Appl Biochem Biotechnol **98-100**: 641-654.

Luterbacher, J. S., et al. (2014). "Nonenzymatic sugar production from biomass using biomass-derived γ-valerolactone." Science **343**(6168): 277-280.

Mackie, K., et al. (1985). "Effect of sulphur dioxide and sulphuric acid on steam explosion of aspenwood." Journal of Wood Chemistry and Technology **5**(3): 405-425.

Mansfield, S. D., et al. (1999). "Substrate and enzyme characteristics that limit cellulose hydrolysis." Biotechnology progress **15**(5): 804-816.

Morton, T. (1935). "The dyeing of cellulose with direct dyestuffs; the importance of the colloidal constitution of the dye solution and of the fine structure of the fibre." Transactions of the Faraday Society **31**: 262-276.

Mosier, N., et al. (2005). "Features of promising technologies for pretreatment of lignocellulosic biomass." Bioresource technology **96**(6): 673-686.

Moxley, G., et al. (2008). "Efficient sugar release by the cellulose solvent-based lignocellulose fractionation technology and enzymatic cellulose hydrolysis." J Agric Food Chem **56**(17): 7885-7890.

Murnen, H. K., et al. (2007). "Optimization of ammonia fiber expansion (AFEX) pretreatment and enzymatic hydrolysis of Miscanthus x giganteus to fermentable sugars." Biotechnology progress **23**(4): 846-850.

Naik, S., et al. (2010). "Production of first and second generation biofuels: a comprehensive review." Renewable and Sustainable Energy Reviews **14**(2): 578-597.

Nakagame, S., et al. (2011). "The isolation, characterization and effect of lignin isolated from steam pretreated Douglas-fir on the enzymatic hydrolysis of cellulose." Bioresour Technol **102**(6): 4507-4517.

Nakamura, K., et al. (1981). "Studies on bound water of cellulose by differential scanning calorimetry." Textile research journal **51**(9): 607-613.

Nakazawa, H., et al. (2009). "Directed evolution of endoglucanase III (Cel12A) from Trichoderma reesei." Applied microbiology and biotechnology **83**(4): 649-657.

Nguyen, M. T., et al. (2009). "Hydrothermal acid pretreatment of Chlamydomonas reinhardtii biomass for ethanol production." Journal of microbiology and biotechnology **19**(2): 161-166.

Nguyen, Q. A., et al. (2000). "Two-stage dilute-acid pretreatment of softwoods." Appl Biochem Biotechnol **84-86**: 561-576.

Ochoa-Villarreal, M., et al. (2012). "Plant cell wall polymers: function, structure and biological activity of their derivatives."

Overend, R. P. and E. Chornet (1987). "Fractionation of Lignocellulosics by Steam-Aqueous Pretreatments." Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences **321**(1561): 523-536.

Pan, X., et al. (2005). "Strategies to enhance the enzymatic hydrolysis of pretreated softwood with high residual lignin content." Appl Biochem Biotechnol **121-124**: 1069-1079.

Pandey, K. and A. Pitman (2003). "FTIR studies of the changes in wood chemistry following decay by brown-rot and white-rot fungi." International Biodeterioration & Biodegradation **52**(3): 151-160.

Park, S., et al. (2010). "Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance." Biotechnol Biofuels **3**: 10.

Payne, C. M., et al. (2013). "Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose." Proceedings of the National Academy of Sciences **110**(36): 14646-14651.

Percival Zhang, Y.-H., et al. (2006). "Outlook for cellulase improvement: screening and selection strategies." Biotechnology advances **24**(5): 452-481.

Petrus, L. and C. Petrus-Hoogenbosch (2006). Process for organosolv pulping and use of a gamma lactone in a solvent for organosolv pulping, Google Patents.

Pinkert, A., et al. (2011). "Extracting wood lignin without dissolving or degrading cellulose: investigations on the use of food additive-derived ionic liquids." Green Chemistry **13**(11): 3124-3136.

Puri, V. P. (1984). "Effect of crystallinity and degree of polymerization of cellulose on enzymatic saccharification." Biotechnology and bioengineering **26**(10): 1219-1222.

Qing, Q., et al. (2010). "Impact of surfactants on pretreatment of corn stover." Bioresource technology **101**(15): 5941-5951.

Qiu, W. and H. Chen (2012). "Enhanced the enzymatic hydrolysis efficiency of wheat straw after combined steam explosion and laccase pretreatment." Bioresource technology **118**: 8-12.

