The Potential of using Carbohydrate-Binding Modules (CBMs) for the Characterization and Modification of Cellulose Surfaces

by

Kevin Aïssa

B.Sc., Université Paris-V (Descartes), 2010

M. Sc. Université de Sherbrooke, 2013

M. Sc. A. Université de Sherbrooke, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(FORESTRY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

The Potential of using Carbohydrate-Binding Modules (CBMs) for the Characterization and Modification of Cellulose Surfaces

submitted by Kevin Aïssa in partial fulfillment of the requirements for

the degree of Doctor of philosophy

in Forestry

Examining Committee:

Dr Jack (N) Saddler, Forestry
Supervisor

Dr. Scott Renneckar, Forestry
Supervisory Committee Member

Dr. Vikramaditya Yadav, Chemical and Biological Engineering
Supervisory Committee Member

Dr. Steven Hallam, Microbiology and Immunology
University Examiner

Dr. Mark Maclachlan, Chemistry
University Examiner

Abstract

Cellulose is a structural material that, through its association with lignin and hemicellulose, is recalcitrant to degradation. Although the effectiveness of cellulose hydrolysis is usually assessed via glucose release, typically, cellulase accessibility to the cellulosic substrate is the key limitation that restricts effective enzymatic hydrolysis and has proven much harder to quantify. A novel method, which has the potential to better elucidate the mechanisms involved, involves the use of carbohydrate-binding modules (CBMs). In the work described here, CBM production was optimized, yielding $g.L^{-1}$ quantities of the specific proteins, which were subsequently used to both characterize the surface morphology of lignocellulosic substrates and functionalize cellulose surfaces. A combination of type A and type B CBMs (CBM2a and CBM17) were primarily employed, as they showed binding preferences towards different morphologies within the cellulosic structure. Compared to more established methods the CBM method more accurately predicted enzyme accessibility, indicating that refining did not significantly improve enzyme accessibility at the microfibril level of the cellulosic substrate. In subsequent work, fluorescence-tagged carbohydrate binding modules (CBMs), which specifically bind to crystalline (CBM2a-RRedX) and paracrystalline (CBM17-FITC) cellulose, were used to differentiate the supramolecular cellulose structures in bleached softwood Kraft fibers during enzyme-mediated hydrolysis. Quantitative image analysis, supported by ^{13}C NMR, SEM imaging, and fiber length distribution analysis, indicated that enzymatic degradation predominated in the more disorganized zones during the initial phase of the hydrolysis reaction. This resulted in rapid fiber fragmentation and an increase in cellulose surface crystallinity. Drying decreased the accessibility of enzymes to

these disorganized zones, resulting in a delayed onset of degradation and fragmentation. The use of fluorescence-tagged CBMs with specific recognition sites provided a quantitative way to elucidate cellulose morphology and its impact on enzyme accessibility. This in turn provided novel insights into the mechanisms involved in enzyme-mediated cellulose deconstruction. As well as using CBMs as an analytical tool, the affinity of CBMs for cellulosic surfaces was also used to introduce functionality. When CBM2a-alkyne bioconjugation was used to link polyethylene glycol (PEG) to CNC surfaces via Click reactions, the CBM-PEG modification of cellulosic surfaces increased CNC redispersion after drying and improved suspension stability.

Lay Summary

The global drive towards “greener products” has highlighted interest in renewable materials. Although many biofuels and biomaterials and biochemicals have used sugar or starch as the initial feedstock the use of lignocellulosic residues as more sustainable feedstocks should be advantageous from an energy and environmental perspective. However, effective deconstruction of biomass-to-sugar is hindered by the recalcitrant nature of lignocellulosic materials. This limits the efficiency of enzymatic deconstruction with the accessibility of cellulose to enzymes proving to be a major contributor to recalcitrance. Highly specific cellulose binding modules (CBMs) were used to better assess cellulose accessibility and to better correlate accessibility with the effectiveness of enzyme-mediated hydrolysis. The CBMs were also used to confer functionality to cellulose surfaces in a more environmentally friendly fashion.

Preface

Section 5.1 and **5.2** were in part published in Aïssa K, Novy V, Nielsen F, and Saddler J.N. (2018). *"Use of Carbohydrate Binding Modules to Elucidate the Relationship between Fibrillation, Hydrolyzability, and Accessibility of Cellulosic Substrates."* ACS sustainable chemistry and engineering. I participated in the experimental work, data interpretation, design and drafted the manuscript. I focused on CBMs production, optimization and depletion assay, while Dr Novy and Dr Nielsen focused on hydrolysis and fed-batch fermentation. Dr Novy and Prof. Saddler contributed to the planning, data interpretation and the writing of the manuscript.

Section 5.3 was in part published in Novy V, Aïssa K, Nielsen F, Strauss S, Cieizielski C, Hunt C, and Saddler JN (2019). "*Quantifying cellulose accessibility during enzyme-mediated deconstruction using two fluorescence-tagged carbohydrate-binding modules"* PNAS. I participated in the experimental design and work, data interpretation, and writing of the manuscript. I focused on CBMs production, depletion assay, bioconjugation, fluorescent imaging and NMR. Dr Novy focused on hydrolysis, imaging and data interpretation. Dr Nielsen and Hunt focused on image analysis and coding, Prof Strauss assisted with the NMR analysis and Dr Cieizielski was in charge of the SEM imaging. Dr Novy, Prof. Saddler and myself contributed to the planning, data interpretation and the writing of the manuscript.

Section 5.4 was in part published in Aïssa K., Karaaslan M, Renneckar S, and Saddler JN (2019) "*Functionalizing Cellulose Nanocrystals with Click Modifiable Carbohydrate-Binding Modules*" Biomacromolecules, I designed and performed the experimental work, Dr Karaaslan assisted with

AFM imaging, Prof Renneckar and Prof. Saddler contributed to the planning and the writing of the manuscript.

The introduction will be in part published as a review "*The potential of CBMs as analytical tools for carbohydrates structure investigation*", Biotechnology trends

The materials and methods were in part pulled out from these different publications.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BET	Brunauer–Emmett–Teller
BSA	bovine serum albumin
CBH	cellobiohydrolase
CBM	carbohydrate-binding module
CBU	cellobiohydrolase units
CD	catalytic domain
CFP	cyan fluorescent protein
CLSM	Confocal laser scanning microscopy
CNC	cellulose nanocrystals
Cr	crystalline
CrI	crystallinity index
EG	endoglucanase
FITC	fluorescein isothiocyanate
FPLC	fast protein liquid chromatography

FPU	filter paper units
FQA	Fiber quality analyzer
FTIR	Fourier-transform infrared
HPLC	High-performance liquid chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
Kan	Kanamycin
LB	Lysogeny broth
LCC	Lignin carbohydrate complex
LPMO	Lytic polysaccharide mono-oxygenase
MW	Molecular weight
Na_2CO_3	Sodium carbonate
NaOH	Sodium hydroxide
NBSK	Northern bleached softwood Kraft
NHS	N-Hydroxysuccinimid
NMR	Nuclear magnetic resonance

OD	Optical density
PASC	Phosphoric acid swollen cellulose
PBS	Phosphate saline buffer
PEG	Polyethylene glycol
RMP	refiner mechanical pulp
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SSA	Specific surface area
STEX	Steam explosion
ΔS^o	standard entropy
TB	Terrific broth
TMP	thermomechanical pulp
WRV	water retention value

Acknowledgements

I would like to thank my supervisor, Prof. Jack Saddler, for his philosophy and support throughout my PhD. I would like to express my gratitude to Prof. Scott Renneckar for his guidance and insightful discussions, as well as my other committee members, Prof. Yadav, Prof Withers for their insightful advices and input during this research.

To my dear Post docs, Dr. Vera Novy and Dr. Fredrik Nielsen, whom by their love of science and impeccable work ethic truly inspired me to be a better person, thank you immensely. To Dr. Jinguang Hu and Dr Keith Gourlay, thank you for your precious advices and moral support. I would also like to thank my collaborators that enhanced my PhD experience: Emily Kwan, Kevin Hodgson and Derrick Horne, Prof. Suzana Strauss, Christopher Hunt, Dr. Muzaffer Karaaslan and Dr Mijung Cho. I would also like to thank NSERC and the Van Dusen fellowship for the financial support. Thank you to Dr Vinay Khatri for helping me until the end.

My time at the Forest Products Biotechnology and Bioenergy group has been made all better due to all the great students and staff that are such wonderful people and make bad days look decent and good days look great. Special thanks to my dear friends for the infinite patience and support.

To my extraordinary family and to the best wife in the world, thank you for everything.

“ On ne voit bien qu'avec le cœur. L'essentiel est invisible pour les yeux.”

“It is only with the heart that one can see insightfully. What is essential is invisible to the eye."

- Antoine de Saint-Exupéry

Chapter 1: Introduction

In the early 1900s, most dyes, solvents and synthetic fibers were produced from trees and agricultural crops (Zhang, 2008). However, by the late 1960s, many of these bio-based products had been displaced by petroleum derivatives. With an energy demand projected to grow by more than 50% by 2025 in addition to finite petroleum resources, long term solutions are needed if the world is to break free from its addiction to fossil feedstocks. To try to minimise the dramatic effect that another petroleum crisis would have on the world's energy security and environment, many countries have invested in research into biomass conversion to biofuels, biochemicals and other bioproducts. Shifting away from fossil carbon source to renewable biomass resources will be crucial for the development of a sustainable industrial system and for better management of greenhouse gas emissions.

The concept of a modern biorefinery can be imagined as a parallel to a present day petroleum refinery, but where a sustainably produced feedstock (i.e. renewable polysaccharides and lignin) are fractionated and converted into a mixture of products via different processes and products (Agbor et al., 2011). This could include transportation fuels, co-products, direct energy and high value materials (e.g. cellulose nanocrystals). Biomass conversion to value added products is typically achieved by three different pathways: thermal, chemical and biochemical. Thermal techniques such as pyrolysis and gasification typically convert biomass into solid (char), gas (biogas) and liquid (biocrude/bio-oil) fractions (Ma et al., 2012). However, these techniques suffer from the disadvantages of low selectivity and heterogeneous products. Alternatively, chemical ways of converting biomass into high value-added chemicals are typically hindered by the cost of

chemicals or catalysts as well as their environmental impact. In contrast, the enzyme or microbial mediated "bioconversion" of biomass has the advantages of good selectivity and a low environmental impact (Ragauskas et al., 2006). However, the relative low efficiency of this type of conversion has often been attributed to the heterogeneity and recalcitrance (*i.e.* accessibility) of the biomass substrate (Mansfield et al., 1999). Therefore, to better overcome these bottlenecks, and to better understand the complexity of the biomass substrate would be beneficial.

As mentioned earlier, it is particularly challenging to characterize and measure the recalcitrance of lignocellulosic substrates. Therefore, developing representative and reproducible methods that would allow us to assess the potential hydrolyzability of a substrate would be of considerable value (Mansfield et al., 1999). With the evolution of new microscopy techniques and an ever-growing library of high selectivity probes selected from hydrolytic enzyme sequences, it is now possible to map carbohydrate structures as well as quantify changes that may occur on the substrate surfaces. These probes are called carbohydrate-binding modules (CBMs) due their selectivity in binding to various areas of the cellulosic substrate. Imaging and characterizing substrates with CBMs have provided useful insights into cellulose modification and hydrolysis that will be discuss within the thesis. However, although CBMs have considerable potential, they are not frequently due to the limited availability of the proteins as well as the lack of standardized methods. In the work described here CBM production and purification was optimized and a simple and efficient process to promote their production was developed. A standardized protocol for their utilization to quantify cellulose accessibility was also developed and applied to better understand the relationship between refining, changes in substrate characteristics and overall cellulose hydrolysis. By using

fluorescent-tagged CBMs and confocal laser microscopy, an innovative qualitative and quantitative technique for substrates characterization was developed. The technique is based on the different specificity of two CBM probes that bind non-competitively to crystalline and paracrystalline regions, allowing tracking of variations in the structural organization of the substrate. With this technique, both micro-scale imaging and bulk experiments were conducted leading to a better understanding of the enzymes and substrates behaviour. It was shown that confocal microscopy methods together with the development of analytic codes for image processing could be used to guide the development of cellulase cocktails. This, consequently, greatly improved our understanding of the key mechanisms and rate-limiting steps that occur during cellulose hydrolysis.

In addition to enzyme-mediated deconstruction of cellulose to sugars, another key product from an effective biorefinery is to produce innovative, high performance biomaterials which will be competitive with fossil derived products. This is exemplified by the significant, recent growth of nanocellulose based materials (Dufresne, 2013). Over the last two decades, a growing number of studies have reported the production and utilization of cellulose fibrils or crystals in the nanometer range with applications from biomedical and drug delivery applications to the carbon fiber and battery fields (Abitbol et al., 2016). However, cellulose nanoparticles tend to self-associate because of the presence of interacting surface hydroxyl groups. This property, which contributes to paper sheet strength, makes it challenging to disperse these cellulosic materials in a polymer matrix. For example, these inter-particle interactions can cause aggregation during the preparation of the desired material leading to lesser stability. For aqueous based applications (e.g. hydrogel,

biomedical and drug delivery) and emulsions, green and biocompatible modification is typically required. Specific functionalization requires the drying of the cellulose and the use of organic solvent, which is usually problematic due to the relatively inert nature of cellulose surface in water. In the work described within this thesis, bioconjugated-CBMs were used to bring functionality to cellulose surfaces in aqueous media. As mentioned earlier, CBMs are compact, stable and high affinity proteins that should prove ideal to functionalize cellulose surfaces. Nanocellulose-based materials are typically carbon- neutral, sustainable, recyclable and non-toxic and have the potential for many useful applications.

1.1 Enzymatic hydrolysis of cellulose

In this section, the enzymes and mechanisms involved in enzymatic degradation of cellulose are briefly described. Enzyme factors that influence and limit hydrolysis are highlighted.

1.1.1 Enzymes involved in biomass/cellulose deconstruction

Cellulose hydrolysis can be catalysed by "simple system" enzymes such as exoglucanases (also called cellobiohydrolases, CBHs), and endoglucanases (EGs), or by enzyme complexes called cellulosomes (Zhang and Lynd, 2004). CBHs release cellobiose units from chain ends and degrade cellulose in a processive manner. In comparison, EGs act more randomly along the cellulose chain, creating new chain ends. Once cellobiose is released, β-glucosidases (βG) hydrolyse the cellobiose to glucose. Cellulolytic enzymes often contain one (or more) carbohydrate-binding modules (CBM) and at least one catalytic domain (CDs). These units are structurally and functionally

independent and are connected by an inter-domain linker peptide (Boraston et al., 2004). As described in more details later (**section 1.6),** it is thought that carbohydrate-binding modules promote hydrolysis by increasing the time of contact between the catalytic domain and the substrate as well as specific targeting the substrate. Non-hydrolytic enzymes, such as lytic polysaccharide monooxygenases (LPMO), add a new element to the classic view of hydrolytic cellulose degradation (Hemsworth et al., 2015). It has been shown that the LPMO can attack crystalline areas of the cellulose and oxidize the glycosidic bonds. They use an external source of electrons, facilitating the degradation of crystalline region. In the presence of hemicellulose, accessory enzymes, such as xylanases, are also required (Hu et al., 2011).

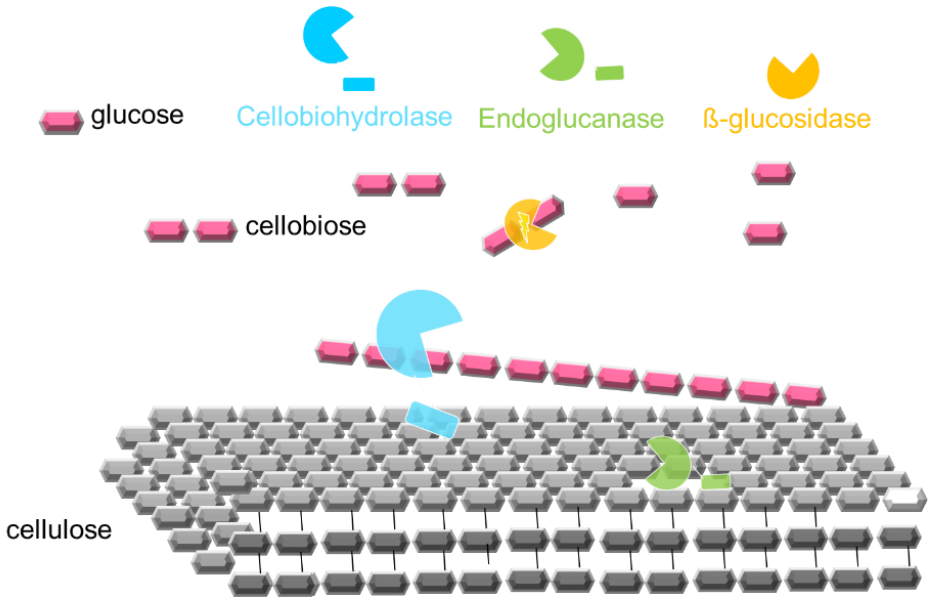


Figure 1: Schematic representation of a simple cellulolytic system. The action of cellobiohydrolase, endoglucanase and b-glucosidase working on their respective carbohydrate substrate

It is commonly observed that the heterogeneous structure of cellulose gives rise to a rapid decrease in rate as hydrolysis proceeds, even when the effects of cellulase deactivation and product inhibition are taken into account (Zhang and Lynd, 2004). Explaining this observation at a mechanistic level is an outstanding issue, with important fundamental and applied implications.

1.1.2 Enzyme factors influencing and limiting the hydrolysis of cellulose

Several factors associated with the nature of the cellulase enzyme system have been suggested to be influential in the hydrolysis process. These include end-product inhibition of the cellulase complex, thermal inactivation, and irreversible adsorption of the enzymes. The problems of end-product inhibition have been largely dealt with through the addition of β-glucosidase, which hydrolyzes cellobiose to glucose, thereby preventing inhibition of cellobiohydrolases by cellobiose.

1.1.2.1 Synergism

Enzymatic synergism occurs when the combined action of two or more enzymes leads to a higher rate of action than the sum of their individual actions. Quantitative representation of the extent of synergism is usually expressed in terms of a ‘‘degree of synergism’’ (DS), which equal to the ratio of the activity exhibited by mixtures of components divided by the sum of the activities of separate components. Different types of synergisms have been proposed in the cellulose hydrolysis literature that include: endoglucanases and exoglucanases; exoglucanases and exoglucanases; endoglucanases and endoglucanases; and exoglucanases or endoglucanases and β-glucosidases (Zhang and Lynd, 2004). It has also been shown that hemicellulolytic enzymes work

synergistically with the cellulolytic enzymes (Hu et al., 2011). By removing the hemicellulose shield these enzymes can enhance enzyme accessibility to the cellulose. Further, hemicellulases and swollenins, (*i.e.* a non catalytic accessory enzymes) have been shown to exhibit some degree of synergism.

1.1.2.2 Traffic Jam

The "roughness" of the crystalline cellulose surface has been shown to lead to the formation of traffic jams of productively bound cellulases (Igarashi et al., 2011). Thus, flattening the surface, removing hindrance, and/or increasing the number of active binding sites by means of pretreatment or combined use of synergistically acting enzymes should reduce the "congestion", improving the kinetics of the cellulase molecules and increasing the efficiency of hydrolysis (Igarashi et al., 2011).

1.1.2.3 Unspecific Adsorption

Unspecific adsorption is dependant on the substrate composition and enzyme physical properties. Lignin and hemicellulose negatively influence the cellulose hydrolysis by irreversibly adsorbing the enzymes, thus preventing their desired action (Palonen et al., 2004). In the case of lignin, especially in steam-exploded substrates, high protein adsorption can be observed on condensed lignin. In previous work where various lignin sources were studied, it was apparent that the extent to which lignin adsorbs enzymes depends very much on the nature of the lignin (Palonen et al., 2004). The work in this thesis expands on this theme in **section 5.3**, showing that, depending on

the substrates and enzyme loading, lignin predominantly acts as a barrier or as an irreversible inhibitor.

When cellulases encounter lignocellulosic substrates, they need to overcome a hierarchy of structural hindrance that restricts the enzymes ability to access and hydrolyse the cellulose. These different structural levels will be described in **section 1.6** on lignocellulosic substrates. The "accessibility" in known to be a key factor that needs to be optimize for better cellulose hydrolysis. However, methods to quantify "accessibility" present their limitations (section 1.8). That is why, in this thesis work, we developed a novel method using CBMs to strengthen our understanding of the changes affecting cellulose "accessibility" to enzymes.

1.2 CBMs: structures, roles and production

The first characterizations of CBMs were published in 1988 (Gilkes et al., 1988; Tomme et al., 1988). Originally, the terminology "CBDs" (cellulose binding domain) was used. However, studies showed that non-catalytic modules also bind to other carbohydrates (Boraston et al., 2004). Thus, the more inclusive term of carbohydrate-binding modules (CBMs) was proposed (Boraston et al., 2004). CBMs are named according to their family (e.g. the family 17 CBM from *Clostridium cellulovorans* Cel5A would be called CBM17). The name of the organism and/or the enzyme can also be included (e.g. CBM17 may be defined as CcCBM17 or CcCel5ACBM17). If a protein contains more than one CBM, a number corresponding to the position of the CBM in the enzyme relative to the N-terminus is typically included. For example, the first CBM is referred to as CBM00-1, the second as CBM00-2, the seventh as CBM00-7 and so forth (Boraston et al., 2004).

1.2.1 Classification

As of 2018, 81 families of carbohydrate-binding modules (CBMs) have been isolated with members of different families showing little if any homology concerning their amino acid sequences (*CAZy database*). Members of the same family have primary and tertiary structures that can be computationally aligned, suggesting that they are descended from common protein ancestors (Carvalho et al., 2014). Fold similarities between CBM families have been demonstrated, which is why "super families" called CBM "tribes" have been suggested. CBMs can also be classified into 3 types according to their substrate specificity (Carvalho et al., 2014). Type A CBMs bind crystalline cellulose, type B CBMs bind amorphous cellulose, type C CBMs bind soluble carbohydrates. The diversity of substrates recognized by CBMs suggests several binding mechanisms (Boraston et al., 2004).

1.2.2 CBM Structures

In nature, CBMs are produced by bacteria and fungi as the binding module of an enzyme that attacks cellulose or hemicellulose. While the amino acid sequences can be very divergent between bacterial and fungal CBMs, the 3D structures are often quite similar, especially regarding the binding site (Blake et al., 2006). Comparisons between similar CBMs from bacteria and fungi have also been studied previously (Tomme et al., 1995). For example, comparisons between CBM1 from CbhI (fungus) and CBM2 from Cex (bacteria) showed that both CBMs had an affinity to bind to crystalline cellulose. However, the binding affinity for bacterial CBM2 was observed to be twice as strong as fungal CBM1 (Tomme et al., 1995).

In this section of the thesis different types of CBMs as well as the specificity of binding of each type are described. The two main CBMs used in the thesis work, CBM2a and CBM17, are discussed in more detail below.

1.2.2.1 Type A CBMs

It has been shown that type A CBMs bind to insoluble, crystalline cellulose and/or chitin. Although many CBMs display aromatic amino acid residues at their binding sites, type A CBMs exhibit a planar binding site, which is complementary to the flat faces presented by cellulose crystals. However, the exact location of CBM binding to the cellulose fiber is controversial and may be influenced by the nature of the cellulosic substrate. Tormo et al. (Tormo et al., 1996) proposed that the binding occurs on the hydrophobic 110 surface, while McLean et al. (McLean et al., 2000) have contested this claim, saying that in a perfect cellulose crystals, the surface area presented by the 110 surface is too small for the binding site of CBMs. In contrast, the authors proposed that the binding sites are more likely to occur on the 110 and 010 surfaces. A study by Lehtio et al. (Lehtiö et al., 2003) used transmission electron microscopy to probe the location of the CBM binding site on crystalline cellulose. They also concluded that CBM1 and CBM3a bind to the hydrophobic 110 surface and suggested that these regions are often disrupted, thus providing a larger surface area that could be more accessible. Although there is no consensus on where type A CBMs exactly bind, it is accepted that they will have a greater affinity for the most hydrophobic cellulose surface.

The binding of CBM to insoluble microcrystalline cellulose is influenced by its structure and the lack of conformational freedom of the substrate. As the conformation of each cellulose chain on

the fiber surface of crystalline cellulose is essentially fixed, the chain is unlikely to undergo a loss in conformational entropy upon binding to CBMs. This, combined with the negative ΔCp (heat capacity) values that indicate a significant dehydration effect, provides an explanation of the large favorable ΔS^{o} (standard entropy) observed for type A CBM adsorption to crystalline cellulose (Boraston et al., 2004). This factor indicates the fundamental difference between binding of soluble and insoluble polysaccharides and might explain why type A CBMs have a low affinity for soluble carbohydrate. As the tight binding of soluble oligosaccharides requires a loss in conformational entropy and the type A CBM binding surface is not flexible, Creagh et al.(Creagh et al., 1996) proposed that, when the water molecules released from the protein when CBMs bind to their target, this increases the entropy of the system. It has also been shown that thermodynamic forces (Van Der Waals) play an important role at the CBM/cellulose interface (Creagh et al., 1996). Van Der Waals interactions stabilize the CBM/fiber binding through a number of hydrogen bonds and C-H/π interactions. However, it has been suggested that, if there is a physical blockage (by lignin, hemicellulose or other enzymes) along the cellulose surface, the Van Der Waals interactions prevent the diffusion of the CBM from the fiber surfaces, causing a decrease in enzyme processivity (Igarashi et al., 2011). The binding of type A CBMs can be considered to be irreversible since its constant of dissociation is a lot smaller than that of the binding constant (McLean et al., 2000). Advances in analytic chemistry, such as neutron crystallography and neutron reflectography, have allowed researchers to better understand the interaction between a protein and insoluble substrate. Recently, neutron crystallography work has suggested that water molecules play a key role in the binding site for different CBMs (Fisher et al., 2015). Atomic force spectroscopy/microscopy has been used to elucidate the substrate recognition of CBM3a and the

binding energies for some CBMs (King et al., 2015; Nigmatullin et al., 2004; Zhang et al., 2013a; Zhang et al., 2014). However, the exact mechanism has not yet been completely resolved because direct analysis in aqueous media is still a challenge. In the described work, CBM2a was chosen as the probe for crystalline cellulose as, 1) it has such a strong affinity for crystalline cellulose and its binding is considered irreversible. 2) Previous studies concerning its expression, production, and interaction with cellulose provided a base of knowledge which will be advantageous for further development. Furthermore, 3) engineered CBM2aH6 (his-tagged) were available for the project and presented an alternative to native CBM2a for an easier purification (Boraston et al., 2001; Hasenwinkle et al., 1997; McLean et al., 2000).

1.2.2.2 Type B CBMs

Unlike type A CBMs, which show a greater affinity towards crystalline cellulose, type B CBMs preferentially interact with single polysaccharide chains and have a higher affinity for the amorphous regions of cellulose. As a result, the binding site for type B CBMs is also different. NMR studies and X-ray crystal structures have revealed that the carbohydrate-binding sites of Type B CBMs are extended, shaped as “grooves” or “clefts”, and are comprised of several “sub-sites” that are able to accommodate the individual sugar units of the polymeric ligand (Notenboom et al., 2001). The depth of the binding sites is also flexible and can vary from very shallow to being able to accommodate the entire width of a pyranose ring (Boraston et al., 2004). For type B CBMs, the binding energy is correlated to the degree of polymerization (DP) of the carbohydrate substrate. For example, type B CBMs do not bind with substrates equal or smaller than three glucose units. However, the increase in binding energy with an increase in cellulose DP reaches a maximum and

plateaus after 6 glucose units. Similar to type A CBMs, aromatic residues also play a role in ligand binding and the orientation of these amino acids are key to influencing CBM specificity. Direct hydrogen bonds also play a key role in defining the affinity and ligand specificity of Type B glycan chain binders (Zhang et al., 2013a). The binding of type B CBMs is thought to be enthalpy driven, which means that the substrates and enzyme structure are flexible and the binding is driven by the Van der Waals interactions (Boraston, 2005). In the work described here CBM17 was chosen to probe the amorphous cellulose region as CBM17 is known to have high affinity for the amorphous region and it its interaction with the substrate has been studied before (Notenboom et al., 2001).

1.2.2.3 Type C CBMs

In contrast to type B CBMs, that cannot bind to carbohydrates with less than 3 glucose units, type C CBMs demonstrate binding specificity for one to three sugar units (Boraston et al., 2004). The distinction between the structure of Type B and Type C CBMs is often subtle, as they both exhibit a groove-like binding site. Nevertheless, despite folding similarities, crystallography study has shown that the hydrogen bond network between protein and ligand is more extensive in Type C CBMs than in Type B modules. In addition, C-H/π interactions (e.g. interactions between the cellulose and aromatic residues) also play a smaller role in the binding of type C.

1.2.3 The details of CBM2a and CBM17 bindings

As CBM17 and CBM2a have been well studied, previous modeling, binding isotherms and ITC data was able to indicate how these proteins would function and the rationale behind their

specificity (Abbott and Boraston, 2012; Boraston et al., 2000; Boraston et al., 2001; Boraston, 2005).

1.2.3.1 The binding of CBM2a

It has been shown that the binding of CBM2a to crystalline cellulose is irreversible yet dynamic (McLean et al., 2000). It has been shown that the CBMs move in two dimensions over the cellulose surface without ever fully dissociating from it. This is typical for type A CBMs (Boraston et al., 2004). The type A CBM's constant of association (Ka) is around 106 M^{-1} for insoluble cellulose. The binding capability of CBM2a to the cellulose surface is attributed to six amino acid residues on the binding face of the module (**Figure 2**). Of those residues, three tryptophans (W17, W54 and W72) are thought to be of particular importance to the cellulose binding capability of CBM2a. Oxidation studies, which targeted solvent exposed aromatic residues, demonstrated that oxidation of any tryptophan residue led to a decrease in binding capability (McLean et al., 2000). Furthermore, when complete oxidation of all three tryptophan residues was achieved, CBM2a could not bind to cellulose anymore, even though the 3D structure of CBM2a remained the same. The results of this oxidation study support the hypothesis that the hydrophobic tryptophan residues play an important role in the binding reaction. In addition, site-directed mutation and Langmuir-type adsorption isotherm analysis were also used to help determine the role of the other neighboring residues that contribute to the overall energetics of CBM2a binding to cellulose (Boraston et al., 2001; Boraston et al., 2002). Through this study asparagine residues (N15 and N18) were shown to participate in binding via hydrogen bonds and it was suggested that the glycine residue (Q52) might also play a similar role (McLean et al., 2000).

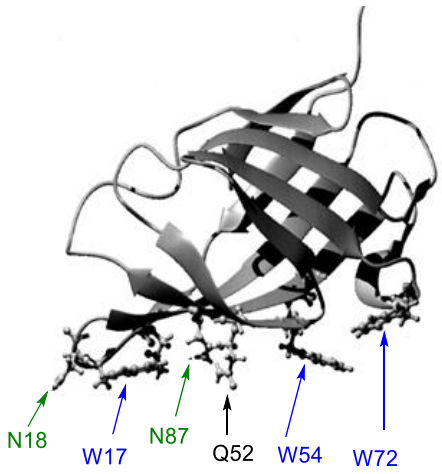


Figure 2: Structure of CBM2a (adapted from McLean et al., 2000), the main amino acid residues involved in the binding are represented here: 3 tryptophans (W, blue), 2 asparagines (N, green), 1 glutamine (Q, black)

1.2.3.2 The binding of CBM17

CBM17 is a type B CBM and, as mentioned earlier, binds preferentially to non-crystalline cellulose with an optimal binding affinity for single cellulose chains of at least 6 glucoses units. Recent analysis of the crystal structure of CBM17 has shown that binding occurs in a cleft on the surface of the molecule that involves two tryptophan residues and several charged amino acids establishing hydrogen bonds (**Figure 3**). Thermodynamic binding studies and alanine scanning mutagenesis allowed mapping of the CBM17 binding site. In contrast to the binding groove characteristic of family 4 CBMs (another type B CBM), the family 17 CBMs appear to have a very shallow binding cleft that may be more accessible to cellulose chains in non-crystalline cellulose than the deeper binding clefts of family 4 CBMs. The structural differences between these two modules may reflect different binding sites on cellulose surfaces (Notenboom et al., 2001).

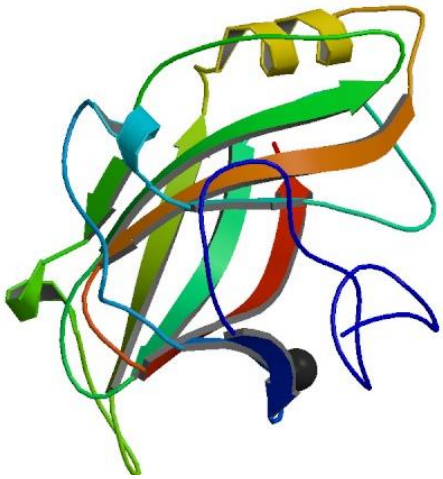


Figure 3: The structure of CBM17 is composed mainly of β-sheet (arrows). CBM17 requires a sodium ion (black sphere) for its stable conformation and optimal binding (adapted from the CAZy database).

1.3 Roles and functions of CBMs

Carbohydrate-binding modules play two essential roles: They both increase enzyme concentration on the carbohydrate surface (proximity effect) and they also preferentially bind to specific substrate characteristics (substrate targeting). Although it has been disputed, it has also been suggested that CBMs play a role in amorphogenesis (fiber disruption, enzyme mediated accessbility). In addition to these mechanisms, CBMs have been suggested to have play other roles in cell metabolism, energy storage, antibiosis and immunological recognition (Levy and Shoseyov, 2002). However, the contribution of CBMs to these mechanisms are still debated.

1.3.1 Non catalytic activities

It has been reported that CBMs increase the effective concentration of the enzymes hydrolytic unit(s) on the surface of the carbohydrate substrate (proximity effect) (Boraston et al., 2004). As mentioned, during hydrolysis, CBMs assist in substrate recognition (targeting) and increase the contact time between cellulases and the substrate. However, CBMs are not mandatory for cellulose hydrolysis to occur. Várnai et al. (Várnai et al., 2013) showed that an increase in hydrolysis yield could still be achieved when using cellulases without CBMs by increasing the substrate concentration. The improvement can be explained by the faster equilibrium between free and bound enzymes without CBMs, avoiding non-productive and irreversible binding by the CBMs and, therefore, allowing for more efficient reuse of the enzymes (Várnai et al., 2013). It is important to point out that the authors used pure amorphous cellulose for their studies, which is very different from more complex lignocellulosic substrate. CBMs binding studies have mainly used model substrates such as Avicel® or amorphous cellulose, which does not represent the

complexity of real lignocellulosic substrate. One of the goals of the thesis, after careful optimization and calibration of the methodology on model substrates, was to expand the CBM binding study to substrates that are more complex. In this way we hoped to obtain increased knowledge concerning biomass recalcitrance and hydrolysis (see **section 5.2**).

1.3.2 Fiber disruption

Fiber disruption has been suggested to occur during cellulose depolymerisation and CBMs were thought to have an effect on cellulose fiber disruption (Din and Al, 1991; Gustavsson et al., 2005; Kataeva et al., 2002; Wang et al., 2008). However, this claim is controversial and numerous papers have contradicted this hypothesis (Carrard et al., 2000; Herve et al., 2010; Josefsson et al., 2008; Várnai et al., 2013). In 1991, when Din et al. used electron microscopy and fluorescent probes to try to demonstrate that CBMs were able to disrupt the substrate, they noticed a production of small particles when they stirred cotton fibers with CBM (Din and Al, 1991). However, no proof at the molecular scale was brought to support the hypothesis. Other results that were published supporting this claim were often not reproducible or had some flaws in the experimental design (e.g. using X-ray to measure crystallinity in a dry state, or questionable control for Quartz Crystal Micro-Balance (QCMB)). These contradictory observations were largely a result of the limitations of current analytic techniques to be performed for liquid/solid interfaces. This has limited the accuracy of measurements on the effect of CBMs on the cellulose structure. However, since then researchers have tried to develop new tools and methods to characterize the interface of insoluble cellulose and water-soluble enzyme (King et al., 2015). To our knowledge of the latest literature,

it seems that the CBMs do not play a role in fiber disruption, leaving the substrate surface unchanged after binding.

Due to their specificity and non-hydrolytic behaviour, CBMs have the potential to be a useful tool to help us better understand the complex structure of lignocellulosic substrates. They can be used to understand the mechanism of hydrolysis as well as the topology of substrate surfaces, (discussed in **section 5.2 and 5.3)**. CBMs have been used for several applications and present significant potential for new biomedical application such as bioassay, drug delivery or biomaterials.

1.4 CBMs: powerful analytic tools

CBM specificity and strong affinity to carbohydrate structures make them ideal probes for lignocellulosic substrates. As shown within this thesis, they can be used to measure substructure ratios or to "map" changes to the cellulosic surface.

1.4.1 Surface characterization: crystallinity/paracrystallinity

Gourlay et al. (Gourlay et al., 2012) used carbohydrate-binding modules, which preferentially bind to specific cellulosic substructures to assess the distribution of crystalline and paracrystalline cellulose at the fiber surface. In addition, CBMs were also used to track changes in surface morphology when Swollenin was added, a protein thought to induce amorphogenesis (Gourlay et al., 2015). Gao et al. (Gao et al., 2014) used CBM adsorption on Avicel® at different time points during the enzymatic hydrolysis reaction to indicate that most of the amorphous cellulose is inside the Avicel® particles and thus could not be readily accessed by the cellulases during the first stage of hydrolysis. This observation supported the proposed "layer by layer" deconstruction mechanism (section 1.1) (Gao et al., 2014). These authors described the relationship between the crystallinity index (CrI) and the hydrolysis rate using Avicel® as a model substrate. They suggested that the CrI might be a misleading concept because it requires a drying step prior to the measurement. Monitoring changes of cCAC (crystalline) and aCAC (amorphous) clearly suggested that during the whole hydrolysis period a significant fraction of amorphous cellulose was hydrolyzed simultaneously with crystalline cellulose rather than in a two-step process. The authors further suggested that the accessibility of cellulose to cellulases was far more important than the CrI in determining the hydrolysis rate (Gao et al., 2014).

1.4.2 CBM adsorption and measurement of cellulose accessibility: kinetics and measurement

As mentioned earlier, the quantitative determination of cellulase enzyme accessibility to cellulose can be measured with CBMs, with or without fluorescent tags (Gao et al., 2014; Gourlay et al., 2015). There are two key experimental approaches for quantifying CBM/insoluble polysaccharide interactions: (1) the adsorption assay, and (2) isothermal titration calorimetry (ITC). As discussed below, the kinetics of binding and equilibrium constant are important parameters that will have to be determined if CBMs are to be used to measure enzyme accessibility to cellulose.

1.4.2.1 Adsorption assay

The methodology describe below is derived from the general binding equation:

$$H + G \rightleftharpoons GH$$

H = Host (cellulose), G = Guest (protein) and HG = CBM-carbohydrate complex. At the equilibrium, the association constant is given by:

$$\mathrm{K_a} = \frac{[GH]}{[G][H]}$$

Abott and Boraston suggested this nomenclature because it provides some flexibility to represent the different CBM–carbohydrate interactions (Abbott and Boraston, 2012).

Considering:

$$[H_{total}] = [H] + [GH]$$

(Where [H] represents the unoccupied cellulose and $[H_{tot}]$ the total cellulose sites). We can rearrange the previous equation:

$$K_a = \frac{[GH]}{[G]([H_{tot}] - [GH])}$$

$$K_a[G] = \frac{[GH]}{[H_{tot}] - [GH]}$$

$$\frac{[H_{tot}]}{[GH]} = 1 + \frac{1}{K_a[G]}$$

Where $[GH]/[H_{tot}]$ is the fraction bound.

Adsorption assays are a simple method used for qualitative and quantitative assessment of CBM interactions with insoluble carbohydrates. For example, cellulose powder can be easily weighted, resuspended in buffer and then analyzed. To perform the experiments, variable CBM concentrations [G] were added at identical cellulose concentrations [H]. The range varies between 1/10 of the dissociation constant to 10 times excess, which ensured substrate saturation. Furthermore, the protein samples can then be used as a baseline or control to compensate for unspecific adsorption and protein precipitation (Abbott and Boraston, 2012). Following binding equilibrium, the samples were centrifuged. The supernatant was removed and analyzed qualitatively by UV absorbance. The protein concentration in the liquid can then be determined at 280 nm using a CBM extinction coefficient. Similar principles apply to the use of fluorescence

spectroscopy to monitor CBM–carbohydrate interactions rather than measuring changes in tryptophan or tyrosine absorbance.

From the previous equation, we can rearrange to the equation below:

$$\frac{[GH]}{[H_{tot}]} = \frac{K_a[G]}{1 + K_a[G]}$$

In saturation conditions, it can be written:

$$[P_{tot}] = [P] + [PH]$$

Where we write [P] instead of [G] for more clarity. $[H_{tot}]$ correspond to the total concentration of available binding site, which is named N_o. So, under saturation conditions, the equation is:

$$\frac{[PH]}{N_o} = \frac{K_a[P]}{1 + K_a[P]}$$

Under conditions where there are two non-interacting CBMs with significantly different binding affinities, the following two-site binding model can be used:

$$[PH] = \frac{N_{o1}K_{a1}[P]}{1 + K_{a1}[P]} + \frac{N_{o2}K_{a2}[P]}{1 + K_{a2}[P]}$$

Where N_{o1} and K_{a1} represent the binding parameters for one class of binding site and N_{o2} and K_{a2} the binding parameters for the other (Abbott and Boraston, 2012). Thus, this equation should be able to model CBM2a and CBM17 competitive binding. In theory, we should be able to quantify disruption areas on the surface of a given substrate.

1.4.2.2 Isothermal titration calorimetry

Calorimetry measures heat changes during a chemical reaction. Constant improvements to microcalorimeters have allowed researcher to refine the methodology (Dam and Brewer, 2002). Isothermal titration calorimetry (ITC) measures the increases (exothermic) and decreases (endothermic) in temperature within a solution, which correspond to the titration of ligand (i.e., carbohydrate, G) into acceptor (i.e., protein, H). It is possible to use ITC to determine Gibb's free energy, enthalpy, entropy and the stoichiometry of a designated interaction. In addition, reaction series at varying temperature can be performed to determine the heat capacity of the system. This method is mostly applied to soluble carbohydrate as it is challenging to measure insoluble substrate.

In the work described within the thesis, we developed an analytic method using CBM adsorption and CBM imaging to measure enzyme accessibility to cellulose and to better understand the structural differences that occur on the cellulose surface. We were able to produce new biomaterials, using bioconjugated CBMs to bring functionality to the cellulose surfaces. Thus, for background, the nanostructures of cellulose and their properties are described below as well as cellulose chemical reactivity.

1.4.2.3 Inactivated enzyme adsorption

Similar to the CBM adsorption, researchers have used cellulases to differentiate between specific substructures of cellulose. There are three main approaches for cellulases adsorption: low temperature, mutant or inhibitor (Lee et al., 1982). Low temperature adsorption typically involves

incubating either a mixture of cellulases or a purified cellulase with the substrate and quantifying the total amount of protein adsorbed. Although this technique may give an accurate representation of the amount of accessible surface area of the substrate, there are two key problems. First, the adsorption study needs to be carried out at 4 °C to prevent any hydrolysis. However, this complicates the experimental set up, especially if microscopy is required. It is possible that cellulase inhibitors could be used to prevent substrate hydrolysis during protein adsorption experiments at relevant temperatures (i.e. 50 °C), however, the inhibition of several cellulases of an unknown commercial cocktail can be tedious. Finally, mutagenesis of a key amino acid in the catalytic site can be performed, making the hydrolysis impossible. This method is mostly used to actually investigate cellulase mechanisms and the use of the mutated enzyme as a probe for lignocellulosic substrate is limited by the little availability of the mutated enzyme and the different field of expertise required. In general, CBMs are more compact and stable than cellulase and a stock solution can be kept in the fridge for an extended period of time; they offer a great versatility without additional work.

1.4.3 CBMs and visualization techniques: targeting effect

The use of CBMs as an analytical tool was first introduced when a bioassay was developed for characterizing pulp fiber surface using cellulases (McCartney et al., 2004). Novel molecular probes for the detection of polysaccharides in plant cell walls were developed using different types of CBMs. In their approach, recombinant CBMs were fused to polyhistidine tags and anti-polyhistidine antibodies were used to detect polysaccharide-CBM interactions (Jamal-Talabani et al., 2004). In the same year, Jamal-Talabani et al. (Jamal-Talabani et al., 2004) proposed that

CBMs could be used for mapping the “glyco-architecture” of plant cells. Considering their specificity towards different substrate characteristics, CBMs were proposed to be capable of mapping wood tissue and lignocellulosic fibers. Therefore, CBMs were conjugated or fused with a fluorescent tag and then imaged using confocal microscopy.

1.4.3.1 Confocal Laser Scanning Microscopy

Confocal laser scanning microscopy (CLSM) is often used in combination with fluorescent-tagged protein to image aqueous system, and has been previously utilized with CBMs to image wood tissue and fibers (Hidayat et al., 2015; Zhu et al., 2011). The CLSM method is described in more detail in **section 5.3** where a quantification method was developed using sub-structure specific fluorescent CBM and automated data analysis. The CLSM provided the capacity for non-invasive imaging in 2 and 3 dimensions.

1.4.3.2 Atomic force microscopy

Atomic force microscopy (AFM) provided image resolution to the fraction of nanometer scale. The AFM imaging was used in combination with CBM and quantum dots to gain insights to CBM binding onto nanocellulose structures (Arslan et al., 2016; Henriksson et al., 2007; King et al., 2015; Lee et al., 2000). With AFM it is also possible image single molecule and scan molecular defect in a given substrate.

Table 1: Summary of previous work that used carbohydrate-binding modules (CBMs) to better elucidate enzyme-mediated changes in cellulose

CBMs	Tag	Substrates	Primary goal of work	Ref
CBM2a(H6) (*C. fimi*)	NA	Avicel, PASC, cotton fiber, CNC, cellulose II and III	Influence on hydrolysis, tracking changes in cellulose accessibility to enzymes during fiber swelling	(Ali et al., 2001; Hu et al., 2016a; Kljun et al., 2011)
CBM3 (*C. thermocellum)*	GFP	Avicel, Kraft pulp, cellulose II	Tracking of changes in cellulose accessibility to enzymes during hydrolysis	(Gao et al., 2014; Hong et al., 2007; Kljun et al., 2011; Široký et al., 2012)
CBM4-1 (*C. fimi*)	*	Cotton fiber	Tracking paracrystalline structures	(Kljun et al., 2011)
CBM15 *(C. japonicus)*	mOrange2	Kraft pulp	Monitoring the xylan	(Hebert-Ouellet et al., 2017)
CBM17 *(C. cellulovorans)*	Non, CFP mCherry	Avicel, Kraft pulp	Monitoring transition structure and paracrystalline cellulose	(Gao et al., 2014; Gourlay et al., 2012; Hebert-Ouellet et al., 2017; Kljun et al., 2011)
CBM28 *(C. cellulovorans)*	CFP	Organosolv pretreated hardwood, cellulose II	Monitoring transition structure and paracrystalline cellulose	(Kawakubo et al., 2010; Široký et al., 2012)
CBM44 *(C. thermocellum)*	NA	Avicel, PASC, cotton fiber, CNC, cellulose II and III	Tracking of changes in cellulose accessibility to enzymes during fiber swelling	(Gourlay et al., 2015; Hu et al., 2016a)
CBM27 *(C. cellulovorans)*	CFP	Kraft pulp	Monitoring the mannan	(Hebert-Ouellet et al., 2017)

NA: not applicable, CFP: cyan fluorescent protein, GFP: green fluorescent protein.

* Some CBMs have been implemented with His-tags and have been analyzed using immunology-based methods.(Kljun et al., 2011)

1.5 Other CBM applications

CBMs have been used in several research fields, mainly as a fusion partner to another protein. The biotechnological applications of fused CBMs have been reviewed. In this section, the major concepts and breakthroughs are reported (Oliveira et al., 2015).

1.5.1 CBMs and protein engineering

The potency of CBMs as analytical tools has also been demonstrated when fused with another protein. For example, in biomedical diagnostics, CBMs were fused with transmembrane protein to immobilize cells. Additionally, Shoseyov (Shoseyov, 1999) developed a system based on CBMs for rapid detection of pathogenic microbes in food samples. Using this method, CBMs were conjugated to a bacterium binding protein such as an epitope-specific monoclonal antibody and then loaded onto a cellulosic matrix. Ofir et al. (Ofir et al., 2005) also published a new strategy for the production of microarrays based on CBMs/cellulose for HIV diagnosis.

The most developed application of CBMs is the application for protein purification strategies. Because large-scale protein purification is often very costly, it is important to effectively isolate and purify specific biomolecules of interest. That is why a methodology which used CBMs to "tag" proteins was developed (Boraston et al., 2001). This strategy consists of fusing the CBM to the biomolecules of interest via an easily cleavable linker. During the purification steps, the mixture containing the molecule of interest is then filtered through a cellulose-based column (Lowe, 2001). Cellulose presents a great advantage as a purification media: it is inert, it does not bind non-specific protein, it is inexpensive and it is biocompatible (Oliveira et al., 2015). As a

result, the specific affinity of CBMs for cellulose will result in the retention of the molecule of interest within the column, which can then be eluted at high purities and recovered by breaking the linker (Levy and Shoseyov, 2002).

Fusion of proteins with CBMs have also been used in cell immobilization technology, for applications from ethanol production and phenol degradation to mammalian cell attachment and whole-cell diagnostics (Levy and Shoseyov, 2002). Whole-cell immobilization to cellulosic materials was first demonstrated using bacterial CBMs (Kleinman, 1987). CBMs successfully helped cell adhesion to cellulose, leading to a good proliferation on carbohydrate-based biomaterials that can be used for biomedical applications (Mordocco et al., 1999).

Another example of application for fused CBM was demonstrated when different enzymes were fused with CBMs to increase their catalytic activity (Zhao et al., 2013). CBMs have been fused to cellulases and lipases to enhance reaction yields (Tang et al., 2014). The enzyme combined with the CBMs allowed for a longer contact time of the enzyme with the cellulosic substrate. Fox et al. (Fox et al., 2013) reported that multiple CBMs coupled to the same cellulase can lead to a synergistic effect on cellulose hydrolysis, suggesting that optimization of substrate targeting improves hydrolysis.

In other research, the fusion of graphene binding proteins with two CBMs was used in applications such as self-assembly nanofibrillar cellulose (NFC) (Varjonen et al., 2011), construction of nanocomposites of graphene and NFC, and the preparation of drug nanoparticles and their immobilization in NFC for increased storage stability (Valo et al., 2011).

In the past, CBMs have mainly be used in tandem with another protein, helping purify the other protein or its activity. More information on fusion proteins with CBMs can be found in the review by Oliveira et al. (Oliveira et al., 2015). However, bioconjugation, which involves the chemical modification of the protein, have not been studied as much. As will be described in more detail in the proposed thesis work, the bioconjugation technique is a versatile method. It can bring a variety of functionalities onto a protein such as fluorescent tags, polymers, or cross-linker, without any genetic modification. In the next section, the way that fused and bioconjugated CBMs have been used to improve physical properties of cellulose fibers is described as this will be the focus of the proposed work in **section 5.4**.

1.5.2 CBMs and cellulose modification

Although only a few studies have reported the use of CBMs for fiber modification or material improvements, this field may increase in coming year considering the increasing interest in carbohydrates-based materials (Abdul Khalil et al., 2014; Babu et al., 2013). Pala et al. (Pala et al., 2001) demonstrated that applying CBMs to secondary paper fibers improved their drainability and resulted in paper with improved mechanical properties. They proposed that CBMs affect the interfacial properties of the fibers in both fiber-water and fiber-air interactions. Bioconjugated CBMs have also been shown to improve other cellulose based materials. Researchers reported on the production of a novel papermaking reagent by covalently linking anionic polyacrylamide (A-PAM) to a CBM originating from *Trichoderma viride* (CBM–A-PAM) (Kitaoka and Tanaka, 2001). They observed that the tensile strength of paper prepared from CBM–A-PAM was increased. Other studies reported the construction of a bifunctional cross-linking molecule

composed of starch and cellulose binding modules (Levy et al., 2002). As expected, the molecule was able to bind soluble and insoluble starch to cellulose.