Rahikainen, J., et al. (2011). "Inhibition of enzymatic hydrolysis by residual lignins from softwood—study of enzyme binding and inactivation on lignin-rich surface." Biotechnology and bioengineering **108**(12): 2823-2834.

Ramakrishnan, S., et al. (2010). "Enzymatic hydrolysis of cellulose dissolved in N-methyl morpholine oxide/water solutions." Bioresource technology **101**(13): 4965-4970.

Ramos, L. (2003). "The chemistry involved in the steam treatment of lignocellulosic materials." Quim. Nova **26**(6): 863-871.

Ravalason, H., et al. (2009). "Fusion of a family 1 carbohydrate binding module of Aspergillus niger to the Pycnoporus cinnabarinus laccase for efficient softwood kraft pulp biobleaching." J Biotechnol **142**(3-4): 220-226.

Reinikainen, T., et al. (1992). "Investigation of the function of mutated cellulose-binding domains of Trichoderma reesei cellobiohydrolase I." Proteins: Structure, Function, and Bioinformatics **14**(4): 475-482.

Rinaldi, R. and F. Schuth (2009). "Acid Hydrolysis of Cellulose as the Entry Point into Biorefinery Schemes." Chemsuschem **2**(12): 1096-1107.

Rollin, J. A., et al. (2011). "Increasing cellulose accessibility is more important than removing lignin: A comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia." Biotechnology and bioengineering **108**(1): 22-30.

Rosenau, T., et al. (2001). "The chemistry of side reactions and byproduct formation in the system NMMO/cellulose (Lyocell process)." Progress in Polymer Science **26**(9): 1763-1837.

Sannigrahi, P., et al. (2010). "Effects of organosolv pretreatment and enzymatic hydrolysis on cellulose structure and crystallinity in Loblolly pine." Carbohydr Res **345**(7): 965-970.

Sannigrahi, P., et al. (2008). "Effects of two-stage dilute acid pretreatment on the structure and composition of lignin and cellulose in loblolly pine." BioEnergy Research **1**(3-4): 205-214.

Sathitsuksanoh, N., et al. (2013). "New lignocellulose pretreatments using cellulose solvents: a review." Journal of Chemical Technology and Biotechnology **88**(2): 169-180.

Sawada, T., et al. (1995). "Effects of fungal pretreatment and steam explosion pretreatment on enzymatic saccharification of plant biomass." Biotechnology and bioengineering **48**(6): 719-724.

Segal, L., et al. (1962). "An empirical method for estimating the degree of crystallinity of native cellulose using the x-ray diffractometer." Textile Research Journal **29**: 786-794.

Shafiei, M., et al. (2010). "Pretreatment of spruce and oak by< i> N</i>-methylmorpholine-< i> N</i>-oxide (NMMO) for efficient conversion of their cellulose to ethanol." Bioresource technology **101**(13): 4914-4918.

Shevchenko, S., et al. (2000). "Optimization of monosaccharide recovery by post-hydrolysis of the water-soluble hemicellulose component after steam explosion of softwood chips." Bioresource technology **72**(3): 207-211.

Shevchenko, S. M., et al. (1999). "The nature of lignin from steam explosion/enzymatic hydrolysis of softwood." Applied biochemistry and biotechnology **79**(1-3): 867-876.

Shoemaker, S., et al. (1983). "Molecular-Cloning of Exo-Cellobiohydrolase-I Derived from Trichoderma-Reesei Strain-L27." Bio-Technology **1**(8): 691-696.

Shoemaker, S., et al. (1983). "Characterization and Properties of Cellulases Purified from Trichoderma-Reesei Strain-L27." Bio-Technology **1**(8): 687-690.

Sierra, R., et al. (2008). "Producing fuels and chemicals from lignocellulosic biomass." SBE special section: biofuels.

Simmons, C. W., et al. (2014). "Bacillus coagulans tolerance to 1-ethyl-3-methylimidazolium-based ionic liquids in aqueous and solid-state thermophilic culture." Biotechnology progress **30**(2): 311-316.

Simons, F. L. (1950). "A stain for use in the microscopy of beaten fibers." Tappi **33**(7): 312-314.

Sims, R. E., et al. (2010). "An overview of second generation biofuel technologies." Bioresource technology **101**(6): 1570-1580.

Singh, S., et al. (2009). "Visualization of Biomass Solubilization and Cellulose Regeneration During Ionic Liquid Pretreatment of Switchgrass." Biotechnology and bioengineering **104**(1): 68-75.