1.6 Lignocellulosic substrates: composition and structure

Cellulosic substrates should be a particularly attractive feedstock for the production of biofuels, chemicals and materials because of their relatively low cost, abundance and sustainable supply (Ragauskas et al., 2006). Although biomass conversion has been a major focus of research and development over the last 40 years (Alonso et al., 2010; Zhang, 2010), the enzyme-mediated conversion has been hindered by substrate recalcitrance that protect the carbohydrate components from degradation (Arantes and Saddler, 2011). In this first section, the lignocellulose architecture and composition are described with the emphasis on the substrate characteristics that limit enzymatic hydrolysis. In the following section, the methods that were developed to counter this natural recalcitrance are discussed. As covered in more detail, to expose the cellulosic component to enzymes, the substrate usually needs to be pretreated. Drying, which is often required for efficient biomass transportation, has been shown to negatively influence substrate accessibility. Drying is particularly problematic when choosing a substrate to model enzymatic hydrolysis. As will be described in more detail, pulp and paper procedures used to generate substrates were showm to be more representative of realistic substrates than model substrates such as Avicel®.

1.6.1 Lignocellulose constituents

Biomass is composed mostly of cellulose, hemicelluloses and lignin. The ratio of cellulose, hemicelluloses and lignin depends on the plant species (softwood, hardwood, agricultural residue) (Cai et al., 2017). Cellulose is a long chain of cellobiose monomers, formed by β1-4 linked glucose units. Glucose rings adopt the "chair" conformation, where all substituents and the glycosidic bonds are in the equatorial position (Trache et al., 2016). The cellulose chains are organized in supramolecular chains of microfibrils with nanometer-scale diameters and micrometer-scale lengths. A cluster of microfibrils forms the macrofibrils, which are the building block of the plant cell walls (**Figure 4**) (Rubin, 2008). The cellulose structure is embedded in the hetero-matrix of plant cell walls, which also includes lignin and hemicellulose. While cellulose is a long linear polymer, hemicelluloses are branched polysaccharides with low degrees of polymerization (100–200 units) (Gírio et al., 2010). Hemicelluloses are made from different sugars such as glucose, mannose, galactose, and xylose and their ratio depends on the plant species (Gírio et al., 2010). The hemicellulose connects the lignin to the cellulose to form lignin-carbohydrate complexes. Although lignin is a non-linear polymer made of monolignols, its complete structure is not yet fully elucidated (Ponnusamy et al., 2019). The lignin composition in monolignols varies depending on the plant source. While the polysaccharide components of plant cell walls are highly hydrophilic, and thus permeable to water, lignin contains hydrophobic groups, which make it much less hydrophilic. As a result, lignin protects the cellulose fibers from degradation and restricts water absorption into the plant cell walls (Rubin, 2008). Hemicellulose and lignin play important

roles in the lignocellulose biomass recalcitrance by restraining the cellulose accessibility to enzymes and binding to the enzymes (Mansfield et al., 1999).

1.6.2 Lignocellulose structure and level of organization

As briefly described in the previous section, there are several scales of biomass organization, from the fiber to the nano-structures of the microfibrils with each affecting enzymatic hydrolysis. The fiber represents the largest scale of organization for pretreated substrates. At this scale, the cell lumen allows the diffusion of enzymes (Meng and Ragauskas, 2014), which is not a critical barrier because its size is normally in the range of micrometers. At the intermediate level or the mesoscale, the macrofibrils are arranged into sheets within the plant cell wall. Pores and cracks are typically only 20–100 nm and can represent a barrier to enzymes, limiting their penetration inside the substrate (Esteghlalian et al., 2001). Depending on the lignin type and hemicellulose content, the enzyme diffusion can be slow (Luterbacher et al., 2013a). Within these macrofibrils, adjacent microfibrils are held together through complex interactions involving hemicelluloses and lignin. It is at the level of the microfibril that the accessibility is the most restricted (Mansfield et al., 1999). At this point, it is important to clarify the vocabulary used. Depending on the field where they are encountered, the microfibril can be called nanofibril (material science) or elementary fibril (plant biology). In this thesis we will use microfibril in the context of the bioconversion (section 5.2 and 5.3) section and nanofibril for the nano-material section (section 5.4).

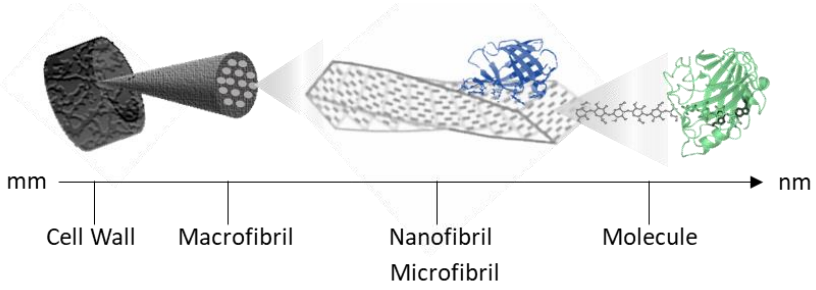


Figure 4: Diagrammatic representation of relative size of a typical pulp fiber, macrofibril, microfibril and cellulose chain (glucan chain) in comparaison to CBM2a and CBM17.

1.6.3 Lignocellulose pretreatments, pulping and drying and summary of their effects on substrate morphology

Substrate-related factors which inhibit the hydrolysis are thought to be a major bottleneck in lignocellulose valorization and limited enzyme accessibility to the cellulose (Rahikainen et al., 2013a; Rahikainen et al., 2013b). In this section, a general summary of different pretreatments and pulping techniques with their effects on the substrates are summarized, as well as the effects of drying.

1.6.3.1 Pretreatments

Numerous physical and chemical pretreatment methods have been developed to try to overcome biomass recalcitrance, including dilute acid, “organosolv” process, steam explosion (STEX), mechanical refining, ionic liquid or ammonia fiber expansion. The changes in lignocellulosic structure during these commonly applied pretreatment technologies have been reviewed (Kumar et al., 2009). Although the chemicals and mechanisms of each pretreatment are different, the final objective is the same, increasing cellulose accessibility to enzymes by opening up this complex

structure. Although, cellulose accessibility could be increased by delignification, the goal is to develop cost-effective pre-treatments for commercially feasible process (Zheng et al., 2009).

Steam explosion and mechanical refining pretreatment make use of little or no chemicals and therefore should have the lowest recycling and environment cost (Dou et al., 2016; Grethlein and Converse, 1991). After steam explosion treatments, the particle size generally decreases. Furthermore, higher steam temperatures and longer retention times result in more homogeneous fiber-like material. In the case of high severity factor, hemicellulose can be lost in the liquid faction and lignin can recondensed into droplet at the surface of the fiber. The mechanical refining of the fibers leads to their fibrillation, which is important to a number of applications (Abdul Khalil et al., 2014; Varjonen et al., 2011). Although fibrillation of the substrate increases the specific surface area, in many cases, lignin is still problematic in mechanically refined substrate as it reduces cellulose accessibility, leading to poor hydrolysis yield.

1.6.3.2 Pulping methods and pulp characteristics

To remain viable, traditional producers of pulp and papers need to increase revenue by producing bioenergy and biomaterials in addition to wood, pulp, and paper products. In the so-called "integrated products biorefinery", the traditional products of the pulp and sector could be increasingly valorized by the development of new product lines. As an example, during the last century, the pulp and paper sector acquired tools to better characterize their substrates, particularly methods related to pulp and paper making. However, these methods also provide useful indicators

regarding substrate property, which are often complementary to other existing characterization methods.

From another point of view, studying enzyme characteristics and structural effect of pretreatment on a system may require the simplification of the substrate to limit the many factors that can influence data interpretation. However, oversimplification may lead to inaccuracy in measurements. For example, microcrystalline cellulose Avicel® is commonly used to represent lignocellulosic substrates. However, many characteristics of Avicel® like its size, fiber structure or homogeneity is far from the types of more realistic substrates that will be used in a biorefinery, likely leading to inaccurate predictions. Therefore, it is crucial to choose model substrate that will keep as many identical properties as a "biorefinery" substrate (e.g. hemicellulose content, dried or not, particle size). As the research in this thesis focussed on the interaction of specific proteins (CBMs) with their cellulosic substrate, particular attention was paid to the different cellulosic substrates that are commonly used to model enzymatic hydrolysis and protein/substrates interaction. One ongoing challenge was to select a model substrate is representative of the more complex lignocellulosic structure

1.6.3.2.1 Mechanical pulping

In mechanical pulping processes, fibers are separated from one another by mechanical energy applied to the lignocellulose matrix. This causes the bonds between fibers to break gradually and leads to the release of fiber bundles, single fibers and fiber fragments. This physical treatment leads to an increase in surface area, which often correlates with accessibility. It has been reported that enzymatic treatment can improve the mechanical pulp properties for paper making. For

example, laccase preparations, an enzyme that modifies pulp lignin, was shown to increase fiber bonding to enhance other strength properties of handsheets (Wong et al., 2000). However, studies on the conversion of mechanical pulps to monomeric sugars generally require a step for lignin removal or modification (Caitriona et al., 1998). It has been shown that mechanical refining is a promising post-treatment to increase cellulose accessibility and improve the bioconversion of lignocellulosics, which can help decrease enzyme loading (Park et al., 2016).

1.6.3.2.2 Chemical (Kraft) pulping

In the chemical pulping processes, the fibers are freed from the cell wall matrix as the lignin is dissolved into the chemical solution at a high temperature. The two chemical pulping methods that predominate use sodium sulfate (Kraft) and sodium sulfite. Today, the Kraft process is the dominant chemical pulping process due to the pulp superior strength properties compared with the sulfite process. Kraft pulp can be bleached to increase their brightness and remove the residual lignin. In bleached Kraft pulp, the carbohydrates are free from the lignin but still conserve some pore structures and are therefore more accessible to enzymes (Mansfield et al., 1997).

1.6.3.2.3 Dissolving pulp

To make a dissolving pulp, the lignocellulosic substrate is hydrolysed with low acid to remove the hemicellulose, yielding a relatively pure cellulosic substrate. Past studies have shown that enzyme treatments can be used to decrease the viscosity and reduce the amount of chemical used in the next steps. Dissolving pulps have been used to produce nanofibrillated cellulose (NFC) with enzyme pretreatment facilitating disintegration of cellulosic wood fiber pulp into nanofibers by

enhancing mechanical shearing in an environmentally friendly process. The enzymatically modified MFC showed more favorable structures than nanofibers resulting from fibers subjected to hydrolysis by strong acid (Henriksson et al., 2007).

1.6.3.3 Drying

To obtain reductions in the cost of transport, lignocellulosic substrates are sometimes dried. However, drying often leads to a collapse of the pore structure and a consequential decrease in accessibility (Esteghlalian et al., 2001). Past studies have indicated that there are significant changes in the pore structure of fibers during drying, some of which are reversible by rehydration, and some of which are not (Welf and Venditti, 2005). When modeling enzymatic hydrolysis on a pure cellulosic substrate, it is important to take into consideration how the substrate may have been dried as it greatly influences enzymatic efficiency. For example, it is known that internal porosity depends on the size and distribution of interfibrillar spaces created by the association of lignin and hemicellulose with cellulose microfibrils. Larger pores, such as the cell lumen or the pit aperture, can more easily accommodate attack by enzymes (Welf and Venditti, 2005). In contrast, the cell-wall capillaries can seriously restrict accessibility to cellulose. Cell-wall capillaries include inter-microfibrillar spaces as well as the interfibrillar defects produced by the removal of hemicellulose and lignin during the pretreatment (e.g. delignification). The collapse or reduction of these internal pores due to drying has been shown to decrease the extent of synergism among enzyme components and reduce the hydrolytic potential of the enzymes.

1.7 Cellulose and cellulosic substrates: structures, properties and model substrates

The hydrogen bond network makes cellulose a relatively stable polymer and results in high axial stiffness. The cellulose fibrils are the main reinforcement material for trees, plants, some marine creatures (tunicates) and algae. To gain insights on the factors influencing the enzymatic hydrolysis, some researchers have carefully selected "model" substrates depending on their structures and properties. Therefore, in this next section some cellulose-based nanomaterials are described emphasising their unique characteristics (*e.g.* cellulose nanocrystals).

1.7.1 Cellulose allomorphs and crystal structures

Pure cellulose can be found or generated in different allomorphic forms (Moon et al., 2011). Several different crystalline structures of cellulose are known, corresponding to different arrangements of hydrogen bonds between and within chains. Natural cellulose is cellulose I, is produced by bacteria, algae and plants. Regenerated cellulose produced from mercerisation of cellulose I fibers is termed cellulose II. With various chemical treatments, it is possible to produce the structures cellulose III and cellulose IV. However, the work described here primarily focussed on the allomorphic forms of native cellulose (cellulose I). Cellulose chains in native cellulose crystallizes in two variations of parallel formation, called cellulose Iα and Iβ. The Iα/Iβ ratio depends on the origin of the cellulose. Cellulose Iβ is more common in plants whereas cellulose Iα is more dominant in organisms such as bacteria or algae. Cellulose Iα has a triclinic unit cell, which includes one chain whereas Iβ has a monoclinic unit cell that includes two parallel chains. The Iβ form is thermodynamically more stable than the Iα form (Moon et al., 2011). Highly organized regions of cellulose are called crystalline regions, whereas less organized regions are

called paracrystalline regions (previously amorphous). Crystalline cellulose microfibrils have two or three distinct faces, which expose different groups and can therefore engage in different Van der Waals interactions.

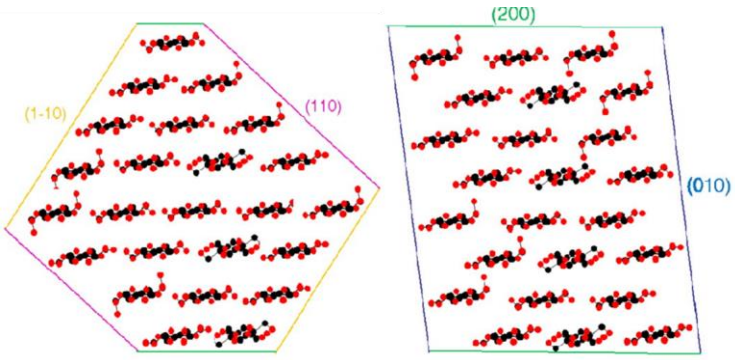


Figure 5: Schematic representation of the supposed arrangements of spruce cellulose microfibrils. (A). Diamond shape. 24 chains, overall dimensions 3.2 × 3.9 nm. Weighted-mean column lengths normal to lattice planes (002) 2.7 nm; (1–10), 3.2 nm, (110), 2.6 nm. (B). Rectangular shape. 24 chains, overall dimensions 3.2 × 3.1 nm. (*Adapted from Fernandes et al. 2011)*

In crystalline cellulose the three lattice planes correspond to Iα lattice planes $(110)_t$, $(010)_t$, and $(100)_t$ and Iβ lattice planes $(100)_m$, $(110)_m$, and $(1\text{-}10)_m$ respectively (**Figure 5**) (Nishiyama et al., 2003; Sugiyama et al., 1991). Several different description of these faces can be found in the literature and the properties of the faces also differ (Oehme et al., 2015). The microfibril crystal structures also vary depending on their source.

1.7.2 Cellulose crystallinity

While originally thought to play a major role in limiting hydrolysis, cellulose crystallinity seems to be less critical (Mansfield et al., 1999). In many studies where the crystallinity was suggested to be important, the substrates used were mechanically pretreated lignocellulosic materials (e.g.

ball milling) or chemically treated (phosphoric acid swollen), therefore the decrease in crystallinity was invariably accompanied by a decrease in other substrate characteristics such as particle size and an increase in available surface area (Yeh et al., 2010). However, even if the degree of crystallinity does not appear to significantly affect hydrolysis yields, it does appear to play a key role in the kinetics of hydrolysis. For example, the paracrystalline region has been shown to generally hydrolyse faster than the highly crystalline region (Beckham et al., 2011). Even at the microfibril level, models have shown that the different cellulose chain constituents of the microfibrils have different energy associated with the hydrogen bond network. Therefore, the thermodynamic barrier will be different when the enzyme is working on a corner cellulose chain compare to a middle chain (Beckham et al., 2011).

X-ray Diffraction (XRD), Fourier transform infrared spectroscopy (FTIR) and nuclear magnetic resonance (NMR ^{13}C solid state) are all often used to measure crystallinity. FTIR is often used as a method for comparison purposes, while the crystallinity index values are more accurately determined by XRD and NMR analyses (Park et al., 2010).

Considering that XRD and NMR are the most used methods to determine cellulose crystallinity, it is important to have a cautious interpretation of what is actually measured. Especially when this data is compared to hydrolysis progress and yields. First, these techniques measures the “bulk substrate” as an average while the enzymes only interacts with the substrates surface. Changes on the surface structures that generally represent 1-2% of the whole substrates (depending of the method used), are often impossible to track with XRD or NMR where these changes will fall into the region of experimental errors. Second, other substrate characteristics affect the NMR or XRD

results. For example, particles size, salt content, fibrillation and macroscopic organization all influence the crystallinity measurement. Therefore, developing a new complementary method such as CBMs adsorption may lead to better substrate characterization and provide new insights on the mechanisms of cellulose hydrolysis. It should be noted that these "classic" techniques require drying the substrate, which can lead to undesirable substrate changes, such as the collapse of pores and structures. More recently, the ratio of amorphous and crystalline surface, which is accessible to the enzymes, have been measured using CBM adsorption in aqueous environment (Gourlay et al., 2012).

1.7.3 Cellulosic substrates

Although some cellulosic substrates might have a similar chemical composition, often they have a distinct size, aspect ratio, morphology, crystallinity, crystal structure and properties. Thus, it is often very difficult to isolate one parameter of interest from other properties and it is important to take in account the whole picture while interpreting data.

1.7.3.1 Cotton

Cotton and derivatives are naturally occurring cellulose in almost pure form and have often been used as model substrates for enzymatic hydrolysis comparisons (*e.g.* Gourlay et al., 2012). These long fibers are usually used for the textile industry, they and cotton linters have often been used as model substrates. Cotton linter are short fibers containing 80% of holocellulose and are relatively homogenous and are consequently used for many cellulose-based products (i.e. cellulose acetate) (Morais et al., 2013).

1.7.3.2 Wood fiber, Pulp fiber

Wood fibers are the main materials used for the paper, textiles and biocomposites industries (Ramamoorthy et al., 2015). The fibers (i.e. bleached Kraft pulp, dissolving pulp) are microns in diameter, millimetres in length and contain a high percent age of cellulose. As these wood fibers conserved their cellular organization (e.g. fiber, macrofibril, microfibril), pulp and paper techniques have been used to "tune" these substrates to match desired characteristics (e.g. lignin content, fibrillation) (Liu et al., 2016). Wood fibers represent a credible, realistic substrate for a "modern" biorefinery (Ragauskas et al., 2014).

1.7.3.3 Bacterial microcrystalline cellulose (BMCC)

Bacterial microcrystalline cellulose (BMCC) is secreted by various bacteria. The purified microfibrils are microns in length, have a large aspect ratio (e.g. 50) with a morphology depending on the specific bacteria and culturing conditions (Esa et al., 2014). BMCC hasbeen used as a model substrate for enzymatic hydrolysis due to their purity and homogeneity (Luterbacher et al., 2013b).

1.7.3.4 Algal cellulose

Algal cellulose microfibrils are extracted from the cell wall of various algae by acid hydrolysis and mechanical refining. The resulting microfibrils are microns in length, have a large aspect ratio (greater than 40) with a morphology depending on the source of the algae (e.g. Valonia and Micrasterias) (Lehtiö et al., 2003). They have proven to be ideal substrates for studying processive enzymes due to their long microfibrils and are easily imaged using different microscopy techniques (Henrissat et al., 1985).

1.7.3.5 Microcrystalline cellulose

Microcrystalline cellulose (MCC) is a commercially available material (one brand name is Avicel®) used for applications in the pharmaceutical and food industries, and is prepared by acid hydrolysis of wood fiber, neutralized with alkali, and spray-dried (Trache et al., 2016). The particles are porous. Often considered mostly crystalline, MCC also contains accessible paracrystalline region (Gao et al., 2014).

1.7.3.6 Micro- and nanofibrillated cellulose (MFC and NFC)

Microfibrillated cellulose (MFC) can be produced via mechanical refining of delignified Kraft fibers. The MFC particles are considered to contain multiple microfibrils. They have a high aspect ratio (10–100 nm wide, 0.5–10 mm in length), and contain paracrystalline and crystalline regions (Lavoine et al., 2012).

NFC are reminiscent of the microfibrils. They can be obtain chemically with tempo oxydation and sonication or with high energy homogeniser. They have a high aspect ratio (4–20 nm wide, 500–2000 nm in length) (Missoum et al., 2013). The differentiation of NFC from MFC is based particle diameters. However, in the literature, the MFC and NFC terminology are sometimes used interchangeably, which may lead to some confusion. In the work reported here, we refer to NFC fibers as having diameters of 5-10 nm. Bigger fiber are considered to be MFC.

1.7.3.7 Cellulose nanocrystals (CNCs)

Also named nanocrystalline cellulose or cellulose whiskers, CNCs have a high aspect ratio (3–5 nm wide, 50–500 nm in length), are 100% cellulose and highly crystalline (Klemm et al., 2018). CNCs have been suggested to be representative of the crystalline regions within the microfibrils of the wood and plant cellulose biosynthesis process. They consist of 36 cellulose chains arranged in Iβ crystal structure and have a square cross-section. However, alternative models have also been proposed (Yoo and Youngblood, 2016).

1.7.3.8 Phosphoric acid swollen cellulose (PASC)

Usually produced to mimic a mostly paracrystalline substrate. Usually 100 units long with a low crystallinity index and very high surface area 240 m^2. g^{-1}, it is commonly acknowledged that PASC hydrolysis is easy and fast, due to its high accessibility and disorganized structure. However, nanocrystalline structures have been identified within PASC (McLean, 2002).

1.8 Substrate characterization: cellulose accessibility to enzymes and how to measure it

Choosing the best pretreatment conditions to increase cellulose accessibility to enzymes is crucial to improving hydrolysis yields. However, measuring cellulose accessibility has been challenging for researchers and some methods often do not represent the real substrate morphology. In the following section the different techniques that have been used to try to relate cellulose accessibility are described. In the literature, cellulose accessibility often refers to its specific surface area (SSA) (Meng and Ragauskas, 2014). However, for deconstruction of lignocellulose substrates only the

carbohydrate surface needs to be measured. Because lignin often covers the cellulose macrofibrils, unspecific measurement can be inaccurate, and the data needs to be interpreted with caution.

1.8.1 Pulp and paper methods that relate to cellulose accessibility

Methods borrowed from the pulp and paper field, such as water retention value (WRV) and fiber quality analyser (FQA), are often used to assess the approximate behavior of the substrate surface area. These techniques are usually fast and easy but are to crude to effectively measure accessibility or predict hydrolysis behavior as they are heavily influenced by other factors such as charge groups (i.e. carboxylic groups).

1.8.1.1 Water retention value (WRV)

The water retention value (WVR) provides an indication of fibers' ability to take up water and swell. In some cases, the WRV has been used to measure accessibility with various degree of success (Aïssa et al., 2019; Arantes and Saddler, 2011). However, as discussed in **section 5.2**, this value provides a different physical phenomenon that cannot be directly correlated to enzyme accessibility to cellulose. The limitations of this technique includes the heavy dependence of the calculated WRVs on test conditions combined with its inaccuracy for analyzing highly swollen pulps due to the different factors involve in the hydrophilicity of a substrate. Additionally, the nature of the protocol make it a very crude method, which prevents the application of this technique for quantifying minor subtle changes in the degree of swelling.

1.8.1.2 Fiber quality analyzer (FQA)

The fiber quality analyzer is a technique that can be used to quantify macroscopic properties of lignocellulosic fibres. The fibre samples is suspended in water and passed through a flow cell, where images are captured and analyzed using image analysis software (Clarke et al., 2011). The length and width of the fibres passing through can then be determined. While this technique is simple and easy to use, it measures the particle size average, which often relate to the surface area. Although fiber quality analysis could potentially be used to detect macroscopic swelling or fragmentation of the substrate, it is unable to give information on internal surface area and microstructures. For example, fiber modifications, such as pitting and roughening of the fibre surface could greatly enhance accessibility, without necessarily altering the macroscopic fibre dimensions.

1.8.2 Techniques to measure cellulose accessibility

The surface area of fibers can be divided into interior surface area (e.g. crack, pores), and exterior surface area (e.g. particle size and fibrillation). The various methods commonly used to asses these parameters are summarised below.

1.8.2.1 Nitrogen adsorption

Nitrogen adsorption is one of the most common methods used to measure the total surface area of cellulose substrates (Wiman et al., 2012). This method requires drying, degassing and cooling of the samples in the presence of nitrogen gas, allowing nitrogen gas to condense on the surfaces and within the pores. The Brunauer–Emmett–Teller (BET) model is used to correlate the physical

adsorption of gas molecules on solid surface. However, the drying step is problematic, because, as shown by Esteghlalian et al. (Esteghlalian et al., 2001), it can modify the structure of the fibers and thus change the substrate. It was also shown that more than 88% of the nitrogen accessible to cellulose is not accessible to cellulases as the pores are too small for the proteins to diffuse into (Gao et al., 2014). Therefore, nitrogen adsorption is not representative and should not be used to measure enzyme accessibility to cellulose.

1.8.2.2 Mercury porosimetry

Similar to nitrogen adsorption, mercury is added to dried and degassed biomass samples and an increase in pressure is used to force mercury into the cellulose pores. The volume of mercury entering the pore is measured as the pressure increased, indicating the cumulative volume of all available pores (Westermarck, 2000). As with the nitrogen adsorption, mercury porosimetry requires a drying step, and therefore introduces a bias. Considering the high toxicity of mercury, this technique is not recommended.

1.8.2.3 Solute exclusion

This technique measures accessibility of probe molecules of different sizes to the substrate pores. The method uses a known concentration of a solute molecule solution that is added to the substrate. The probe molecule solution is diluted by water that is contained in the substrate (Lin et al., 1987). The water present in the pores that was not accessible to the probe molecules does not contribute to the dilution. As a result, the substrate pore size and volume distribution can be determined using the concentration of a set of different solute solutions with various molecule sizes. This method is

known to be very tedious, unspecific to cellulose and does not measure the external surface area (Wang et al., 2012).

1.8.2.4 Simons' stain

The Simons' stain method involves the use of a blue and an orange dye, which are differentially adsorbed by the lignocellulosic substrate. The blue dye has a low molecular weight, 100Da, and binds to every constituent regardless of its structure. The orange dye has a high molecular weight, and a higher affinity towards the cellulose structure. The Simons' stain methods was further refined by Chandra et al. (Chandra et al., 2008) to improve the throughput and sensitivity of Simons' stain method using the Langmuir isotherm with the orange dye. The modifications to the original technique greatly reduced he incubation time required for dye impregnation. Based on this method, structural changes in the biomass, induced by pretreatment, were successfully measured (Chandra and Saddler, 2012).

1.8.2.5 Nuclear magnetic resonance (NMR)

This technique has been used to determine pore size distribution, taking advantage of the formation of small crystals within pores. The internal crystals melt at a lower temperature than the bulk liquid. This phenomenon is known as melting point depression, which can be related to the pore size through the Gibbs–Thompson equation (Östlund et al., 2013). NMR-relaxation experiments can provide information on the molecular mobility within a porous system. However, the pore size determination range is limited by the temperature control and NMR techniques require expensive and complex set up and long experiment times (Meng et al., 2013).

In general, the methods used to measure cellulose surface area have been adapted from other fields (e.g. material science), where they are used to measure the specific surface area of different materials. Thus, most of these methods are not specific to cellulose, which is problematic for the measurement of carbohydrate surface in a multi-component substrate. Therefore, this was one of the major reasons that motivated the development of the novel, specific method (differential CBM binding) used to measure cellulose accessibility to enzymes.

Table 2: Summary of analytic method to measure cellulose accessibility

Techniques	Measures	Advantage and disadvantage	ref
➢ **Water retention value (WRV)**	• Substrates interaction with water	✓ Fast and easy set up × Crude	(Gu et al., 2018; Hoeger et al., 2013; Nakagaito and Yano, 2004)
➢ **Fiber quality analyzer (FQA)**	• Substrates average sizes	✓ Fast and easy set up × Crude	(Clarke et al., 2011)
➢ **Nitrogen adsorption**	• Surface Area accessible to a nitrogen molecule	✓ Allows pore size analysis ✓ Versatile × Measure dry sample × Overestimate the accessible surface area for enzyme	(Westermarck, 2000)
➢ **Mercury porosimetry**	• Surface Area accessible to Mercury	✓ Allows some pore analysis × Toxic × Tedious	(Westermarck, 2000)
➢ **Solute exclusion**	• Internal surface accessible to the probe used	✓ Measure in wet stage × Tedious × Only internal surface × Affected by pore size and osmotic pressure	(Lin et al., 1987; Wang et al., 2012)
➢ **Simons' stain**	• The ratio of dye adsorbed can be used to calculate accessible surface area	✓ Measure in wet stage × Not fully quantitative × Depends on dye quality × Depends on substrate composition	(Esteghlalian et al., 2001; Yu et al.,)
➢ **Protein adsorption (CBMs)**	• Accessible surface area to proteins	✓ Only method specific to cellulose ✓ Measure in wet stage ✓ Fully quantitative × May required lignin blocking	(Kawakubo et al., 2010; Wang et al., 2012)

Chapter 2: Objectives, work hypothesis, background

The thesis work tackles four milestones: 1) CBMs production and purification, 2) the use of CBMs to measure accessibility, 3) following the structure changes of cellulose surfaces during the initial stage of the hydrolysis using fluorescent CBMs and 4) functionalization of cellulose surfaces with bioconjugated CBMs.

2.1 CBM production, purification and modification

Typically, the production of CBMs have been a major bottleneck in their utilization. Thus, we focused on the production of heterologous protein from modified *E. coli*. A successful protein production needs to address different problems such as expression lost. Higher production efficiencies and, and consequently lower costs of the final product are needed for obtaining a commercially viable process. Common problems in recombinant protein production were addressed (*e.g.* IPTG concentration) and strategies for their solution were discussed. To continue our work, we chose the CBMs that were produced in large quantities and that were easily conjugated (*i.e.* CBM2a and CBM17).

2.2 CBM adsorption to measure accessibility and predict enzymatic hydrolysis

Following the production of CBMs, a reproducible and accurate methodology to quantify cellulose accessibility, a key parameter in enzymatic degradation of lignocellulose substrates, was developed. The standardization of depletion assays was conducted with different model substrates.

Thus, the method was compared to other commonly used method used to determined cellulose accessibility. The method was applied on a case study where, a type A (crystalline cellulose) and a type B (paracrystalline) CBMs were used in parallel with microscopy, fiber analysis and water retention values to determine if the observed and anticipated changes in differentially prepared microfibrillated cellulose (MFC) substrates were similar.

2.3 Visualization of cellulose surface with fluorescent CBMs using CLSM

The quantification of accessible cellulose and cellulose substructures was further refined using a fluorescent marker on the CBM probes. Fluorescence-tagged CBMs were used to differentiate the cellulose substructures during enzyme-mediated hydrolysis. Differences in CBM adsorption were elucidated using confocal laser scanning microscopy (CLSM) and the structural changes occurring during enzyme-mediated deconstruction quantified via the relative fluorescence intensities of the respective probes. Quantitative image analysis, supported by ^{13}C NMR, SEM imaging, and fiber length distribution analysis, showed that enzymatic degradation predominates at these zones during the initial phase of the reaction, resulting in rapid fiber fragmentation and an increase in cellulose surface crystallinity. The use of fluorescence tagged-CBMs with specific recognition sites provided a quantitative way to elucidate cellulose structures and their impact on enzyme accessibility. This in turn provided novel insights into the mechanisms involved in cellulose deconstruction.

2.4 The use of CBMs to functionalize cellulose surface

Chemical functionalization has been shown to improve the overall utilization of cellulosic polymers. Although a wide variety of reactions for cellulose modification have been studied (Cunha and Gandini, 2010) common use of solvents increase cost and environmental challenges. In this part of the thesis, CBM2a (crystalline cellulose) was functionalized with NHS-alkyne making use of the terminal amine. Following this bioconjugation, a Click reaction with polyethylene glycol (PEG) was conducted to modify CNC surfaces. This provided a strong and non-covalent modification of cellulose surfaces. It had the added advantage of being a “one-pot reaction” in aqueous media. The CBM-PEG modification of cellulose surfaces was shown to increase CNC dispersion after drying, thus, its improved suspension stability. It was apparent that polysaccharide–protein hybrid and self- assembled nanoparticles could be effectively produced with CBM providing a versatile vector for cellulose functionalization.

Chapter 3: Material and Methods (Part 1) – Substrates characterization, microscopy and bioconjugation.

3.1 Substrates used and their preparation (section 5.2)

Avicel was purchased from Sigma-Aldrich (St Louis, USA). Phosphoric acid swollen cellulose (PASC) was produced from Avicel, as described previously (Gourlay et al., 2012). In brief, 4 g Avicel was suspended in 100 mL of phosphoric acid. After stirring (1 h at 4 °C), the reaction was quenched with 2 L of cold water. After stirring at 1 h at 4 °C, the extracted swollen cellulose was collected by filtration. The swollen cellulose was washed four times with ultrapure water, and two times with 1 % $NaHCO_3$ for neutralization, and then 3 more times with ultrapure water.

The MFC substrates were produced by treating bleached kraft pulp with different refining energies using an Aikawa single disc 14" refiner (Advanced Fiber Technologies Inc., Petaluma, USA). Refining was performed with a of 3.5 % (w/v). During refining power levels of 20 kW for about 6 h total, the substrate consistency may drop to about 3 % (w/v) because of the added cooling water. Each pass through the refiner added a refining energy of about 25 $kWh.ton^{-1}$. The MFC0 sample was collected after 10 passes, corresponding to a refining energy of 250 $kWh\ ton^{-1}$. The MFC2, MFC4 and MFC6 substrates were additionally treated 10, 30, and 50 times, resulting in refining energies of 500 $kWh.ton^{-1}$, 1000 $kWh.ton^{-1}$ and 1500 $kWh.ton^{-1}$, respectively.

3.2 Substrates and chemicals

All chemicals were purchased from Thermofisher Scientific (Massachusetts, USA) and Sigma-Aldrich (st Louis, USA) (Unless mentioned differently), enzymes were provided by Novozymes (Denmark). Soaked northern softwood Kraft pulp (NBSK) from handsheets and never-dried bleached Kraft pulp were used as substrates. NBSK handsheets were soaked in water at 1% consistency and the fibers redispersed non-disruptively with a standard disintegrator (Robert Mitchell Company Ltd, QC, Canada) for 15 min at 3000 rpm (≤ 45000 revolutions, *TAPPI T205 standard*). Filtration and washes with 2 L of water for 30 min underwent under gentle stirring. Afterwards the fibers were again filtered and washed with 2 L of water, and then manually pressed out up to 40% dry matter content. Conductivity measurements were performed in the wash fractions to confirm the absence of soluble salts in the NBSK. The never-dried softwood Kraft pulp was prepared as described previously (**section 1.7**). The chemical composition of two substrates was determined following the protocol of the National Renewable Energy Laboratories (NREL).

3.3 Cellulose nanocrystals

CNC suspensions were prepared as follows: cellulose powder (40 g of dried Whatman cotton powder) was added to 64% w/w sulfuric acid (700 mL) and allowed to react with stirring at 45 °C for 45 min. To quench the reaction, the reaction mixture was diluted with deionized water. To remove the excess of acid the CNC were placed in a dialysis membrane and stirred for 48 hours with regular change of the external media (deionized water) (Karaaslan et al., 2013). The

suspension was sonicated (Watt) repeatedly (5 cycles of 7 min at 15% output) to create cellulose crystals of colloidal dimensions. During the ultrasonic treatment, the suspension was cooled in an ice water bath to avoid overheating and the sonication process was performed in a plastic beaker to avoid possible ion release from the glass containers. The carboxylic groups were calculated via titration following the procedure of Kloser et al. 2010. The width and length of CNC were calculated using AFM based on the selection of 50 CNC and averaging their dimensions.

3.4 Compositional analysis

The chemical composition of the microfibrillated cellulose were determined following the two-step sulphuric acid hydrolysis method. Glucose, xylose, galactose, arabinose, and mannose were quantified by a high-performance liquid chromatography (HPLC; ICS-3000, Dionex, Sunnyvale, USA) equipped with a CarboPac PA1 anion exchange column (Dionex). Nanopure water was used as isocratic eluent at a flow rate of 1 $mL.min^{-1}$. All analyses were performed in triplicates.

3.5 Measurement of the aspect ratio

The aspect ratio represents the ratio of fiber width versus fiber length. The aspect ratio was determined using an Optest Hi-Resolution fiber quality analyzer (LDA02-series, OpTest Equipment Inc., Hawkesbury, Ontario, Canada). About 5000 fibers were collected to calculate the average length over the range of 0.07 and 10.0 mm. For fibers between 0.5 and 10 mm length, the average width was calculated over the range of 7 to 60 µm.

3.6 Fiber length analysis

The fiber length distributions of the hydrolysis samples were measured by a HiRes Fiber Quality Analyzer (LDA02-series, OpTest Equipment, ON, Canada). The sample size was 10000 fibers and the data acquisition range was bounded by the lower detection limit, 0.07 mm, and 2.5 mm. Observations outside the data range was censored during data analysis. The arithmetic mean, median, and a non-parametric skewness (Pearson's second skewness coefficient) of the fiber length distributions were derived from the resulting frequency tables (histograms). Data processing was performed in MATLAB R2017a (MathWorks Inc., MA, USA), using the statistics and machine learning toolbox.

3.7 Accessibility techniques

3.7.1 Measurement of the Water Retention Value (WRV)

The Water Retention Value (WRV) represents the amount of water retained by fibers after centrifugation relative to the dry weight of the substrate. It was determined and calculated using a slightly modified version of TAPPI UM 256 (Water retention value (WRV), Useful Method UM 256 (2015)). In brief, 0.1 g of MFC was incubated in water for 2 h, then centrifuged at 900 g for 30 min and finally oven-dried at 105 °C for 3 days to reach constant weight.

3.7.2 Simons' Stain

The Simons' Stain method is derived from a staining technique used in the pulp and paper industry to examine changes in the physical structure of pulp fibres. It was adapted to try to estimate the overall surface area of cellulosic materials. It was performed according to the modified procedure of Chandra et al. Pontamine fast orange (direct orange) and Pontamine fast blue (direct blue) dyes were used (Pylam Products Co. Inc., Garden City, NY, USA). Fractionation of DO was performed according to Esteghlalian et al. 2001 (Esteghlalian et al., 2001).

3.7.3 NMR analysis

Solid-state NMR performed on a Bruker 500-MHz instrument (Milton, Ontario, Canada) and ^{13}C CPMAS experiments were performed on a 2.5 mm HCN MAS probe, using RAMP-CP. Spinal-64 was used during the decoupling. All spectra were acquired with 10240 scans, a recycle delay of 4 s, a 90° 1H pulse of 2.75 μs and a cross-polarization time of 3 ms. The spectra were processed with a line broadening of 50 Hz and referenced to external adamantane.

3.8 Fourier Transform Infrared Spectroscopy (FT-IR)

Qualitative analyses of CNC and modified CNC were conducted by FT-IR transmission spectra deposing 25 µl on a Zirconium support. (FTIR Spectrum 100, PerkinElmer, Waltham, MA, USA) All spectra were collected per sample vary from 450 to 4000 cm^{-1} with a resolution of 4 cm^{-1} and 10 scans.

3.9 Enzymatic hydrolysis

3.9.1 Analysis of hydrolyzability

The substrates MFC0, MFC2, MFC4, and MFC6 were hydrolyzed using the commercial enzyme mixture Cellic Ctec3 (novozymes). The substrate and enzyme loading used was 1 % (w/v) dry mass and 5 Filter paper units per g^{-1} of substrate dry mass, respectively. Reactions were performed in sodium acetate buffer (20 mM, pH 4.8) in 250 mL screw cap Erlenmeyer flasks. The reaction was weighted at 50 g. The substrate suspensions were autoclaved, and the enzyme solution was added aseptically. Incubation was at 50 °C, 200 rpm for 24 h in an orbital incubator shaker (IST-4075, Thermo Fisher Scientific, Hampton, USA). During hydrolysis samples were taken regularly. Immediate sample work-up included boiling (100 °C, 10 min) and centrifugation (17,000 g, 5 min, accuSpin Micro 17, Thermo Fisher). The supernatant was stored at -20 °C prior to quantification by HPLC, as described above. Hydrolysis was evaluated by the amount of glucose released (in $g.L^{-1}$).

3.9.2 Enzymatic hydrolysis of NBSK and never-dried pulps

The hydrolysis reactions of the substrates were performed in 100 mL screw cap shake flasks with 25 g total reaction weight. Incubation was at 50 °C and 150 rpm in an incubator shaker (25 gIST-4075, GMI Inc, MN, USA). Never-dried pulp was hydrolyzed in 2 mL screw cap tubes with 1.5 g total reaction weight at 50 °C in a thermo-block. The reactions were stirred hourly by vortexing. All reactions were performed in 20 mM sodium-acetate buffer (pH 4.8) at a substrate loading of 1

% (w/w). The commercial enzyme mixture used was Ctec3 (total cellulase and β-glucosidase activity of 205 $FPU.mL^{-1}$ and 6400 $CBU.mL^{-1}$, respectively). The enzyme loading of all experiments was 10 $FPU.g^{-1}$ dry mass. To stop the reaction, the mixtures were brought to 100 ℃ for 10 min, and then stored at -20 ℃ for future analysis.

3.9.3 Analysis of enzymatic activities and glucose concentrations and determination of hydrolysis yields

Total cellulase activities (FPU) were measured as described by Ghose et al. β-glucosidase activities were determined with p-nitro-phenyl β-D-1.4-glucopyranoside as substrate, following a previously established protocol. Released glucose was analyzed using a YSI Biochemistry Analyzer (2700 Select, Fisher Scientific, MA, USA).

3.10 Microscopy

3.10.1 SEM imaging of the MFCs

The oven dried MFC substrates were mounted on aluminum stubs using double-sided tape. After sputter-coating with 10 nm Au/Pd (Gold/Palladium, 80:20 mix), the MFCs were imaged using scanning electron microscopy (SEM, Hitachi S-2600 VP-SEM, Tokyo, Japan).

3.10.2 Sample preparation for confocal laser scanning microscopy (CLSM) imaging

In samples taken from enzymatic hydrolyses, enzymes bound to the cellulose fibers were removed prior to CBM binding. The fibers were washed three times, by centrifuging (5000 rpm, 10 min), and subsequently re-suspended in deionized water. After the last run, the fibers were taken up in sodium dodecyl sulfate solution (SDS, 1 % w/v) and incubated for 15 min at 85 ºC. After cooling, the substrate samples were washed two times with absolute ethanol and one time with water. The washed samples were then suspended in PBS buffer (8 $g.L^{-1}$ NaCl, 0.2 $g.L^{-1}$ KCl, 1.44 $g.L^{-1}$ Na_2HPO_4, 0.24 $g.L^{-1}$ KH_2PO_4) and stored at 4 ℃ for CLSM analysis. Shortly before imaging, the fluorescence-tagged CBMs were mixed with the fibers, resulting in a protein loading of approximately 5 µg of each CBM per g dry mass substrate. After a short incubation period (~15 min), excess protein in the supernatant was removed by centrifuging the fibers and resuspending them in PBS buffer. A small amount of sample was then mounted on objective glass, covered with a cover slip, and sealed with paraffin wax. Sample preparations and sample storage before CLSM imaging were conducted in the dark to prevent fluorescence quenching.

3.10.3 Confocal laser scanning microscopy (CLSM) imaging

Confocal laser scanning microscopy (CSLM) imaging was performed using an Olympus FV1000 (Olympus, Japan), a 10x (NA 0.3-0.4) air objective and a Zeiss 710 LSM (Carl Zeiss AG, Germany), using a 10x (NA 0.3-0.4) air objective and a 40x (NA 1.2) water objective. Following the specifications of the dyes, CBM17-FITC (green) was excited at 570 nm and emissions detected at 590-650 nm. The excitation and emission of CBM2a-RRed X signal (red) was at 490 nm and

510-550 nm, respectively. Images were acquired in 5-6 µm (10x) and 0.4-0.5 µm (40x) thick optical sections. Bleed-through of fluorescence emission was tested and can be neglected for the presented image quantification method.

3.10.4 Quantitative Confocal laser scanning microscopy (CLSM) image analysis

Quantitative image analysis of the acquired CLSM images was performed to assess changes in concentrations of fluorophores and the distribution of structural features in the specimens with increasing cellulose conversion. High throughput could be achieved with a high degree of automation of the image analysis. The image analysis was performed in a Python environment (Build 2.7.10, Python Software Foundation) using open source image processing toolboxes. A detailed description is given in the Supplementary Information. In brief, the red and green fluorescence layers of the acquired CLSM images were imported to the processing environment, the background values of the layers were subtracted from the fluorescence intensity data, and the data was scaled by the laser intensities. Point features, i.e. negative and saturated pixels, in the layers were removed from the corrected intensity data sets. Fibers were identified by semi-automated detection of the region of interest (ROI). In brief, a median filter with a 7×7 aperture was applied to reduce spatial noise in each layer. The filtered layers were segmented based on contrasting intensities between fibers and background by unsupervised thresholding using Otsu's method for the red layer and the triangle method for the green layer. The triangle method was used for increased robustness for the latter because of the weaker bimodality of the intensity histogram. The ROI was defined as the union of the identified fiber subset of each layer. The ROI was used

to mask the corrected fluorescence intensity layers and the overall red to green ratio (R/G-ratio) was calculated as the arithmetic mean of the spatially distributed ratios of the masked intensity layers.

3.10.5 Analysis of pulp/fibers by Scanning Electron microscopy (SEM)

Hydrated samples were washed and freeze-dried to prevent dehydration artifacts prior to imaging. Samples were mounted on aluminum stubs (VWR) using double-stick carbon tape and sputter coated with 7 nm of iridium. Scanning Electron microscopy (SEM) imaging was performed using a FEI Quanta 400 FEG with accelerating voltages ranging from of 10-15 keV.

3.10.6 Atomic force microscopy (AFM)

Images of CNC samples were captured using Multimode AFM Nanoscope-VIII from Veeco Instruments (Santa Barbara, CA, USA) with PeakForce Tapping mode. A freshly cleaved mica was chemically modified with a cationic polymer (3-Aminopropyl)triethoxysilane (APTES) according to the previously described method (Karaaslan et al., 2013). Briefly, a droplet of 0.1% w/v APTES aqueous solution was placed on a cleaved mica surface for 1 min, thoroughly rinsed with nanopure water and dried with pressurized air. Small aliquots (30 µL) of 1% w/v CNC suspensions was dropped on the modified mica surface and rinse with 1 ml of nanopure water and dried under vacuum in an oven operating at 40 °C. AFM images were captured with scan rate of 0.7 Hz and at a scan size of 3 * 3 µm, 1*1 µm and 300 * 300nm using RTESPA-150 cantilever probes with a nominal spring constant of 2.45 $N.m^{-1}$. The AFM images including height, peak

force error, adhesion and deformation channels were analyzed using NanoScope Software 8.10 (Veeco, Santa Barbara, CA, USA). Peak force tapping mode was calibrated using sapphire standards.

3.11 Click reaction chemistry

CBM-alkyne were mixed with $CuSO_4$ (0.05 mol%) and THPTA, then sodium ascorbate (50 µL, 100 mM) was added the completion of the reaction was monitored on polyacrylamide gel, XT-MES buffer (biorad). conjugated CBM2a-PEG was purified using a desalting column HiPrep 26/10 eluted with 50mM phosphate buffer (pH 7).

3.12 Other Chemicals

Propargyl-N-hydroxysuccinimidyl ester (NHS-propargyl), Methoxypolyethylene glycol azide (PEG-azide), Tris(benzyltriazolylmethyl)amine THPTA, copper sulfate ($CuSO_4$), sodium ascorbate were purchased from (Millipore Sigma, St Louis) and used as it is.

Chapter 4: Materials and Methods (Part 2) - CBM production and utilization

The aim of this thesis section was to design a robust production process that can be easily reproduced in different laboratories, which will allow the use of CBMs across the bioenergy and biomaterial fields. *Escherichia coli* (*E. coli*) was the organism of choice for the expression of a wide variety of recombinant protein for diagnostic and industrial applications. First, shake-flask cultures were used to screen parameters to select optimal growth condition. The CBM production was then scaled up using a fed-batch approach. The CBM2a and CBM17 from *Cellulomonas fimi* and *Clostridium cellulovorans*, respectively, were utilized. The CBM2a and CBM17 genes were on pTug and pET expression vectors, respectively, and transformed into the host strain Escherichia coli BL21. The strains were denoted BL21-CBM2a and BL21-CBM17. The *E. coli* strains JM101, BL21 and R1360 have been used to produce wild-type CBMs and histidine-tagged CBM2a. The kanamycin resistant *E. coli* strains were graciously provided by Prof. Withers (UBC, Faculty of Science).

4.1 Seed and starter cultures

All chemicals were from Fisher Scientific. The BL21-CBM2a and BL21-CBM17 were stored in 20 % (v/v) glycerol stocks at -80 ºC. Prior to cultivations, 50 µL of stock was streaked on LB-Kan plates (10 $g.L^{-1}$ tryptone, 10 $g.L^{-1}$ NaCl, 5 $g.L^{-1}$ yeast extract, 15 $g.L^{-1}$ agar, 50 $\mu g.mL^{-1}$ kanamycin) and incubated at 37 ºC overnight. One colony was then used to inoculate 25 mL TB-Kan medium (12 $g.L^{-1}$ tryptone, 24 $g.L^{-1}$ yeast extract, 2.3 $g.L^{-1}$ KH_2PO_4, 12.5 $g.L^{-1}$ K_2HPO_4, 50 $mg.L^{-1}$ Kan) with 5 $g.L^{-1}$ glycerol (TB-Kan) in 250 mL baffled shaken flasks. The flasks were incubated

overnight at 37 ºC and 150 rpm in an orbital shaker. Cells were then transferred to 50 mL of TB-Kan medium in 250 mL baffled shaken flask. The starting OD_{600} was 0.05. Incubation was at 37 ºC and 150 rpm until an OD_{600} of 0.5 was reached.

4.2 Shaken flask cultivation

Shake flask cultures were run in batch, all components were already added at the start of the cultivation. There was no monitoring and control of pH or the level of dissolved oxygen. The CBM production in shaken flasks was optimized following a standard protocol of one single colony selected form a petri dish used to inoculate 25 mL of terrific broth media in a 100mL shaken flask (French et al., 1996). Terrific broth (TB), an enriched medium, helped the recombinant strains of *E. coli* maintain an extended growth phase. This media was developed especially to increase cell yields (Krause et al., 2010). Glycerol was added as a carbon source (4.0 g/L) as it is known to reduce the accumulation of overflow and fermentative metabolites (Lowry et al., 1971). Potassium phosphate buffer (pH 7) was added to prevent acidification of the medium by excessive acetic acid formation. In order to prevent the *E. coli* from losing the plasmid and to protect *E. coli* cultivation from contaminant, a constant concentration of the antibiotic Kanamycin (50µg/mL) was maintained during each step of the production.

Table 3. Media composition for 1 L of terrific broth, final pH 7.2±0.2 at 25ºC

Component	Amount (g)
Tryptone	12.0 g
Glycerol	4.0 g
Yeast Extract	24.0 g
Potassium Phosphate, Dibasic	9.4 g
Potassium Phosphate, Monobasic	2.2 g

From the intermediate cultivation, a 10 mL cell suspension was then added to 1 L baffle flask containing 300 mL of TB media. It was found that more volume led to a decrease in protein production, which was likely due to a lack of oxygen in the media (Losen et al., 2004). Once the optical density at 600nm (OD_{600}) reached its desired value, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce the CBM production. Previous work demonstrate that the time of induction can be crucial to optimize heterologous protein expression (Hasenwinkle et al., 1997).

4.3 Fed-batch approach to CBM production

A primary goal of fermentation optimization was the cost-effective production of desired products. Biotechnology companies have been producing protein through high cell densities culture (HCDC) to improve their productivity (Johnston et al., 2002). It also provided other advantages such as a reduce culture volume, enhanced downstream processing, reduced wastewater, lower production

costs and reduced investment in the needed equipment (Lee, 1996). Fed-batch culture uses inocula growing at the maximum specific growth rate that can be sustained using the nutrients initially present in the fermenters. Heterologous protein production in fed-batch processes should provide high cell density, high protein production and good protein quality (Lee, 1996). In practice, it is often more complicated to reach these goals due to several challenges such as heat and mass transfer issues, which affect the oxygen level and pH control. This typically needs to use low induction cell densities to impact protein yields (Krause et al., 2010). In this section, the various challenges of HCDC were investigated (e.g. oxygen level, time of induction) to establish an efficient CBM production. This was a key requirement if the thesis project was to succeed as efficient production will allow CBMs to be used for larger scale applications or as routine analytical tools.