Soderstrom, J., et al. (2004). "Effect of washing on yield in one- and two-step steam pretreatment of softwood for production of ethanol." Biotechnol Prog **20**(3): 744-749.

Söderström, J., et al. (2002). "Two-step steam pretreatment of softwood with SO2 impregnation for ethanol production." Applied biochemistry and biotechnology **98**(1-9): 5-21.

Söderström, J., et al. (2003). "Two-step steam pretreatment of softwood by dilute H< sub> 2</sub> SO< sub> 4</sub> impregnation for ethanol production." Biomass and Bioenergy **24**(6): 475-486.

Sood, Y., et al. (2010). "Surface charge of different paper making raw materials and its influence on paper properties." J. Sci. Ind. Res **69**(4): 300-304.

Srisodsuk, M., et al. (1993). "Role of the interdomain linker peptide of Trichoderma reesei cellobiohydrolase I in its interaction with crystalline cellulose." Journal of Biological Chemistry **268**(28): 20756-20761.

Sun, L., et al. (2013). "Unveiling high-resolution, tissue specific dynamic changes in corn stover during ionic liquid pretreatment." Rsc Advances **3**(6): 2017-2027.

Swatloski, R. P., et al. (2002). "Dissolution of cellose with ionic liquids." Journal of the American Chemical Society **124**(18): 4974-4975.

Tadesse, H. and R. Luque (2011). "Advances on biomass pretreatment using ionic liquids: An overview." Energy & Environmental Science **4**(10): 3913-3929.

Taherzadeh, M. J. and K. Karimi (2008). "Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review." Int J Mol Sci **9**(9): 1621-1651.

Tang, C. D., et al. (2013). "Fusing a Carbohydrate-Binding Module into the Aspergillus usamii beta-Mannanase to Improve Its Thermostability and Cellulose-Binding Capacity by In Silico Design." PloS one **8**(5).

Taylor, C. B., et al. (2012). "Computational investigation of glycosylation effects on a family 1 carbohydrate-binding module." Journal of Biological Chemistry **287**(5): 3147-3155.

Teramoto, Y., et al. (2008). "Pretreatment of eucalyptus wood chips for enzymatic saccharification using combined sulfuric acid-free ethanol cooking and ball milling." Biotechnology and bioengineering **99**(1): 75-85.

Thakur, V. K., et al. (2014). "Progress in green polymer composites from lignin for multifunctional applications: A review." ACS Sustainable Chemistry & Engineering **2**(5): 1072-1092.

Thongekkaew, J., et al. (2013). "Fusion of cellulose binding domain from Trichoderma reesei CBHI to Cryptococcus sp S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates." Enzyme and Microbial Technology **52**(4-5): 241-246.

Tsao, G. (1999). Recent progress in bioconversion of lignocelluosics, Springer.

Tu, M., et al. (2009). "Adsorption of cellulase on cellulolytic enzyme lignin from lodgepole pine." J Agric Food Chem **57**(17): 7771-7778.

Varnai, A., et al. (2013). "Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs." Biotechnol Biofuels **6**(1): 30.

Verd á, P., et al. (2014). "Fractionation of lignocellulosic biomass with the ionic liquid 1-butylimidazolium hydrogen sulfate." Green Chemistry **16**(3): 1617-1627.

Vittadini, E., et al. (2001). "< sup> 1</sup> H and< sup> 2</sup> H NMR mobility in cellulose." Carbohydrate Polymers **46**(1): 49-57.

Voutilainen, S. P., et al. (2013). "Engineering chimeric thermostable GH7 cellobiohydrolases in Saccharomyces cerevisiae." Appl Microbiol Biotechnol.

Wallace, G., et al. (1991). "Lignin-carbohydrate complexes in graminaceous cell walls in relation to digestibility." Animal Feed Science and Technology **32**(1): 193-199.

Wang, K., et al. (2009). "Influence of steaming explosion time on the physic-chemical properties of cellulose from Lespedeza stalks (Lespedeza crytobotrya)." Bioresour Technol **100**(21): 5288-5294.

Wilson, D. B. (2009). "Cellulases and biofuels." Current opinion in biotechnology **20**(3): 295-299.

Worrall, J. J., et al. (1997). "Comparison of wood decay among diverse lignicolous fungi." Mycologia: 199-219.

Wyman, C. E., et al. (2005). "Comparative sugar recovery data from laboratory scale application of leading pretreatment technologies to corn stover." Bioresource technology **96**(18): 2026-2032.