As mentioned earlier, CBM17, CBM2a and CBM2aH6 have been successfully expressed in *E. coli* host BL21 from pET and pTug vectors, respectively (Boraston et al., 2000; McLean et al., 2000). Using the established strains and previous shake-flask screening, a fed-batch approach with predetermined exponential feeding was designed. In contrast to previously published work (Hasenwinkle et al., 1997), this strategy eliminated the need for complex feed-back control. The cultivations were run in two steps, a batch phase that allowed unregulated cell growth (μ_{free}) followed by a fed-batch phase, in which the growth rate was controlled by carbon-source limitation at approximately 0.12 h^{-1} (μ_{set}, see materials and methods).

Bioreactor cultivation: For bioreactor cultivations, Applikon bioreactor system was used equipped with an ADI-1025 BioConsole and ADI-1010 controller (2 L, Applikon Biotechnology,

Delft, Netherlands). The cultivation was run in two phases, a batch followed by a fed-batch phase. For the batch phase, 1 L TB-Kan medium with 20 $g.L^{-1}$ glycerol (TB-Kan) was prepared and inoculated with cells from the starter culture to an OD_{600} of 0.05. Initially the cultivation conditions were 37 ºC, pH 7, and 30 % dissolved oxygen (pO2). The regulation was performed with a cascaded controller on agitation and aeration with pressurized air or pure oxygen. At OD_{600} of 0.7, protein expression was induced with 0.3 mM IPTG (Corning, New York, USA). At the same time the temperature was reduced to 30 ºC and the pO2 increased to 60 %. During batch phase, samples were taken regularly to measure the growth rate as determined by the increase in OD_{600}. The batch phase was continued until the glycerol was depleted. Growth was calculated assuming a specific glycerol update rate (qGlycerol) of 2.67 $g.g^{-1}$ cell dry weight.h^{-1} and a cell dry weight to OD600 ratio of 0.52. After modelling, the calculated feeding trajectory was used to program the peristaltic feed pump. A 10-fold concentrated TB-Kanamycin medium containing 100 $g.L^{-1}$ glycerol was used as feed. The fed-batch phase was run at 30 ºC, pH 7, and 60 % pO2. To keep the latter stable throughout the fed-batch phase, the air inflow was fortified with pure O_2 (Aïssa et al., 2019).

4.3.1 Oxygen level

The bioreactor was adjusted to sustain a consistent oxygen level during the entirety of the cultivation. Oxygen is an important nutrient that is used by microorganisms for growth, maintenance and metabolite production, meaning that the scarcity of oxygen will affect the process performance. In initial tests, the oxygen levels in the fed-batch cultures were insufficient and challenging to control. Despite increasing agitation and with a gas stream of pure oxygen, the

oxygen level stayed too low to provide optimal cell growth. To solve this issue, the bubble diameters and geometric parameters of the bioreactor were optimized. In mechanical agitated bioreactors, the stirrer is the main gas-dispersing tool and stirrer speed and design both had a pronounced effect on mass transfer (Garcia-Ochoa and Gomez, 2009). Two flat blade paddles were installed instead of a single curved blade paddle to increase the oxygen level in the bioreactor (Garcia-Ochoa and Gomez, 2009), which was especially problematic at the latter stages of the cultivation. Because the interfacial area is key for efficient oxygen transfer, the system was improved to create smaller bubbles, increase their numbers and their residence time. Thus, three fixed baffle paddles were installed to maintain a turbulent flow and a larger sparger was added to generate more and smaller bubbles. These improvements successfully provided the cells with enough oxygen for their growth.

4.3.2 Acetate inhibition

As mention previously, *E. coli* produces acetate in oxygen-limiting conditions or in the presence of an excess of glucose. Concentration of acetate over 5 $g.L^{-1}$ reduces growth rate and biomass yield (Luli and Strohl, 1990). The expression of the recombinant protein was also greatly affected by the acetate concentration, even though it is unclear what are the exact mechanisms involved (Lee, 1996). The designed fed-batch process allowed the recombinant *E. coli* to be continuously fed with a substrate solution so that one substrate was growth rate limiting. The latter approach has been shown to effectively prevent acetate formation caused by the overflow-metabolism and favours protein production (Korz et al., 1995). To further slowdown metabolic rates and thus,

decrease the likelihood of the onset of the bacterial Crabtree effect, glycerol was used as carbon source (Korz et al., 1995; Luli and Strohl, 1990; Yee and Blanch, 1992). In order to control the pH that may decrease due to residual acetate formation, the system was equipped by ammonium chloride control line that increased the pH in the media in case of acidification. However, the designed glycerol feed efficiently prevents this phenomenon. Therefore, the system only required one-pump to be operational.

4.4 Protein stability and IPTG concentration

A substrate-limited fed-batch cultivation was used, comprising of a batch process continuously fed with a substrate solution so that one substrate component is growth rate limiting. In this way, the growth rate was controlled to match the oxygen transfer rate, allowing the cultivation to be run in aerobic cultivation mode. Controlling the growth rate was particularly important because uncontrolled growth is often associated with uncontrolled protein synthesis rate and incorrect protein folding. Thus, high local concentration of recombinant protein may lead to formation of insoluble protein aggregates (inclusion bodies). This problem is commonly addressed by reducing the protein synthesis rate by the use of lower IPTG concentrations and/or lower induction temperatures. To ensure ideal refolding IPTG concentration and the time of induction were also investigated for the fed-batch approach.

We next investigated different strategies for CBM isolation from the biomass and media. As some CBMs bind irreversibly to cellulose and others bind reversibly, different separations and purifications were required.

4.5 Purification of CBMs

The isolation methods used to purify soluble protein depended on their intrinsic physiochemical property. Therefore, each CBM purification is described below. In this work, we focused on CBM2a (with and without an affinity tag) and CBM17. Interestingly, CBM2a was found inside and outside the cells, while CBM2a-His and CBM17 were almost exclusively inside the cell membrane. Although the fed batch approach resulted in excellent protein yields, it brought new challenges downstream with regard to protein purification. The quantity of cells generated was problematic to the extraction of the desired protein and it was important to develop a method that only requires "standard" equipment that will not request a massive investment for "regular" laboratories.

4.6 Cell disruption and CBM purification

Cell disruption on small scale (<10 mL) was performed chemically with a BugBuster (Sigma Aldrich), following the instructions of the manufacturer. Larger scale cell disruption was accomplished by sonication using a Q500 Sonicator (Qsonica, Newtown, USA) equipped with a 19 mm tip. Sonication was run with settings of 45 % amplitude and 15 sec and 45 sec on- and off-time, respectively. Cells were continuously kept on ice. After cell disruption, the cell debris was separated from the crude cell extract by centrifugation (6,600 g, Avanti, Beckman Coulter, Brea, USA). The CBM17 from the E. coli crude cell extract was purified with Avicel applying a ratio of 100 mg protein per 25 g Avicel. Binding was performed at 4 ºC overnight. The Avicel was then collected by filtration and the CBMs eluted by washing sequentially with PBS buffer (50 mM

phosphate, 100 mM sodium chloride, pH 7). Sodium phosphate buffer (50 mM, pH 7) and nanopure water. The CBM2a-histidine tagged was purified by affinity chromatography from the E. coli crude cell extract using Ni-NTA beads (Thermo Fisher) using 500 mM Imidazole as elution buffer. For both CBMs, the final step was a buffer exchange (HiPrep Desalting 16/20 column, GE Healthcare Life Sciences, Little Chalfont, UK) performed on an AKTA prime (GE Healthcare) with 5mL min-1 PBS buffer. The preparation of the bacterial lysate is a critical step as optimal conditions maximize cell lysis and the fraction of the recombinant protein that is extracted while minimizing protein oxidation, unwanted proteolysis and sample contamination.

4.6.1 Breaking cells

For the smaller volume cultivation, we opted for a commercial lysis product, bugbuster® (Millipore). This protein extraction reagent is formulated for gentle disruption of *E. coli* cell wall to liberate active proteins. It provides a simple, rapid, way of releasing expressed target protein in preparation for purification or other applications. The proprietary formulation utilizes a Tris-buffer based mixture of non-ionic and zwitterionic detergents that is capable of cell wall perforation without denaturing protein.

Two physical methods were tested for larger volume, high pressure homogenization and sonication. Because the volumes handled were often over 250 mL, these methods were really challenging. During homogenization, we installed a loop system allowing the cells to go through the homogenizer three times. This should lead to the complete rupture of the cell membrane. However, the protein yield obtain after purification were 20% inferior to the bug buster® treatment.

The heat generated during homogenization, which can damage the protein, may explain the yield drop. It was found that the homogenizer over-heat during the process, which may have led to protein degradation. Therefore, we investigated the sonication technique, in which the temperature was controlled using an ice bath. The volume treated needed to be divided in 100 mL increments to achieve maximum membrane rupture. The lysis buffer should contain a strong buffer (50–100 mM phosphate or HEPES) to overcome the contribution of the bacterial lysate, high ionic strength (equivalent to 300–500 mM NaCl) to enhance protein solubility and stability. However, CBMs are compact protein that are not particularly sensitive to protease, at least in short term. To confirm this assumption protease inhibitor was added to the cell lysate, without affecting yields.

Breaking the cells lead to the release of DNA that drastically increased the viscosity of the lysate. It was important to add the benzonase to facilitate the handling of the material. The lysate was centrifuged and clarified before affinity column or extraction. For the CBM2a cultivation, the supernatant and cell clarified cell lysate were combined to maximize CBM recovery.

4.6.2 Affinity extraction of CBM2a (type A)

Microcrystalline cellulose (25 g/L) was added to the supernatant of the *E. coli* culture and the cell lysate, the mixture was stirred for 4 hours at 4 ºC. The cellulose was centrifuged down (4500 rpm) and washed with phosphate buffer (more details can be found in *material and method*). As the CBM2a binding is “irreversible” ($k_{on} >> k_{off}$), its elution required the denaturation of the protein. The denatured protein is thought to go through a “molten globule” state, a compact denatured state with a significant amount of secondary structure but with substantially disordered tertiary

structures. Guanidine-HCl is the most popular protein denaturant because it is thought to stabilize this molten globule state and allow maximum recovery of the refold protein. For the case of CBM2a, 6 M guanidine were used to further linearize the protein. The CBMs were refolded slowly by adding, dropwise, the phosphate buffer (50 mM).

Several attempts of affinity cellulose column showed promising results. Ideally, swollen cellulose should be used as is has a high binding potential (250 mg/g). However, the pressure inside the column was not stable with this packing material, and ultimately was not functional. Therefore, homogenous microcrystalline cellulose was used. However, with the large volumes, due to the extracellular proteins, we opted to "fish out" the protein instead.

4.6.3 Affinity extraction of CBM17 (type B)

In contrast to type A CBMs, Type B CBMs bind reversibly to cellulose in mild condition. As a result, they can be eluted by decreasing the ionic force of the eluent. The CBM17 was eluted with nanopure water from the cellulose substrate. The intrinsic salt concentration acted to shield the ionic interactions affecting macromolecular stability as well as intermolecular binding reactions. Without salt, the proteins were unstable and might rapidly precipitate. It was essential the keep the eluted CBM17 in an ice bath and add the concentrated buffer after elution.

4.6.4 Affinity extraction of CBM2a-His (type A)

High capacity affinity column is often the favoured method for recombinant protein purification. It is simple and rapid compared to alternatives. These procedures make use of a particular property of a protein moiety or short amino acid sequence that is fused to the recombinant protein as a 'tag'. Development of nitrilotriacetic acid (NTA), which has four chelating sites, allows a stable interaction between Ni^{2+} and column matrix leaving two metal coordination sites free to interact with functional groups of proteins. A stretch of 6 histidine residues (His-tag) linked to the N- or C-terminal part of the recombinant protein is enough to allow a high affinity interaction with the Ni^{2+}-NTA resin.

Characterizing the purified protein in some detail reduces the risk of wasting resources on protein material of inadequate quality. It also provides a means to ensure that different batches of the same protein have similar properties.

4.7 SDS-PAGE analysis

After protein purification, the samples were resolved by denaturing SDS-PAGE. When stained with Coomassie brilliant blue, the intensity of the bands approximates the amount of protein. This allows an estimation of the purity of the sample and confirms the purified proteins molecular weight. The polyacrylamide gel was performed on precast gel (4–12% Criterion™ XT Bis-Tris Protein Gel, 18 well, 30 µl, Bio Rad, Hercules, CA, USA) using XT-MES buffer, XT reducing

agent (Bio Rad) The gel was run at 200V, 2A for 40min. Proteins where revealed using brilliant comassie blue (Bio Rad) and destaining solution.

4.8 Protein concentration using UV absorption spectroscopy

To quantify the amount and concentration of purified protein, the simplest and most common method is the Bradford assay, which measures the binding of Coomassie brilliant blue to the protein. However, better measurement were obtained using the CBM extinction coefficient at A280, which were previously reported in the literature (McLean, 2002). By taking a UV absorption spectrum, it was also possible to uncover contamination with DNA or RNA, or reveal common co-purifying cofactors (for example, NAD, FAD) (Structural Genomics Consortium et al., 2008).

4.9 Protein storage

To avoid CBM loss due to the freeze-melt cycles and precipitation, aliquots of the protein were stored, frozen in liquid nitrogen and stored at –80 °C. Concentrated protein (>1 g/ml) aliquots were unstable at 4 ºC and tend to be less stable to freeze-thaw cycles. We opted to store the CBM under diluted conditions and concentrated using centrifuge-driven filter devices with adequate molecular weight size cut-offs when needed. When we explored the stability of the protein to concentration and freeze-thaw cycles before processing the entire batch it was difficult to assess any trends or factors that would impact the protein. The frozen and thawed sample was compared with never-frozen proteins for biochemical activity. The CBMs were stored at 4 °C in PBS buffer with the addition of protein stabilizer (Millipore). The proteins were stable for several weeks (10 weeks

minimum) before any loss of binding. Before utilization, buffer exchange was performed to minimize any impact, likely due to the stabilizer. Overall, depending on the frequency of use of the protein, we recommend either 4 ºC storage + stabilizer (high frequency) or -80 ºC aliquots for longer storage.

4.10 CBM quantification

The concentrations of purified CBM17 and CBM2a were determined spectrophotometrically at 280 nm (Cary50 Bio, Varian, Palo Alto, USA), using a calculated molar extinction coefficients of 31010 M-1 cm-1 and 27625 M-1 cm-1, respectively.

4.11 CBM adsorption assays

The amount of bound CBMs was determined using depletion assays. Triplicate samples of CBMs, at concentrations ranging from 0.5 to 15 µM, were incubated in 1.5 mL tubes with 1 to 10 mg of insoluble cellulose in a final volume of 1 mL in PBS buffer. This ensured saturation of the substrate sites. Samples were incubated for 30 min at room temperature. The substrates were centrifuged (17,000 g, 20 min), the supernatant was collected and measured spectrophotometrically at 280 nm (Cary 50 Bio).

4.11.1 CBM adsorption on hydrolysed substrates analysis

The fibers were washed with SDS to remove bound enzymes, as described earlier. Afterwards a depletion assay using CBM2a and CBM17 was performed, as detailed later. In brief, CBMs (0.5 to 15 μM) were added to the NBSK (10 mg dry mass) in a final volume of 1 mL in PBS buffer. Samples were incubated for 30 min at room temperature, the supernatant separated from the fibers by centrifugation, and the concentration of CBMs in the supernatant quantified spectrophotometrically at 280 nm (Cary 50 Bio, Agilent Technologies, CA, USA). The molar extinction coefficients of CBM17 and CBM2a were 31 010 M^{-1} cm–1 and 27 625 M^{-1} cm^{-1}, respectively.

4.11.2 CBM adsorption on Cellulose nanocrystals

Carbohydrate-binding Module (CBM), Polyethylene glycol (PEG), (CBM-PEG) and CNC were mixed in a 10 mmol phosphate buffer and gently stirred for 1hour. Close to 100% of the CBM was adsorbed, based on the concentration of CBM left in the solvent after the removal of the CNC.

4.12 Bioconjugation

CBM2a (10 $mg.mL^{-1}$) were conjugated with Propargyl-N-hydroxysuccinimidyl ester (1.2 eq) in sodium carbonate buffer (pH 8.3) with 100 mM sodium chloride for 1 hour. The conjugated CBM2a-alkyne was then purified using a desalting column (HiPrep Desalting 16/20 column, GE

Healthcare Life Sciences, Little Chalfont, U.K.) performed on an AKTA prime (GE Healthcare) eluted with 5 mL.min^{-1} 50 mM phosphate buffer (pH7).

Chapter 5: The use of CBMs for characterization and functionalization of cellulose surfaces

5.1 Production optimization of CBMs

CBM production was optimized, yielding $g.L^{-1}$ quantities of the specific proteins, which were subsequently used to both characterize the surface morphology of lignocellulosic substrates and functionalize cellulose surfaces.

5.1.1 Flask cultivation optimization

Typically, the induction time should be between 0.6 and 0.8, which was consistent with the data obtained. It was found that CBM2a was produced in higher yields when the induction time was closer to an OD_{600} of 0.6 while CBM17 was produced consistently at OD_{600} 0.6 or 0.8. It seemed that there was little variation between the two strains. However, as will be described in the next section, the fed-batch cultivation was more sensitive to the induction time.

Next, IPTG concentrations were investigated, as they have been shown to influence protein yields (Donovan et al., 1996). If too little IPTG is added this can result in the failure to initiate protein production while too high IPTG concentration can inhibit cell growth, and therefore, the recombinant protein yields. For both CBM2a and CBM17 the optimal yields were obtain for IPTG concentration of 0.5 mM (**Figure 6**), This was consistent with the previous literature (Hasenwinkle et al., 1997). Because IPTG is costly, it may be better to only use 0.3 mM as the production is only

slightly lower. Once induced, the cultivation was allowed to stir for 24 hours before removing the cells from the media via centrifugation. At the end of the cultivation, the pH was determined to be close to pH4, indicating that the initial buffer was consumed by "overflow metabolites" such as acetic acid. *E. coli* cells produce acetic acid as an extracellular co-product of aerobic fermentation. The rate at which acetate forms is directly related to the rate at which the cells grow or the rate at which they consume their carbon source. To maintain an optimal pH of around 7 for cell growth, increments of 5 mL of phosphate buffer was added at 5 hours, 10 hours and 12 hours of cultivation. Although the final pH was determined to be pH6, there were no improvement in protein yields. This result can be explained by the fact that the maximal cell growth was reached before the pH decreased, and only the final stage of the protein production correlate with the pH drop. Acetate formation is even more critical during high cell density cultivation. In the next section, the fed-batch strategy chosen to minimize the effect of acidification and to obtain high yields with a relatively simple design is discussed.

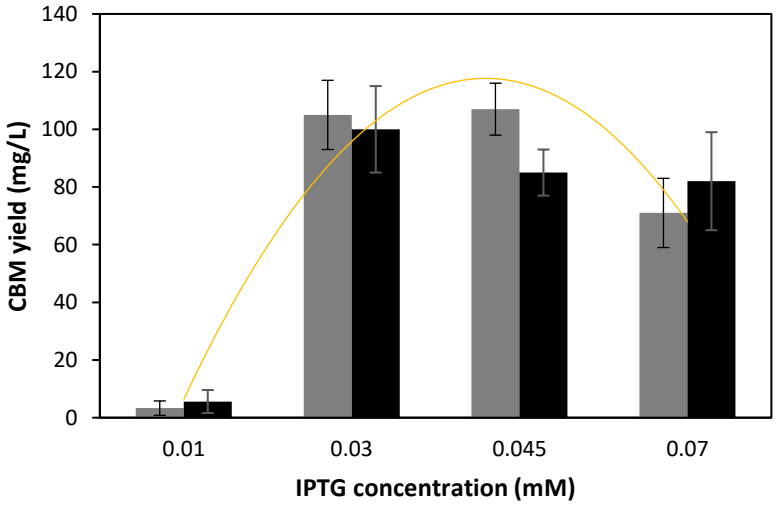


Figure 6: Influence of the Optical Density at the time of induction on CBM2a (black) and CBM17 yields (grey)

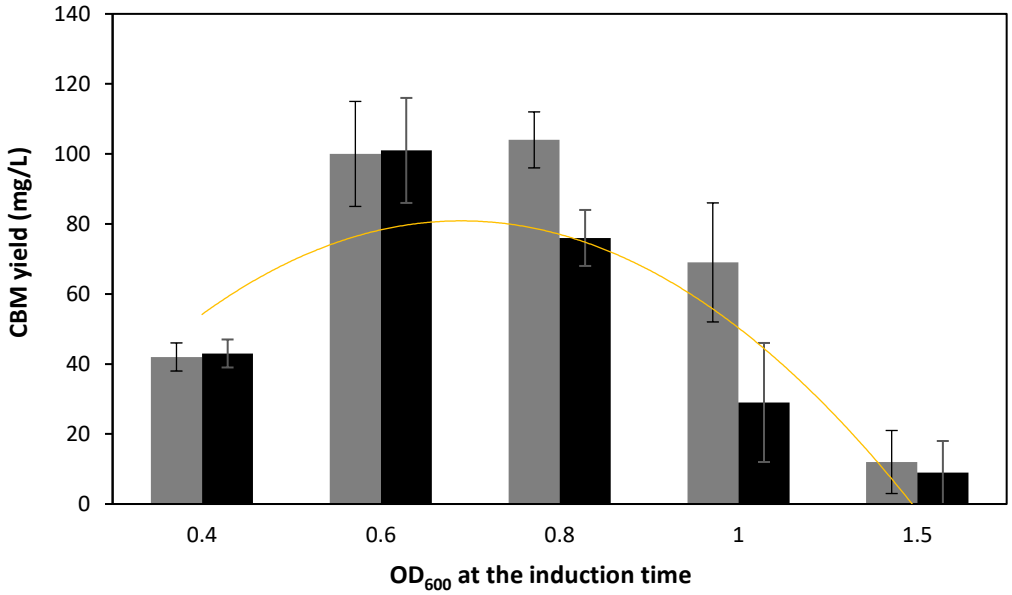


Figure 7: Influence of the IPTG concentration on CBM2a (black) and CBM17 yields (grey)

Using shaken-flask cultivation, a range of CBMs were produced (50-100 mg/L) . The shaken-flask parameters were optimized and provided the groundwork for the fed-batch approach.

5.1.2 Fed-batch process optimization for CBM production

In the designed set up, *E. coli* was successfully cultivated to high cell densities by applying the fed-batch principle. The growth rates, the final cell OD_{600}, and the CBM yields are summarized in **Table 4**. In both cases the μ_{set} was reduced to below 0.2 h^{-1} and a final OD_{600} of 105 (CBM17) and 145 (CBM2aH6) was obtained, resulting in successful, high cell density cultivations (Luli and Strohl, 1990; Riesenberg et al., 1991; Yee and Blanch, 1992).

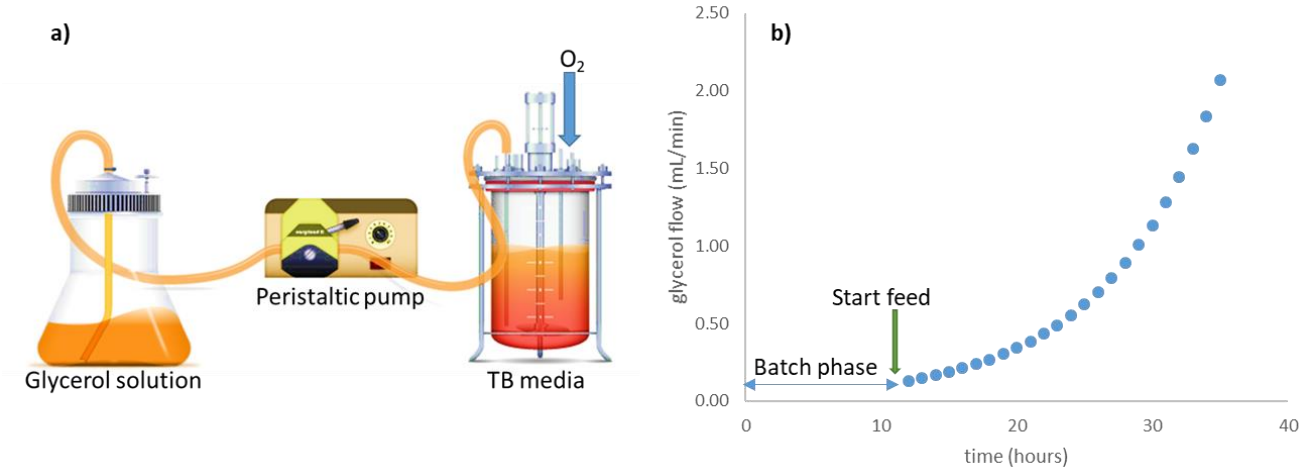


Figure 8: a) Schematic representation of the fed-batch reactor, the glycerol solution is exponentially fed to the cultivation, calculated in function of the estimated cell growth (Aïssa et al., 2019). b) Glycerol feed during cultivation.

Like the shaken-flask cultivation, IPTG concentration was ideal around 0.5 mM. However, the time of induction seems to be more sensitive in fed-batch cultivation. Ideally, the IPTG should be added to the culture when the OD is no more than 0.6. On average, the designed fed-batch cultivation produced 10 times more CBMs that the shake-flask cultivation, with biomass yields also 10-fold higher that shake-flask.

After cultivation, the CBM concentrations were estimated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) prior to purification. The CBMs sizes were between 11 kDa and 28 kDa. Therefore, the gel and elution were calibrated to separate relatively small proteins.

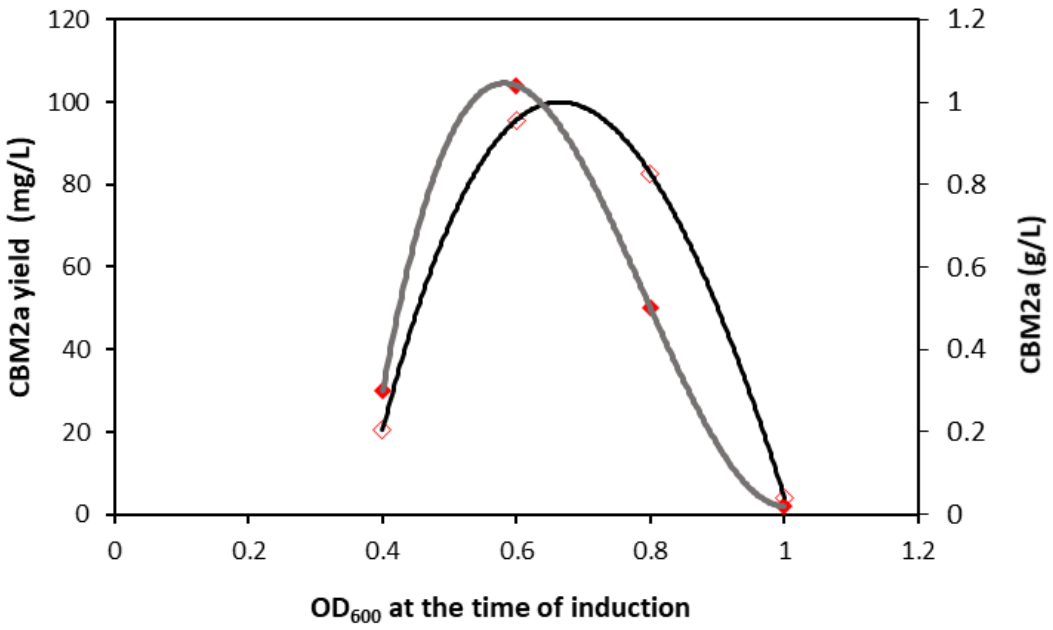


Figure 9: Influence of the time of induction on CBM yields, comparison between shaken-flask (black line, left axis) and fed-batch (grey line, right axis)

Table 4. High cell density fed-batch cultivation for CBM production. Summary of μ_{free}, μ_{set}, final OD_{600}, and CBM yields

Strain	μ_{free} (batch) [h^{-1}]	μ_{set} (fed-batch) [h^{-}	Final	CBM yield [a)] [g L^{-1}]
CBM17	1.34	0.12	105	1.04 (*0.10*)
CBM2a	1.26	0.14	136	1.00 (0.05)
CBM2aH6	1.11	0.16	145	0.32 (*0.05*)

a) Corresponding CBM yields achieved in shaken flask cultivations are shown in brackets

The fed-batch cultivation allowed the production of 0.5-1 $g.L^{-1}$ of CBMs, using a minimalized system that should be easily accessible. We developed an easy purification and stable storages of the proteins that facilitated the use of CBMs as analytic probes and as a vector of functionality for cellulose surfaces.

5.2 The use of CBMs to elucidate the relationship between structure, hydrolyzability, and accessibility of cellulosic substrates

One of the major impediments limiting the commercialization of enzyme-mediated processes for conversion of biomass to sugars, fuels, and chemicals is the relatively poor efficacy of the enzymes and proteins that are required for the hydrolysis of the cellulosic fraction. Although the efficiency of hydrolysis is often measured by the amount of sugar released, it is generally acknowledged that it is the restricted enzyme accessibility to the cellulosic substrate that is the rate-limiting step (Meng and Ragauskas, 2014). However, in-depth analysis of cellulose accessibility to enzymes (CAE) is challenged by the lack of quantifiable, reproducible, and relatively easy methods that allow quantification of CAE (Chandra et al., 2008). Previous work has used methods such as water retention (Bendzalova, 1996), microscopy (Daniel, 1994; Hildén et al., 2003; Nieves et al., 1991), and fiber characterization (e.g. aspect ratio)(Mansfield et al., 1999) as ways of trying to predict cellulose accessibility and, consequently, the effectiveness of enzyme-mediated hydrolysis.

In order to determine the role of accessibility at different levels of organization on influencing enzymatic hydrolysis rates, five different techniques were used to quantify cellulose accessibility on a range of pretreated biomass. These included Simons' Stain, FQA, water retention value, CBM adsorption, and nitrogen adsorption. These techniques have been described in previous sections.

5.2.1 Substrate characteristics

Substrate preparation will influence the structural parameters that influence enzyme-mediated hydrolysis. Model substrates were prepared with minimal amount of lignin to focus on the

carbohydrate structures. However, several parameters will still influence the hydrolyzability of cellulosic substrate including the degree of organization (paracrystalline vs crystalline region), the degree of hornification due to an eventual drying step, the particle size (FQA) and the amount of residual hemicellulose. For lignin-containing substrate, the previously cited parameters still play an important role in the substrate hydrolyzability. However, their respective effect can be undermined by the lignin properties and distribution. In addition to the hindrance effect, lignin can introduce bias in the measure of accessibility, mainly through unspecific adsorption or leeching effects. Thus, we wanted to develop a method using CBMs applied to "realistic biorefinery substrate" to try to determine the influential role of lignin.

The following section is divided between carbohydrate-only (model) and lignocellulosic substrates. With the model substrates, we standardized the method to provide consistent results, normalizing the CBM adsorption on Avicel and PASC. We also compared our results with previously published work. The CBM method was used to elucidate the relationship between fibrillation, hydrolyzability and accessibility of cellulosic substrates. After this initial work, lignocellulosic substrates were prepare using steam explosion pretreatment (STEX) or thermomechanical pulp (TMP) and refined mechanical pulp (RMP).

5.2.2 "Model" substrates including phosphoric acid swollen cellulose (PASC), microcrystalline cellulose (Avicel) and nanocrystalline cellulose (CNC)

CBM depletion assays, using type A (CBM2a) and a type B (CBM17) CBMs were initially used to compare CBM binding onto crystalline (Avicel) and more paracrystalline (PASC) cellulosic

substrates. Thus, it was possible to compare the reproducibility of their binding and the production quality between batches and with the previous literature (Table 5) (Boraston et al., 2001; Notenboom et al., 2001).

Table 5: CBM17 and CBM2aH6 adsorption on microcrystalline cellulose (Avicel) and phosphoric acid swollen cellulose (PASC)

Substrates	**Adsorption CBM2a ($mg.g^{-1}$)**	**Adsorption CBM17 ($mg.g^{-1}$)**
Avicel	29 (Boraston et al., 2001)	5.2 (Boraston et al., 2003)
	26 [a)]	4.6 [a)]
PASC/Ball-milled cellulose	178 (Boraston et al., 2001)	171 (Boraston et al., 2003)/ 96 (Araki et al., 2009)
	224 [a)]	196 [a)]/ 84 [a)]

a) Data measured in this study, as described in the material and method's section

It is important to connect the binding mechanism of the two CBMs (i.e. CBM2a and CBM17) to Avicel, CNC and PASC structures. The binding of CBM2a to crystalline cellulose is considered irreversible in a strict thermodynamic sense. The dilution of free CBM2a does not lead to the release of the already bound CBMs, which suggests a kinetic barrier to their desorption (McLean 2002). Given the demonstrated specificity of CBM2a for crystalline cellulose (Boraston et al. 2001), it is possible that its binding to PASC is restricted to highly accessible microcrystalline surfaces within the sample. PASC probably contains a high surface area of microcrystalline regions while exhibiting relatively low bulk crystallinity (McLean 2002). This would explain the binding capacity of the PASC to CBMs that specifically bind crystalline cellulose. Designing an

insoluble cellulose I substrate that would be completely paracrystalline (disorganized) will be difficult to achieve because the interaction between the cellulose chains will lead to a new arrangement in water to try to find a thermodynamically favorable state. In contrast, CNC represent a close to purely crystalline substrate, exhibiting no affinity for paracrystalline specific CBM17. With regard to the binding site of CBM17, the cellulose chain interacts is accommodated in the groove of the binding site. Thus, these differing binding site architectures are responsible for the ability of the CBMs to distinguish fine paracrystalline structure in cellulose.

Substrates with controlled swelling were prepared to confirm CBM specificity. Kraft fibers were swollen with different phosphoric acid concentration. As the Kraft fibers swell, the hydrogen-bonding network is disrupted, loosening cellulose chains. At a 64% acid concentration, the binding of the CBMs significantly increase (**Figure 10**). As it is likely that some less organized structures within the Kraft fiber are more sensitive to the acid treatment, the substrates will therefore display an apparent increase in accessibility. At 78% phosphoric acid, features such as peeling and delamination, roughening and fibrillation are apparent (SEM pictures, **Figure 11**). The NMR spectra also indicated a decrease of crystallinity of the substrate (*i.e.* CrI goes from 62 to 40) (**Figure 11**). Although the CrI decreased significantly, it seems that the measurement indicated a trend which was also be influenced by other factors (*e.g.* salt concentration, degree of polymerization). In the introduction X-ray Diffraction (XRD) and nuclear magnetic resonance (NMR ^{13}C solid state) were described as the most common methods to measure cellulose crystallinity. Several studies have questioned how the cellulose crystallinity varies during hydrolysis with various, sometimes contradictory conclusions, with the different type of substrates

as well as the analytical methods employed contributing to the lack of resolution. Furthermore, as mentioned earlier, *in situ* measurements of cellulose structure under reaction conditions (aqueous buffer) are challenging for these methods. Another element of why these methods cannot clearly elucidate this phenomenon is due to their non-specificity, with only "bulk" information obtained.

In the present work (**section 5.2),** we analyzed the reliability of the CBM adsorption to estimate the substrate crystallinity (i.e. organization) on carefully prepared model substrates. In the **section 5.3**, we hoped to better elucidate how the cellulose organization actually varies during hydrolysis.

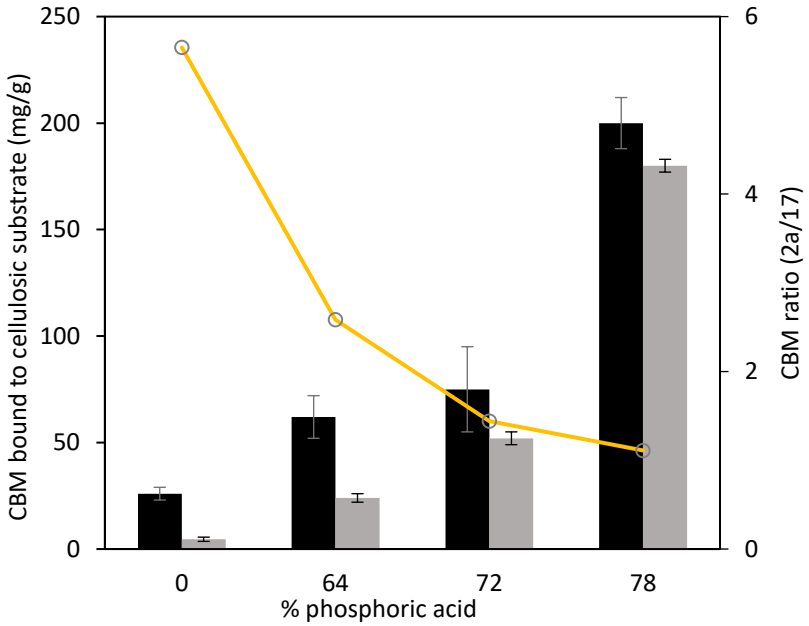


Figure 10: Adsorption of crystalline cellulose-binding CBM2a (black) and paracrystalline cellulose-binding CBM17 (grey) to Kraft fibers swollen with a range of phosphoric acid concentrations. The yellow line represents the ratio of the two CBMs. Experiments were run in triplicate and error bars represent one standard deviation from the mean.

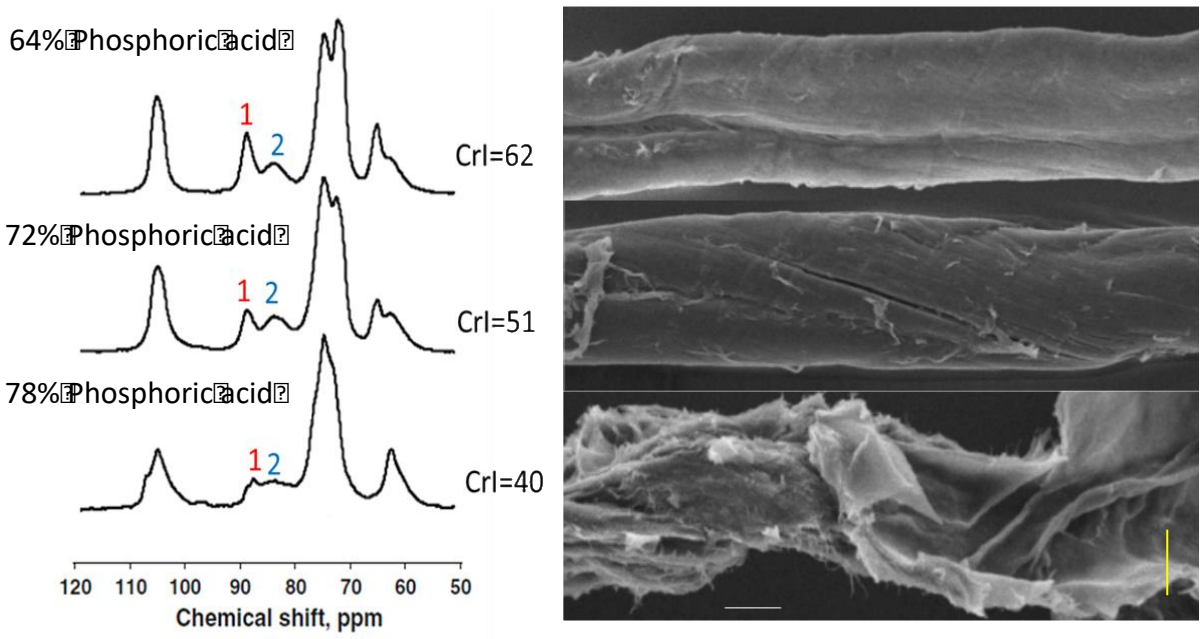


Figure 11: ^{13}C NMR (left) and SEM pictures (right) of Kraft fibers treated with a range of phosphoric acid concentrations. Red 1 and blue 2 are represents respectively the crystalline and paracrystalline peaks used to calculate the crystallinity index (CrI). Scale bar (yellow) is 10 µm.

Further investigation, on CBMs binding specificity was conducted by mixing PASC and CNC (**Figure 12**). As expected, the CBM17 did not bind to CNC and CBM2a bound to around 30 $mg.g^{-1}$ of substrate. Trying different ratios of these two substrates, the increase of binding of the two CBMs was almost linear. Similar to the swollen Kraft fibers, PASC exhibited a high binding capacity for type A and B CBMs (≈190 $mg.g^{-1}$). It seems that mixing the two substrates did not influence the binding of the CBMs compared to their adsorption on each substrate separately. It was concluded that there is little to no interaction between the PASC and CNC that would block the CBM binding to their respective substrates.

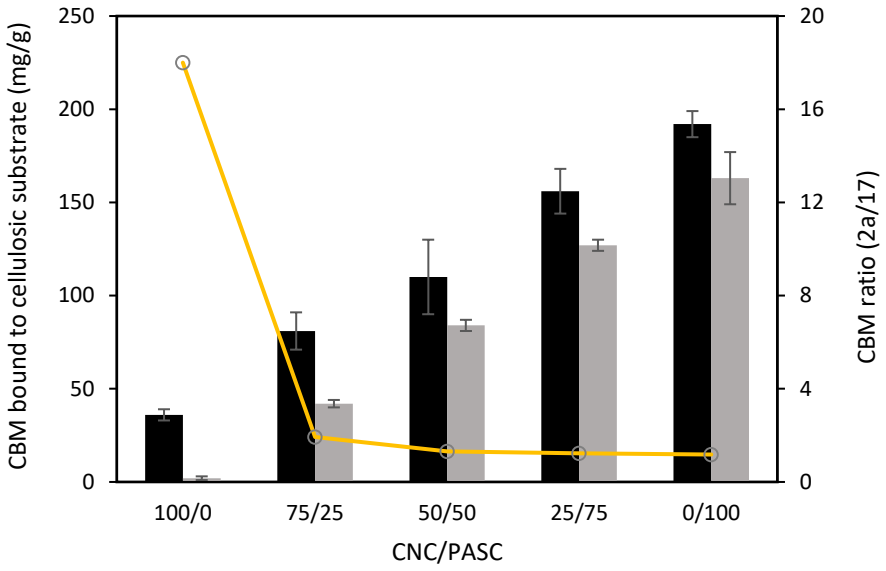


Figure 12: Adsorption of crystalline cellulose-binding CBM2a (black) and paracrystalline cellulose-binding CBM17 (grey) to PASC and CNC. The yellow line represents the ratio of the two CBMs. Experiments were run in triplicate and error bars represent one standard deviation from the mean.

5.2.3 CBM adsorption to elucidate the relationship between fibrillation, hydrolyzability, and accessibility of cellulosic substrates

The primary goal of this part of the thesis was to investigate the relationship between fibrillation, accessibility and hydrolyzability. To minimise the influence of components such as lignin and hemicellulose on substrate accessibility to enzyme, a range of NBSK pulps were prepared. The substrates were differentially refined with increasing refining energy to provide various microfibrillated cellulose (MFC) substrates (Hu et al., 2011; Mansfield et al., 1999). As indicated

in **Table 6**, very little lignin was present in any of the MFC substrates and the overall carbohydrate composition remained largely unchanged.

When each of the MFC substrates was assessed by SEM, the increased fibrillation was apparent with higher refining energy (**Figure 13**). The unrefined MFC0 substrate was primarily composed of longer, intact fibers with few "wrinkles". The refined MFC substrates appeared to contain increasing amounts of fibrillated fibers and the proportion of thin thread-like fibrils increased significantly from 500 kWh ton^{-1} (**Figure 13.b**) to 1000 kWh ton $^{-1}$ (**Figure 13.c**) and were highest at 1500 kWh ton $^{-1}$ refining energy; at this refining energy the original fiber structure almost disappeared (**Figure 13.c**). The increase in apparent fibrillation of the MFC substrates, as shown by the SEM images, suggested improvements in accessibility and hydrolyzability (Nieves et al., 1991).

Table 6: The chemical compositions of the microfibrillated cellulose (MFC) substrates [a)]. Glucan, xylan, galactan, mannan and lignin are expressed in % of dry mass. Data represent the mean value and standard deviation of triplicates

	Glucan	Xylan	Galactan	Arabinan	Mannan	Lignin
MFC0	80.4 ± 0.7	8.7 ± 0.1	0.4 ± 0.0	0.8 ± 0.0	6.6 ± 0.1	2.3±0.1
MFC2	80.6 ± 0.3	8.5 ± 0.1	0.5 ± 0.0	0.8 ± 0.0	6.6 ± 0.2	1.9±0.2
MFC4	78.3 ± 2.1	8.1 ± 0.2	0.4 ± 0.0	0.7 ± 0.0	6.4 ± 0.2	1.9±1.2
MFC6	81.0 ± 0.5	8.6 ± 0.1	0.5 ± 0.0	0.8 ± 0.0	6.8 ± 0.2	2.0±0.2

[a)] The MFC0 sample corresponds to a refining energy of 0 kWh.ton^{-1}. The MFC2, MFC4

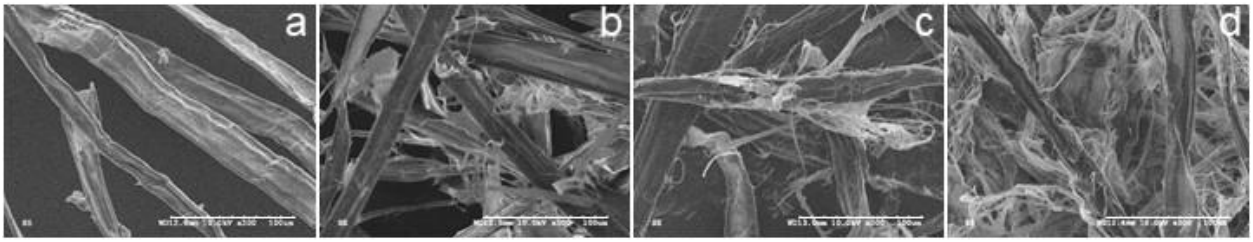


Figure 13: The impact of the refining energy on the fiber morphology. Scanning electron microscopy (SEM) images of MFC0 (panel a, 0 $kWh.ton^{-1}$), MFC2 (panel b, 500 $kWh.ton^{-1}$), MFC4 (panel c, 1000 $kWh.ton^{-1}$), and MFC6 (panel d, 1500 $kWh.ton^{-1}$). The white scale bars represent 100 µm.

To further evaluate the impact of refining on the substrates, the MFC substrates were characterized with regard to their overall fiber dimensions (here given as the aspect ratio, i.e. fiber width vs. fiber length) and the ability to swell by measuring the water retention value (WRV). These two parameters have been shown to influence accessibility and have been successfully used to analyze and predict hydrolyzability (Ali et al., 2001; Arantes and Saddler, 2011).

As shown in **Figure 14**, the WRV of the substrates increased from 18 $g.g^{-1}$ to 23 $g.g^{-1}$, going from low (MFC0) to high (MFC6) refining energies. Based on the SEM images and the observed increase in fibrillation, this was anticipated (Liu et al., 2016). The aspect ratio of the MFCs, also follows a disparate trend (**Figure 14)** decreasing with increasing refining energy. The combined results of these two assays seemed to confirm the SEM observations that the accessible cellulose surface area continuously increased with higher refining energies. This suggested that an increase in accessibility could be anticipated to result in a convergent trend in hydrolyzability. The enzymatic hydrolysis efficiencies were thus expected to increase from MFC0 to MFC6.

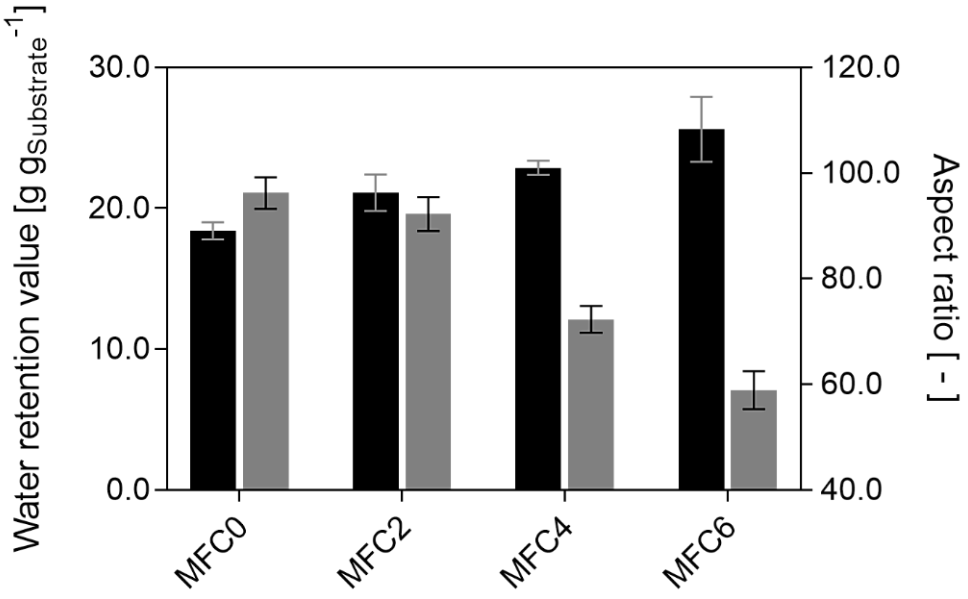


Figure 14: Determining the likely cellulose accessibility of the various microfibrillated cellulose (MFC) substrates by changes in Water Retention Value (black bars) and the aspect ratios (light grey bars). Data represent mean values of triplicates, the error bar indicate the standard deviation.

5.2.3.1 Enzymatic hydrolysis of the MFC substrates – analysis of initial rates and yields

The impact of refining on the MFC substrates hydrolysis was analyzed and the initial rates as well as the yields after 24 h of reaction are summarized in **Figure 15**. A threshold for improved hydrolysis efficiency with increased refining energy has been suggested previously (Liu et al., 2016; Park et al., 2016), where limited or only incremental improvements in hydrolyzability were achieved with further refining above 4000 revolutions of PFI refining. As the initial characterization of the MFC substrates showed a continued increase in apparent fibrillation (SEM, WRV, and aspect ratio), the levelling-off of the enzymatic hydrolysis was unexpected. However, although these characteristics can affect accessibility(Mansfield et al., 1999) it does not provide evidence of the ability of the enzyme to actually access the binding site on the microfibril.

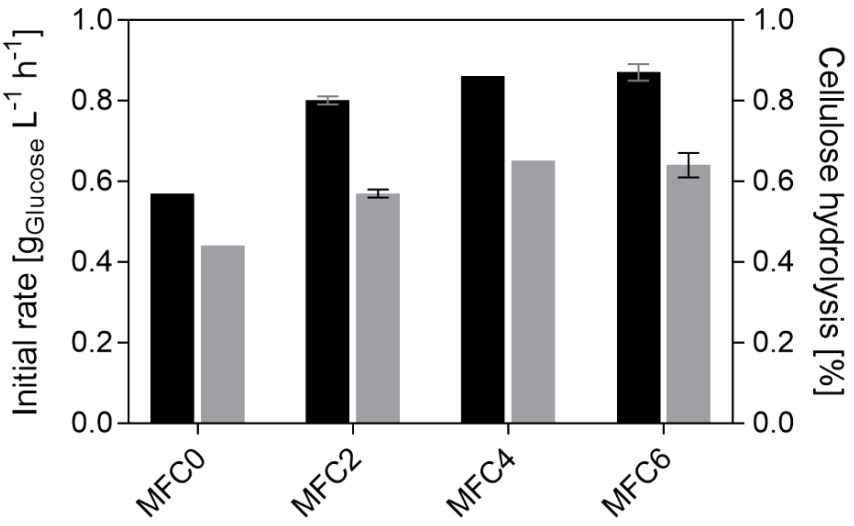


Figure 15: The impact of increasing refining energies on ease of enzyme mediated hydrolysis of the microfibrillated cellulose (MFC) substrates**.** Comparison of the initial glucose release rates (black bars) and cellulose hydrolysis after 8 h (dark grey). Data represent mean values of duplicates, error bars show the spread. Data represent mean values of duplicates, error bars show the spread.

To analyze if accessibility was affected by refining, binding studies with a type A (crystalline cellulose) and a type B (paracrystalline cellulose) CBM were performed.