Wyman, C. E., et al. (2005). "Coordinated development of leading biomass pretreatment technologies." Bioresource technology **96**(18): 1959-1966.

Xu, J., et al. (2014). "An ionic liquid tolerant cellulase derived from chemically polluted microhabitats and its application in in situ saccharification of rice straw." Bioresource technology **157**: 166-173.

Yamashiki, T., et al. (1990). "Characterisation of cellulose treated by the steam explosion method. Part 1: Influence of cellulose resources on changes in morphology, degree of polymerisation, solubility and solid structure." British Polymer Journal **22**(1): 73-83.

Yang, B., Dai Z., Ding. S., Wyman CE (2011). "Enzymatic hydrolysis of cellulosic biomass." Biofuels **2**(4): 421-450.

Yang, B. and C. E. Wyman (2008). "Pretreatment: the key to unlocking low-cost cellulosic ethanol." Biofuels, Bioproducts and Biorefining **2**(1): 26-40.

Yarbrough, J. M., et al. (2009). "Plant cell wall characterization using scanning probe microscopy techniques." Biotechnol Biofuels **2**: 17.

Yu, X. and R. H. Atalla (1998). "A staining technique for evaluating the pore structure variations of microcrystalline cellulose powders." Powder Technology **98**(2): 135-138.

Yu, X., et al. (1995). "Mechanism of action of Simons' stain." Tappi Journal **78**: 175-175.

Zhang, L.-h., et al. (2008). "Effect of steam explosion on biodegradation of lignin in wheat straw." Bioresource technology **99**(17): 8512-8515.

Zhang, S. J., et al. (2006). "Physical properties of ionic liquids: Database and evaluation." Journal of Physical and Chemical Reference Data **35**(4): 1475-1517.

Zhang, X. Z. and Y. H. P. Zhang (2011). "Simple, fast and high-efficiency transformation system for directed evolution of cellulase in Bacillus subtilis." Microbial biotechnology **4**(1): 98-105.

Zhang, Y.-H. P., et al. (2010). Advances in cellulose solvent-and organic solvent-based lignocellulose fractionation (COSLIF), ACS Publications**:** 365-379.

Zhang, Y., et al. (2012). "Pretreatment based on two-step steam explosion combined with an intermediate separation of fiber cells-optimization of fermentation of corn straw hydrolysates." Bioresource technology **121**: 100-104.

Zhang, Y. H. P., et al. (2007). "Fractionating recalcitrant lignocellulose at modest reaction conditions." Biotechnology and Bioengineering **97**(2): 214-223.

Zhang, Y. H. P. and L. R. Lynd (2004). "Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems." Biotechnology and bioengineering **88**(7): 797-824.

Zhang, Y. H. P. and L. R. Lynd (2005). "Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis." Biomacromolecules **6**(3): 1510-1515.

Zhao, Y., et al. (2008). "Enhanced enzymatic hydrolysis of spruce by alkaline pretreatment at low temperature." Biotechnol Bioeng **99**(6): 1320-1328.

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

Zheng, Y., et al. (2009). "Overview of biomass pretreatment for cellulosic ethanol production." International Journal of Agricultural and Biological Engineering **2**(3): 51-68.

Zheng, Y. F., et al. (2013). "Temperature sensitivity of cellulase adsorption on lignin and its impact on enzymatic hydrolysis of lignocellulosic biomass." Journal of biotechnology **166**(3): 135-143.

Zhu, J. Y., et al. (2010). "Pretreatment of woody biomass for biofuel production: energy efficiency, technologies, and recalcitrance." Appl Microbiol Biotechnol **87**(3): 847-857.

Zhu, J. Y. and X. J. Pan (2010). "Woody biomass pretreatment for cellulosic ethanol production: Technology and energy consumption evaluation." Bioresour Technol **101**(13): 4992-5002.

Zhu, Z., et al. (2009). "Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility." Biotechnology and bioengineering **103**(4): 715-724.

VITA

Yuzhi Kang

Yuzhi Kang was born in Jiangxi, China in 1987. She attended South China University of Technology (SCUT), Guangzhou, China in 2004 and obtained her B.S. in Pulp and Paper Engineering in 2008. After graduation, she received first class scholarship to continue her study at SCUT in Pulp and Paper Engineering graduate program, which she later found a lack of challenge. She quitted school in 2009 and came to the United States to pursue a doctoral degree in Paper Science and Engineering (PSE) in School of Chemical and Biomolecular Engineering at Georgia Institute of Technology. In 2012, she transferred her major from PSE to Chemical Engineering.