5.2.3.2 CBM adsorption studies

The CBMs were used to investigate the various MFC substrates to see if their binding was convergent with the results obtained previous methods (*e.g.* WRV, SEM), or if they followed the observed hydrolysis profiles. CBM2a with histidine tag (CBM2aH6) was used for this study. The histidines have been shown to not affect the CBM2a binding. The results are summarized in **Figure 16**.

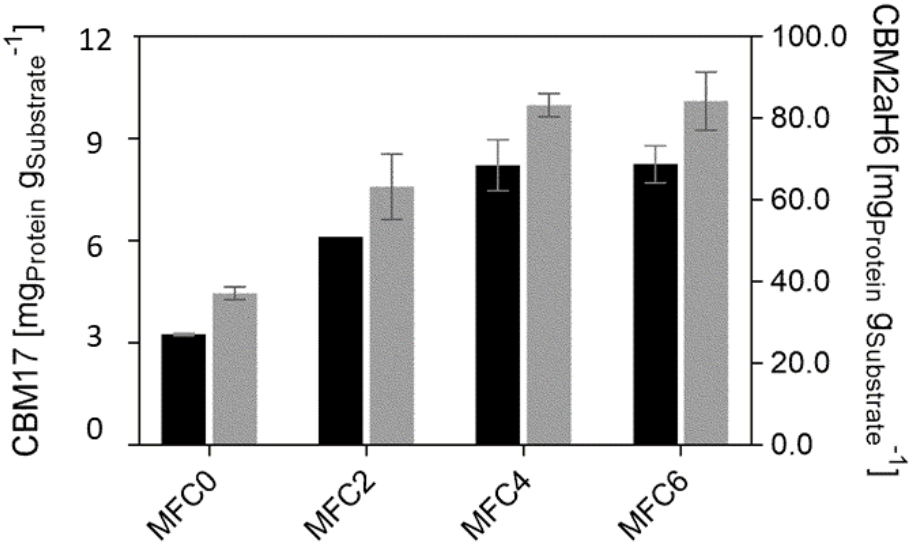


Figure 16: The impact of increasing refining energies on carbohydrate-binding module (CBM) accessibility to the microfibrillated cellulose (MFC) substrates. The extent of CBM17 (black bars) and CBM2a (grey bars) binding to the increasingly refined MFC substrates is indicated. Data represent mean values of triplicate experiments. Error bars show the standard deviation.

It was apparent that the CBM2aH6 and CBM17 adsorption increased from 35 $mg.g^{-1}$ to 80 $mg.g^{-1}$ and from 3 $mg.g^{-1}$ to 9 $mg.g^{-1}$, respectively, when they were added to the MFC0 and MFC4 substrates. However, increasing refining from 1000 kWh^{-1} (MFC4) to 1500 kWh^{-1} (MFC6) did not result in an increase in binding of either of the two CBMs. These results mirrored the hydrolysis profiles of the refined substrates, where no increase in hydrolysis was observed between the MFC4 and MFC6 substrates (**Figure 16**). These results suggested that, although initially beneficial, through the reorganization and dispersion of microfibrils, further mechanical refining did not significantly increase the accessibility of the cellulolytic enzymes to the substrate at the nanoscale. Furthermore, the consistency of the CBM17/CBM2a binding ratio suggested that mechanical refining does not result in modification of the cellulose surface morphology (crystalline vs. paracrystalline cellulose).

Mechanical refining has been suggested as one way to overcome biomass recalcitrance (Jones et al., 2013), facilitating biomass conversion by increasing the biomass accessibility and offering a commercially proven and scalable "front end" for biorefining processes. It was apparent that the never-dried NBSK pulps that were refined with increasing refining energies showed a significant increase in fibrillation as indicated by SEM imaging, as well as WRV and aspect ratio measurements. However, although beneficial for initial hydrolysis, no subsequent increase in cellulose hydrolysis was observed in the pulps that were refined at energy intensities higher than 1000 $kWh.ton^{-1}$. When probes based on the selective adsorption of CBM2a and CBM17 were used to assess the potential ease of hydrolysis of the various MFC substrates, their adsorption followed the same trend as the hydrolysis profile. This suggested that refining results in a dispersion of the microfibrils without affecting enzyme accessibility at the microfibril level (**Figure 17**). It seems that an optimal beneficial treatment can be obtained with refining substrate and further energetic increase would be inefficient. As a result this energy limit depends on the method used to refine the substrate.

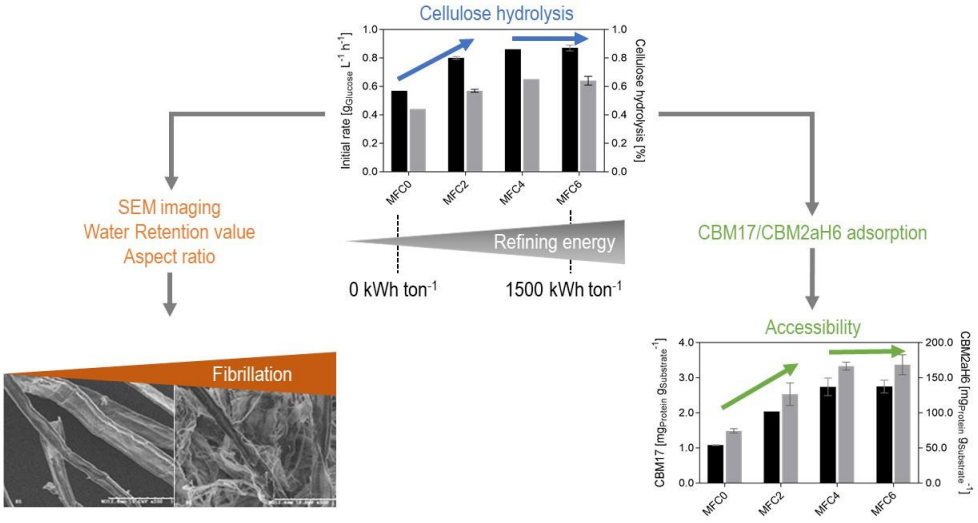


Figure 17: result summary of the CBM investigation on the relationship between fibrillation, hydrolyzability, and accessibility of cellulosic substrates

5.2.3.3 CBM adsorption on steam exploded corn stover (CS-STEX), steam exploded lodge pole pine (LP-STEX), thermo-mechanical pulp (TMP) and mechanically refined pulp (RMP).

Contrary to model substrates, such as Avicel or PASC, biorefinery substrates are likely to contain at least some lignin. Depending on the lignin type and the pretreatment effects, the characteristics and factors that will limit the hydrolysis will be quite different. Thus, four different types of lignocellulosic substrates were adopted in this study, which had different lignin types and amount. The depletion assays were conducted with and without prior treatment with bovine serum albumin protein (BSA), which have shown to reduce the unspecific binding of enzymes to lignin. Previous studies have shown the affinity of CBMs for lignin, which is particularly problematic for type A CBM due to their exposed binding face. As described earlier, type A CBMs bind through their C-H-π interaction with relatively flat and organized carbohydrate structures. Typically, non-modified

lignin occurs in the RMP substrate and should have little affinity with CBM, while condensed lignin (e.g. STEX substrates) will likely exhibit some unspecific binding. This is due to the condensation of the lignin during steam explosion, likely leading to flatter lignin surfaces that allow π-π bonds to be formed with the tryptophans of the proteins. For all substrates it appeared that the BSA treatment effectively reduced unspecific binding, with minimal influence on the RMP and TMP (**Figure 18**).

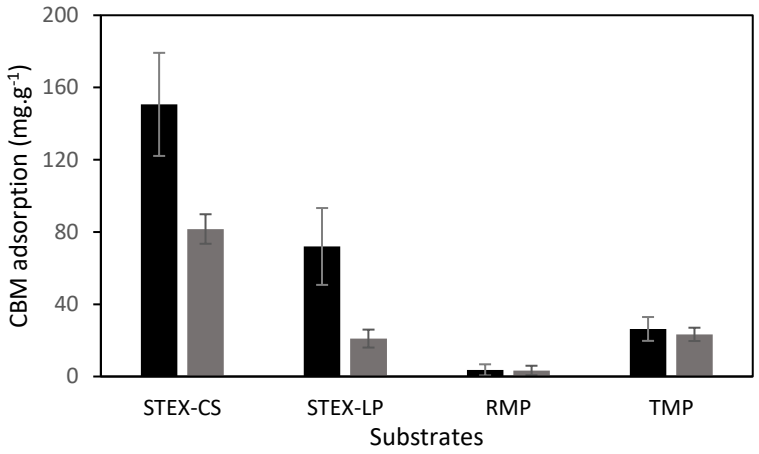


Figure 18: CBM2a adsorption on lignocellulosic substrates with (grey) and without (black) BSA treatment. Error bars show the standard deviation.

As expected, the reproducibility of the CBM adsorption proved to be more challenging than previously achieved using model substrate with up to a 30% error. Even after several washes, it seemed that some lignin was leeching from the substrate in the supernatant, interfering with the protein concentration measurement. However, this problematic issue motivated the development a method using fluorescent-CBM and CLSM to reduce the heterogeneity of the substrate and the interference issue.

5.3 Fluorescent CBM probes as a new quantitative method to study structure specific accessibility and monitoring of cellulose surface changes during enzymatic hydrolysis

In the previous section, CBM adsorption was shown to provide a good estimation of cellulose accessibility. The methodology provided an estimation of the substrate bulk behavior. However, within each substrate, many heterogeneous structures exist at different levels of organization. Being better able to assess these substructures will help us better understand the mechanisms involved during enzymatic hydrolysis. As mentioned earlier, the overall cellulose structure is a key parameter which influences accessibility with methods such as XRD and NMR, which are often used to categorize cellulose into either crystalline or paracrystalline substructures, too limited to fully picture what is happening at the enzyme level.

As discussed earlier, fluorescence-tagged CBMs have been successfully used to assess the glyco-architecture of plant cell walls when combined with confocal laser scanning microscopy (CLSM). As CBMs of different affinity do not compete for binding sites, they have the potential to be used simultaneously to assess cellulose structures as well as its impact on enzyme accessibility with previous studies illustrating the differential binding of CBMs to cellulosic fibers (**Figure 19**). It was apparent that CBM44 (type B) are more specific for disaggregated cellulose regions while type A CBM was bound along the entire fiber.

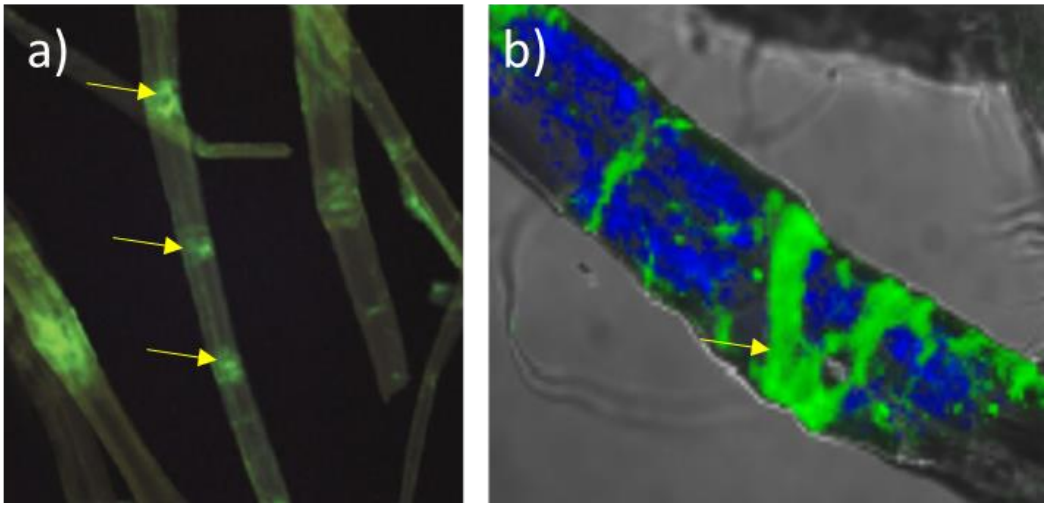


Figure 19: a) Fluorescent imaging of delignified fiber with CBM3a-GFP (adapted from Filonova et al. 2008), b) Fluorescent imaging of delignified fiber with CBM2a-blue and CBM44-tagged (adapted from 2013), and yellow arrows represent dislocation zones.

Building on these previous studies, the work described below used two fluorescence tagged CBMs, a type A (CBM2a) and a type B (CBM17) to assess lignocellulosic substrates using the CLSM method. High throughput computer-aided image analysis was used to provide quantitative data to track the changes happening and assess how key substrate characteristics might influence enzyme-mediated deconstruction of cellulose. The CLSM images were complemented by SEM imaging, ^{13}C NMR analysis and fiber length quantification. As described in more detail, the less organized regions were shown to be more susceptible to rapid degradation, leading to fiber fragmentation and the production shorter fiber fragments.

5.3.1 Imaging lignocellulosic substrate using CLSM

The pretreated substrates were imaged using one or two CBMs depending on the lignin concentration in the substrate (**Figure 20**).

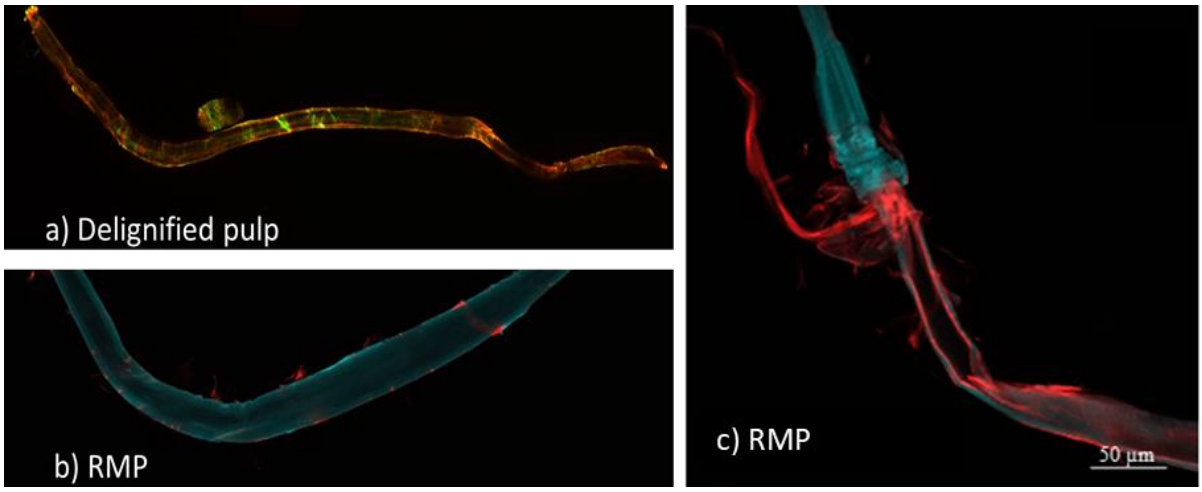


Figure 20: CLSM imaging of differently pretreated lignocellulosic substrates. In red, crystalline cellulose-binding CBM, in green, paracrystalline cellulose-binding CBM. The blue corresponds to lignin autofluorescence a) delignified Kraft pulp, b) and c) mechanically treated hardwood fiber.

These images illustrated the influence of lignin on cellulose accessibility. The delignified pulps showed red and green fluorescence without discontinuity all along the fiber while the RMP sample (b) was entirely shielded by lignin. However, Sample (c) exhibited red fluorescence, suggesting that the disruption of the fiber, through mechanical shear, had "freed-up" some of the cellulose. The lignin exhibits fluorescence emission spectra that peaks at ≈360 nm on excitation at wavelengths ranging from 240 to 320 nm. This can be explained by non-radiative energy transfer from lignin substituted aromatics. In **Figure 21**, steam exploded substrates containing lignin (≈25% wt), were imaged with the two florescent CBMs. Like the delignified sample (**Figure 20.a**) the fiber exhibit red and green fluorescence. However, for the STEX sample, black spots were observed (yellow arrows), likely showing lignin droplets, which had recondensed on the surface of the fiber. This was further observed using the autofluorescence, where the lignin spots showed characteristic pattern (Araya et al., 2015; Donohoe et al., 2008; Shevchenko et al., 1999; Xiao et

al., 2011). As the lignin after steam explosion treatment has been shown to lead to high unproductive binding (Lu et al., 2016), BSA had been added earlier, likely explaining the absence of fluorescence in these zones. These images showed that CBM2a and CBM17 bound specifically to different lignin-containing substrates. In subsequent work we wanted to quantify the fluorescence emitted by the CBMs.

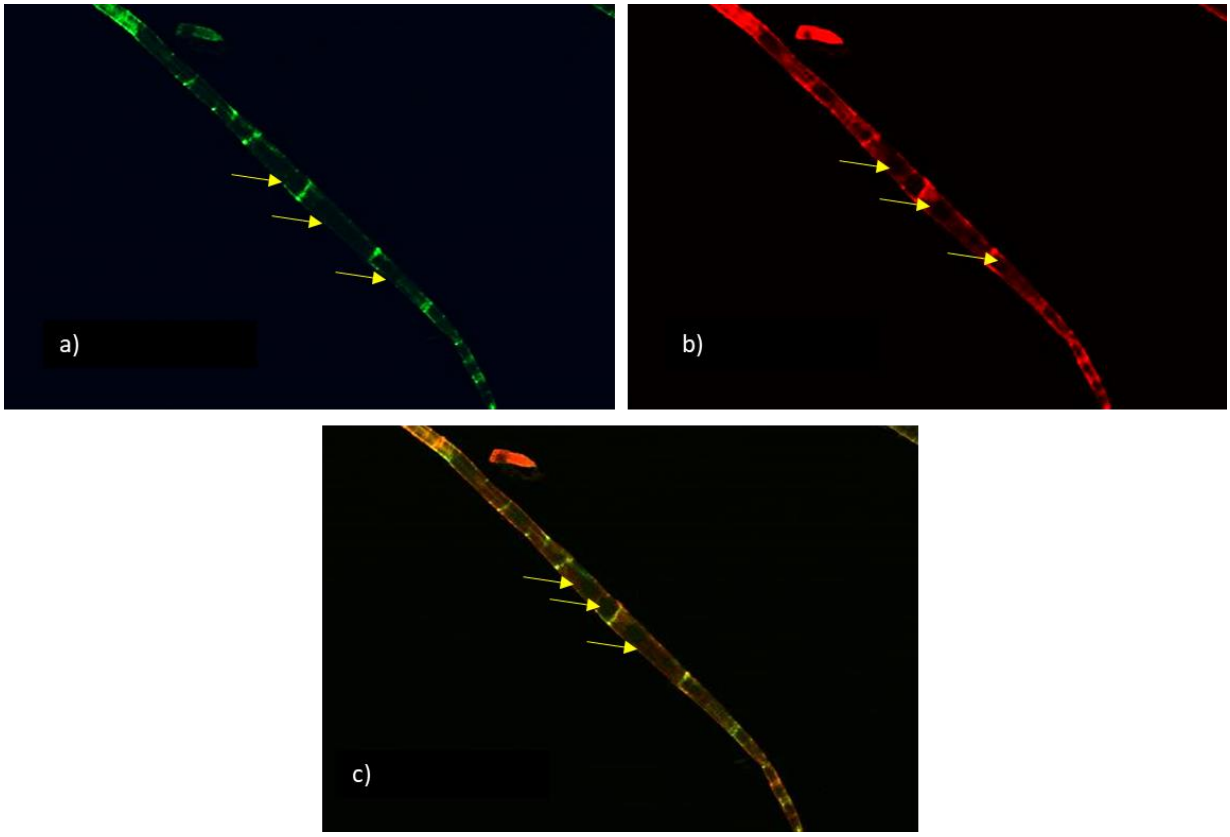


Figure 20: Fluorescence imaging of steam exploded softwood, a) CBM17-, b) CBM2a-Rhodamine and c) superposition of both channels. Yellow arrows indicate lignin droplets.

5.3.2 Structure changes during enzymatic hydrolysis

Previous work has suggested that accessible regions within the cellulose wood and plant fibers are more receptive to enzyme-mediated hydrolysis (Thygesen et al., 2011). These regions, sometimes referred to as dislocation zones, have been shown to be mechanically weak points in the fiber (Hidayat et al., 2012), consequently making them of ongoing interest to the pulp and paper sector (Galbe and Zacchi, 2002). Polarized light microscopy (Clarke et al., 2011; Eder et al., 2008; Hidayat et al., 2012; Thygesen et al., 2006) and scanning electron microscopy (SEM) (Eder et al., 2008; Suchy et al., 2009) have been used to try to better understand the characteristics of these disruptions. These less organised or so-called dislocation zones have been shown to vary from reversible microcracks at the surface to deep holes across the fiber (Hidayat et al., 2012). However, the exact nature of these disruptions is not clear. For example, the birefringence of these regions has been interpreted as indicating high crystallinity (Thygesen et al., 2006) and that the microfibrils continue unbroken through these regions. However, related work has suggested that these localized areas of morphological changes are predominantly paracrystalline, based on their crystallinity index (Hidayat et al., 2015) and susceptibility to acid (Ander et al., 2008) and enzymatic (Thygesen et al., 2011) hydrolysis. An ongoing challenge has been to monitor changes of cellulose structures as the enzyme-mediated deconstruction proceeds. As previously mentioned, the crystallinity of the substrate during enzymatic hydrolysis is still debated and the proposed mechanisms of peeling, delamination or swelling (amorphogenesis) are still debated.

5.3.3 Fluorescence-tagged CBMs analyzed under CLSM to assess the intrafibrillar structures of Northern Bleached Softwood Kraft pulp (NBSK) during the enzymatic hydrolysis

Previous work has used photoactivated localization microscopy to try to the determine binding selectivity of six CBMs, including CBM2a and CBM17. The work experiments on various cellulosic substrates (Fox et al., 2013). As expected, the results suggested that CBM2a bound to areas with a higher degree of structural organization than did the CBM17. However, CBM2a also showed a high degree of binding promiscuity. Although it primarily targeted the cellulose crystal, it also bound to less ordered paracrystalline regions. This observation is consistent with the previous section results and can be explain by the diversity on structures within one substrate. As discussed previously, CBM2a (type A) and CBM17 (type B) were covalently bound to Rhodamine-Red X (red) and FITC (green) respectively to assess the supramolecular structure of Northern Bleached Softwood Kraft pulp (NBSK) fibers.

When the two CBMs were used to follow the initial hydrolysis (**Figure 21**) it was apparent that, at time 0 (**Figure 21.a**) most of the fibers were interspersed with green zones (CBM17-FITC). Earlier work had suggested that these less organized regions (green) were where enzymatic activity was most pronounced, resulting in their more rapid degradation (Hidayat et al., 2015).

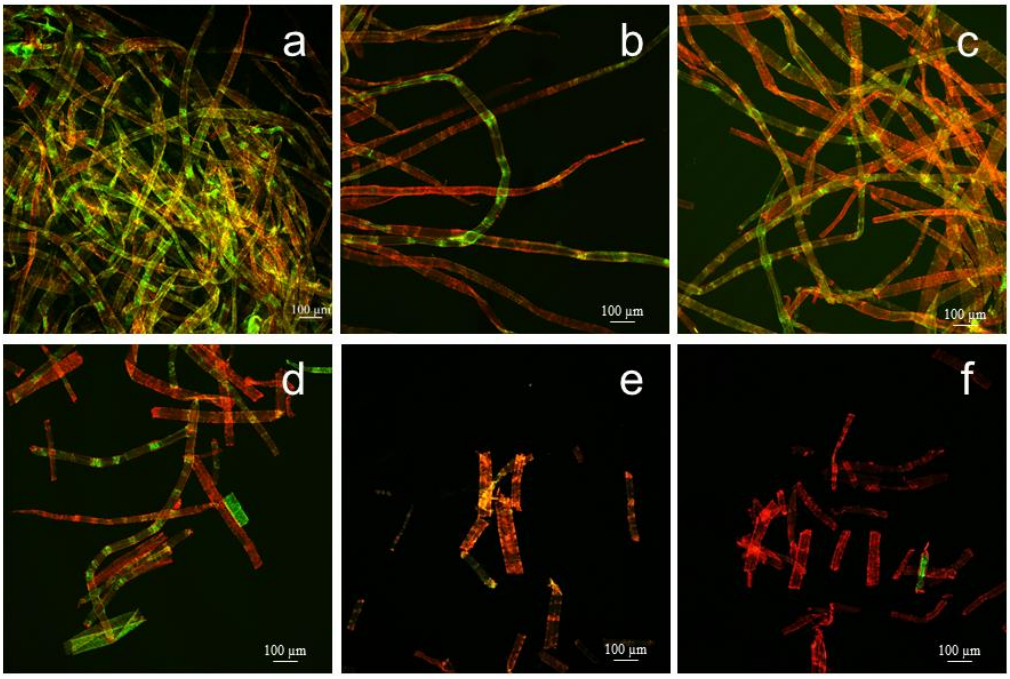


Figure 21: Changes in the morphology of Northern Bleached Softwood Kraft (NBSK) pulp fibers after 0 h (a), 0.5 h (b), 1 h (c), 2 h (d), 4 h (e), and 6 h (f) of hydrolysis, visualized using CBM17-FITC and CBM2a-RRedX binding with CLSM. (Scale bars represent 100 µm).

5.3.4 Quantifying fluorescence: high through-put image analysis

During the acquisition of digital images, the photons are converted to an intensity value and this is correlated to the number detected photons. Thus, it is possible to use the digital images acquired to calculate different spatial information such as distance and area as well as an intensity value. With this intensity value, it is possible to determine local concentrations of a fluorophore.

Following these qualitative results, we assessed the concentration of fluorophores and their distribution within the substrates during cellulose conversion by obtaining batch data and user

supplied inputs retrieved from a spreadsheet of the acquired micrographs. This consisted of a 3-step process involving pre-processing, segmentation and analysis. The workflow of the image analysis is illustrated in **Figure 22** with an example of "masking" of the bleached fiber. The fluorophores intensities were extracted from the acquired 3D numpy array as separate grayscale images. The background intensities of the red and green layers, acquired from blank slide micrographs analyzed with ImageJ, were subtracted from the micrograph of interest.

It is apparent that, after 2 h (**Figure 21.d**), fiber fragmentation occurred. Concurrently, the red fluorescence, which represents the binding of CBM2a-RRedX, proportionally increased while the green fluorescence intensity (from the binding CBM17-FITC) decreased (**Figure 23**). This confirmed that enzymatic activity was most pronounced in the less-organised zones, resulting in their more rapid degradation and fiber fragmentation.

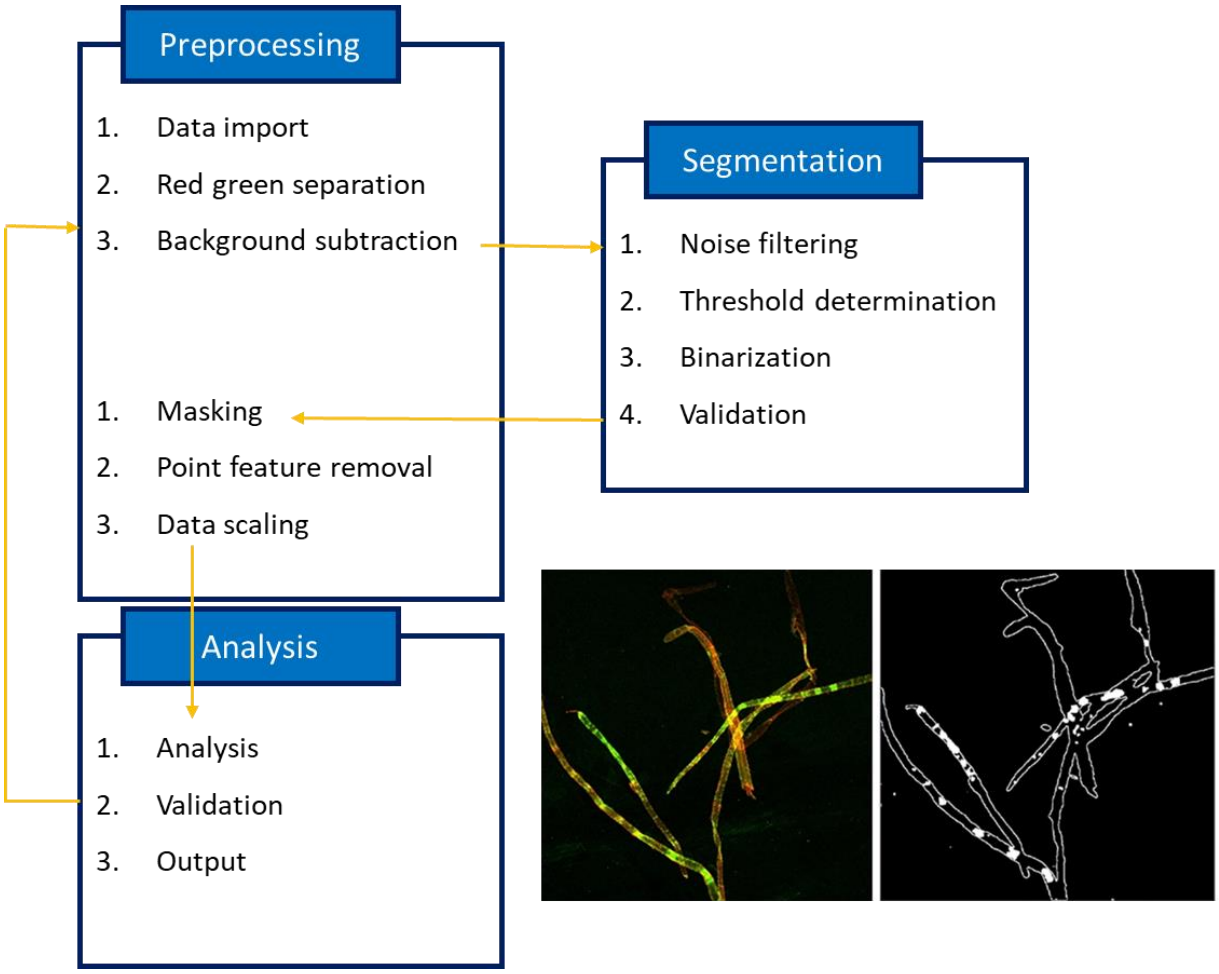


Figure 22: Overview of the quantitative image analysis workflow and fluorescence imaging of softwood delignified pulp and "masking" of the fiber highlighting the high intensity fluorescence of CBM17.

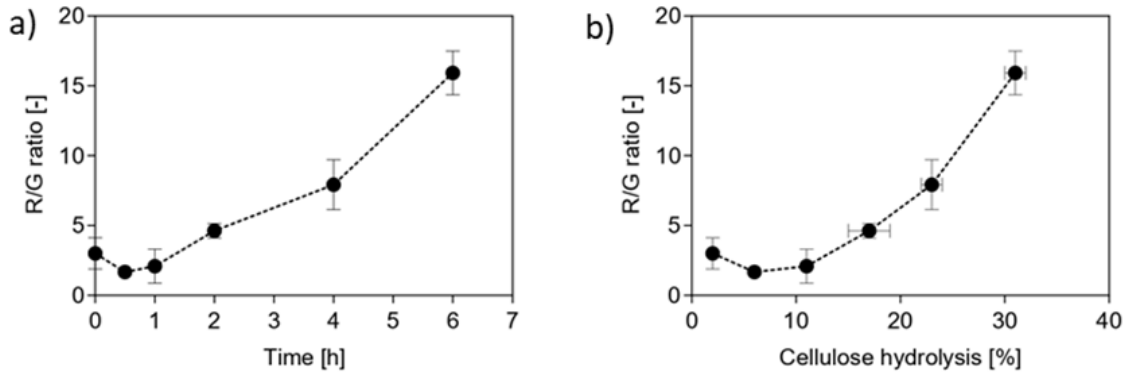


Figure 23: a) change in R/G ratio over time, and d) change in R/G ratio in residual substrates over time. Duplicate experiments were performed. Data in a) and b) show mean values and the spread. Data in a) and b) was derived from 8 to 10 image frames and mean values and the standard deviations are shown.

5.3.5 The influence of enzyme-mediated cellulose hydrolysis on fiber size

When the length and skewness of the NBSK fibers during hydrolysis was assessed using a Fiber Quality Analyzer (FQA) (**Figure 24**), the mean fiber length increased during initial hydrolysis, likely due to the faster degradation of the smaller, "fines" particles. Over the same time period of time the skewness plateaued and then decreased, indicating the fragmentation of the long fibers into shorter more homogenous fibers (**Figure 24**). Once 25 to 30% of the cellulose had been hydrolyzed, the mean-length of the fibers leveled out at ~0.2 mm (**Figure 24**). These results were consistent with previous work (Clarke et al., 2011). The CLSM images indicated that the fiber fragments were located approximately at the same distance between major structural defects, suggesting that fiber fragmentation occurred at the disorganized zones in the different substrate types.

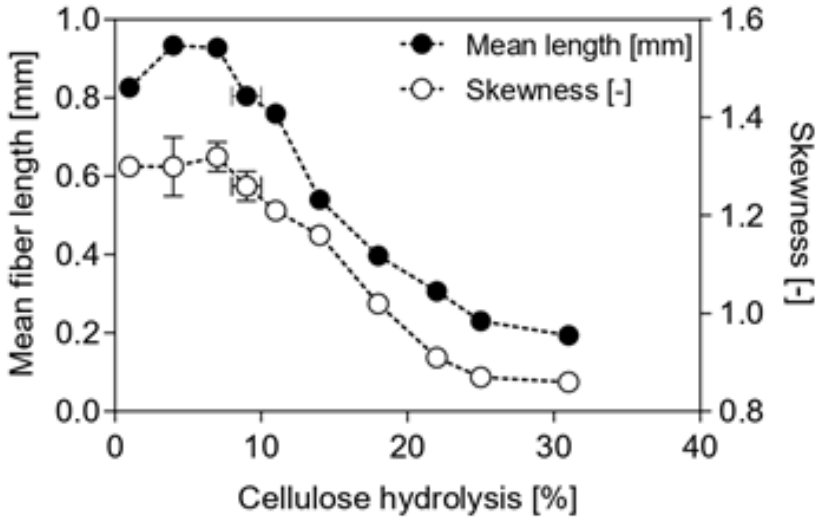


Figure 24: Monitoring of Kraft fibers size during the hydrolysis.

5.3.6 Quantifying changes in the cellulose structures during hydrolysis: image analysis, CBM depletion assay and ^{13}C NMR

When the crystallinity of the pulps was assessed using ^{13}C NMR, a crystallinity index (CrI) of 0.52 was calculated from the NMR spectra according to Park et al. 2010 (**Figure 25**). In addition, the accessibility of the NBSK substrates was assessed using the CBM adsorption assay (**Table 7**). One of the advantages of the CBM/CLSM method over the CBM adsorption and the NMR analysis is that it can provide localized analysis of the fiber of interest (e.g. excluding fines), as well as spatial distribution of the cellulose morphologies within the fiber.

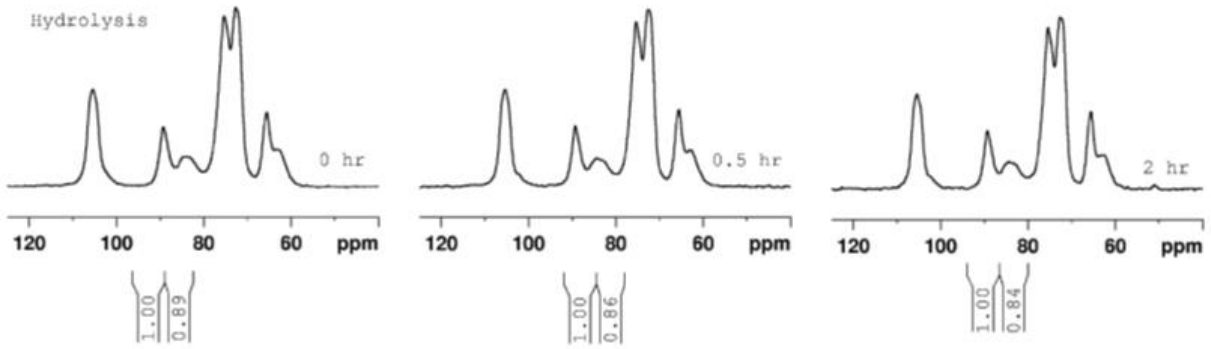


Figure 25: ^{13}C NMR analysis of NBSK fiber hydrolyzed for 0, 0.5, and 2h. After 2 h of hydrolysis, the changes occurring in the substrate drastically changes its size and supramolecular organization. The ^{13}C NMR spectra will reflect these changes, making it impossible to attribute changes in the CI to the enzymatic deconstruction mechanism.

Table 7: The change of CBM17 and CBM2a adsorption on bleached Kraft fibers over hydrolysis time

Time [h]	CBM2a [$mg_{Protein}\ g_{cellulose}^{-1}$]	CBM17 [$mg_{Protein}\ g_{cellulose}^{-1}$]
0	24.9 ± 0.5	2.43 ± 0.04
0.5	21.6 ± 1.5	2.33 ± 0.01
1	21.0 ± 2.0	1.94 ± 0.01
2	18.2 ± 1.3	1.55 ± 0.02
6	15.2 ± 0.9	1.55 ± 0.02

Depicted are mean values and the standard deviation from triplicate experiments.

The most likely reason for the increase in the R/G ratio (**Figure 25**) was due to a loss in CBM17-FITC binding with the CBM adsorption data in **Table 7** also showing a faster decline in CBM17 binding as compared to the CBM2a binding. As an increase in the R/G ratio implied an increase

in surface crystallinity, ^{13}C NMR analysis was used to assess the crystallinity index of the various samples. After 2 h of hydrolysis, the crystallinity had increased from 0.529 to 0.543, which is a minimal change that may support the increased crystallinity observed.

5.3.7 Qualitative analysis of the fragmentation with SEM and CLSM

When CLSM (**Figure 26**) and SEM (**Figure 27**) were performed at higher magnification to investigate localized morphological changes that resulted from the initial phase of digestion, with increasing hydrolysis time these "cracks" appeared to increase in size, resulting in fiber shearing and breakage. Similar observations could be interpreted from the SEM images (**Figure 27**). It appears that the initial "nicks" had developed into larger cracks (**Figure 27a** to **27c**), leading to the fragmentation and disruption of the fiber.

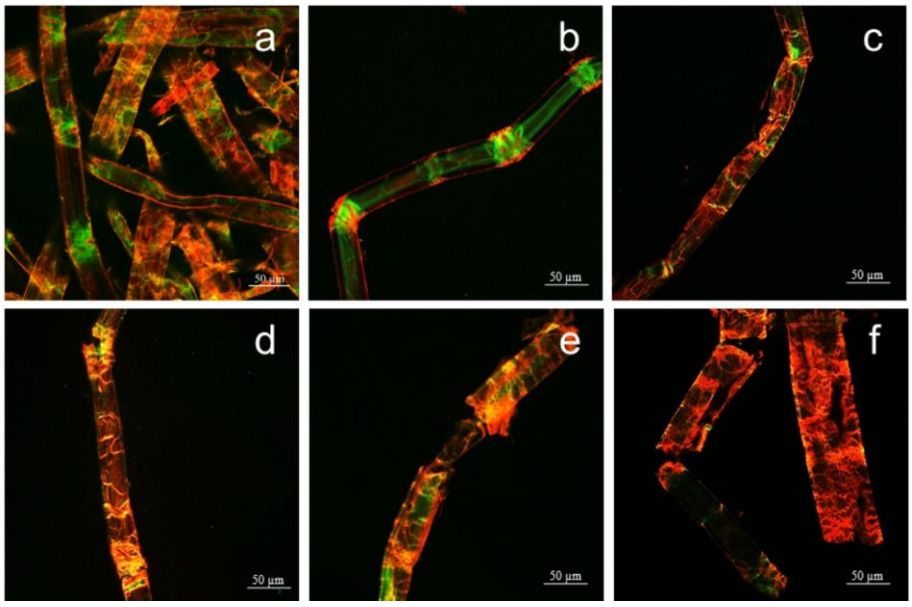


Figure 26: Visualization of fiber breakage of Northern Bleached Softwood Kraft (NBSK) pulp fibers hydrolyzed for 0 h (a), 2 h (b-d), and 4 h (e-f) using CBM17-FITC and CBM2a-RRedX binding with CLSM (Scale bars represent 50 µm).

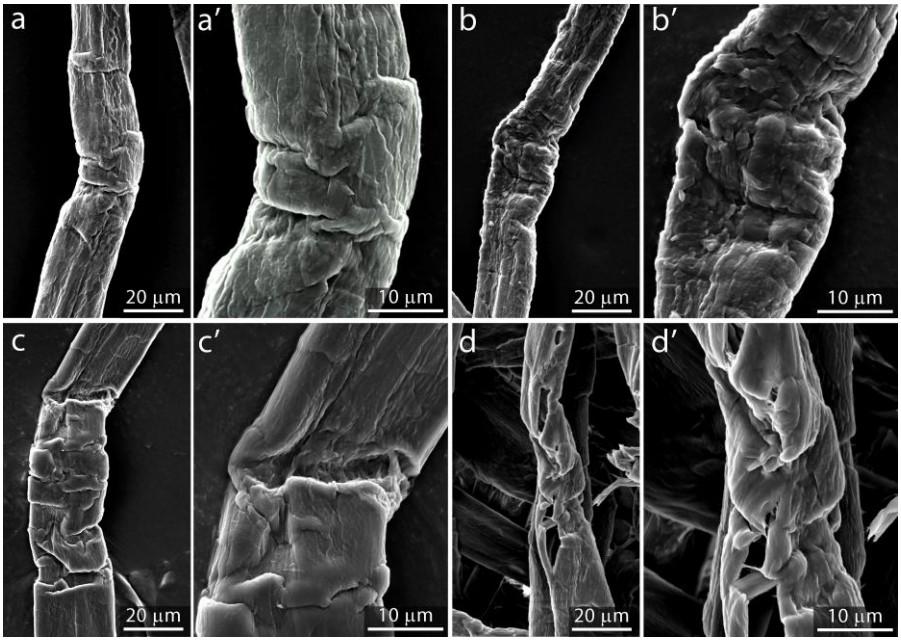


Figure 27: Visualization (SEM) of disorganized zones after 0.5 h (a, a'), 1 h (b, b'), 2 h (c, c'), and 4 h (d, d') of incubation. Magnification of 2000x (a, b, c, d) and 5000x (a′, b′, c′, d′).

Previous work that looked at the thermodynamics involved in enzymatic hydrolysis showed that the energy required for degradation of crystalline cellulose is much higher than the energy needed to hydrolyze a single glucan chain within paracrystalline cellulose (Beckham et al., 2011). Thus, it would appear consistent that the disorganized zones were more quickly hydrolyzed, resulting in the observed fiber fragmentation.

In consideration of the results observed, further investigation on the organization of the microfibrils and changes on their surface (e.g. number of chain ends) is still required to understand the extent of their influence to the fiber morphology. In the previous section, we showed that apparent fibrillation does not always lead to an increase of cellulose accessibility, which was reflected by both the enzymatic hydrolysis and the CBM adsorption. Here, we showed that the simultaneous hydrolysis of crystalline and paracrystalline region in highly accessible zone led to fiber fragmentation, leaving particles with higher surface organization. Although CBM probes have demonstrated great potential to help us understand lignocellulose substrates and their enzymatic degradation, the multifactorial nature of their structures requires a careful interpretation of the results.

5.3.8 Conclusions

Fluorescence-tagged CBMs were successfully used with CLSM and quantitative image analysis to show that, during initial hydrolysis, enzyme-mediated deconstruction predominated at localized regions of cellulose showing lower degrees of organization. In these highly accessible regions, the concentration of the green fluorescence (caused by pronounced CBM17-FITC binding) suggested

readily available paracrystalline structures. When the ^{13}C NMR, SEM and fiber characterization were combined with the CBM adsorption and CBM/CLSM data, it was apparent that fiber fragmentation and the preferential hydrolysis of the more disorganized fiber regions occurred at the same time. This work helps elucidate the nature of these zones; indeed, it was shown that dislocation zones displayed high binding capacity for both types of CBMs. In these regions, paracrystalline and crystalline nano-structures can undergo rapid degradation, which in turns lead to fragmentation of the fiber.

5.4 The potential of CBMs to functionalize cellulose surfaces

While the work described in the previous sections looked at the use of CBMs for imaging and characterization of carbohydrate structures, the work described in the next thesis section explored the potential of CBM to functionalize cellulose surfaces. As mentioned in the introduction, previous work mainly modified the chemical and physical properties of cellulose materials using CBM-protein fusion constructs. Notably, CBM-CBM construct was used to crosslink cellulose fibers, increasing the strength of papers (Levy et al., 2002). In other work, CBM-hydrophobin constructs were used to improve the compatibility of composite material, such as graphene-cellulose (Laaksonen et al., 2011). Another research topic has looked at cellulose functionalization using CBM for the development of biosensors (Pelton, 2009).

5.4.1 Challenges in cellulose surface modification

Nano-scale cellulose materials such as nanofibrillated cellulose or bacterial cellulose, are promising candidates for the development of novel renewable products (e.g. biosensors) (Credou and Berthelot, 2014). These new high-value materials are the subjects of ongoing research and should be commercially interesting for the pulp and paper and agricultural sectors. However, one of the major bottlenecks for the applications of this new material development is the necessity to dry the cellulose and use organic solvent to modify its surface. However, carbohydrate binding modules (CBMs) have the potential resolve some of these drawbacks.

5.4.1.1 Reactivity of cellulose

The reactivity of cellulose hydroxyl groups varies according to the reaction medium in which functionalization is done. For example, the order of reactivity for etherification performed in an alkaline medium is OH-2 > OH-6 > OH-3 while the primary hydroxyl group (OH-6) is the most active in esterification (Varshney and Naithani, 2011). The type, distribution, and uniformity of substituent groups determine the properties of derivatives. The average number of hydroxyl groups replaced by the substituents is the degree of substitution (DS). Chemical functionalization of cellulose includes reactions of hydroxyl groups such as esterification, etherification, intermolecular crosslinking reactions, and free radical reactions, particularly in the formation of graft cellulose copolymers to increase the applications of cellulose by altering its properties (Missoum et al., 2013). Cellulose modification can be performed via covalent link or adsorption, both have advantages and inconveniences.

5.4.1.2 Covalent modification of cellulose

As noted previously, chemical grafting processes have to be mild to preserve the integrity of the nanoparticle (Klemm et al., 2018). Furthermore, as mentioned earlier, the surface modification of cellulose nanoparticles involves the surface hydroxyl groups. The most common surface chemical modifications can be categorized into two groups: (1) substitution of hydroxyl groups with small molecules, (2) polymer grafting. Although it allows an improvement in the dispersion of the modified nanoparticles, it inevitably restricts the interactions between nanoparticles through hydrogen-bonding which is the basis of the outstanding mechanical properties of nanocellulosic

based nanocomposites. Although covalent modification of cellulose surface lead to more stable nanocomposite, the processes required are costly and rarely environmentally friendly. If covalent modification seems mandatory for bioplastics, it seems that a better way can be found for other application, especially in aqueous media (e.g. biomedical application).

5.4.1.3 Non-covalent modification of cellulose

The surface of cellulose nanoparticles can be "tuned" using surfactants or polyelectrolyte adsorption. Surfactants are usually amphiphilic organic compounds (*i.e.* containing both hydrophobic groups and hydrophilic groups). For example, the anionic surface of TEMPO-NFC can be modified with a cationic surfactant. Often the adsorbed layer is used to control the hydrophobicity of cellulose nanofibrils. Lignin and tannic acid adsorption has been shown to enhance cellulose surface hydrophobicity (Hu et al., 2016b). Another way to modify surface properties of NFC is to use a polyelectrolyte solution (PEI). Researchers produced a polyelectrolyte multilayer (PEM) using different polyelectrolytes to improve biocomposite performances (Missoum et al., 2013). The combination of PEI and NFC in deionized water results in the formation of regular layers of NFC and PEI. In this work, we combined the best of two worlds by modifying the cellulose surface with specific and irreversible non-covalent adsorption of functionalized CBMs.

5.4.2 The use of CBMs to functionalise cellulose surfaces

In this application the cellulosic material was just used as the support for a detection kit as CBM-antigens or CBM-antibodies are fixed on the paper to improve sensor sensitivity (Rosa et al., 2014). The fusion approach restricts the cellulose modification only to polypeptides, which may be too restrictive for the different application possible. For example, polymeric grafting on cellulose nanocrystals (CNCs) have been shown to increase their stability as well as their dispersion (Araki et al., 2001).

With the bioconjugation approach, CBMs can be used as a versatile platform for cellulose modification. CBMs often present a free amine at the end of their amino acid sequence, which does not affect the binding of the CBMs (Filonova et al., 2007). Furthermore, it is readily accessible for nucleophile substitution.

Commonly, proteins are modified through their terminal amine or lysine residue. An amine-coupling process can be used to conjugate with nearly all protein or peptide molecules as well as a host of other macromolecules. The primary coupling reactions for modification of amines proceed by one of two routes of acylation or alkylation. Most of these reactions are rapid and occur at high yield to give stable amide or secondary amine bonds. For example, N-hydrosuccinimide esters are one of the most common activation chemistries for creating reactive acylating agents. NHS esters may also be formed in situ to react immediately with target molecules in aqueous reaction media. NHS esters have a half-life on the order of hours under physiological pH conditions. However, hydrolysis and amine reactivity both increase with increasing pH. At 0 °C at pH 7.0, the half-life is typically 4 to 5 h., which may be problematic for long experiment times.

In addition, a high protein concentration is often required for NHS bioconjugation. Thus, it is challenging to achieve this in situ (Hermanson, 2013a).

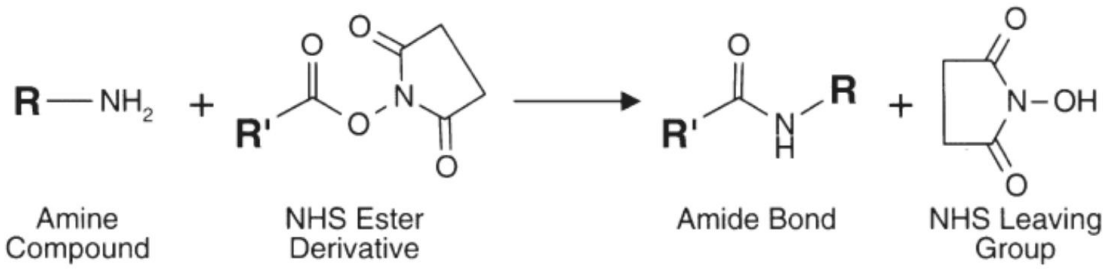


Figure 28: Nucleophile substitution of NHS ester with amine group yielding an amide covalent bond and NHS leaving group

Click chemistry has been used to develop powerful, selective, and modular building blocks, such as azide and alkyne, and this approach can be used at both a small and large scale. It is worth noting that the unactivated azide-alkyne cycloaddition was first discovered by Huisgen in 1963, but was only updated by Kolb et al. in 2009 by employing Cu(I) as a catalyst (Kolb et al., 2009). These authors indicated that the Click reaction must be "modular, wide in scope, give very high yields, generate only inoffensive by-products that can be removed by non-chromatographic methods.'' The starting materials and reagents should also be readily available, and the reaction should proceed under "friendly" reaction conditions, such as room temperature and a benign solvent (i.e., water).

With the rise of cellulose nanomaterials and the increasing interest on carbohydrate-based biomedical compounds, new ways to introduce functionality on the otherwise "inert" cellulose surface would be beneficial (i.e. in water). In the work reported here we opted for a two-step NHS-

Click strategy, which allowed the efficient, specific, as well as versatile functionalization of cellulose. As mentioned earlier, cellulose nanocrystals (CNCs) are being increasingly used in various fields due to their unique characteristics such as being renewable and having a high specific surface area, high aspect ratio, low coefficient of thermal expansion, high thermal conductivity and outstanding mechanical properties such as high strength and stiffness (Park et al., 2013). CNCsuspensions have unique characteristics such as forming nematic liquid crystalline domains and are known to stabilize emulsions due to the pickering effect (Zoppe et al., 2012). Aqueous utilization of CNCs have shown potential in various fields of application such as medicine, food and cosmetics (Nascimento et al., 2018). Although CNCs have shown great potential, surface modification can further enhance their utilization (Tang et al., 2017). For example, grafting helps redisperse agglomerated particles, *e.g.* grafting polymer chains (Khoshkava and Kamal, 2014). This is an important characteristic as it is not easy to redisperse agglomerated CNCs, resulting from the strong hydrogen bonding of the crystals. Grafting can also help reorganize the CNC network within gels, consequently improving gel properties (Way et al., 2012). However, to date, most cellulose surface modification methods that have been used require drying/solvent exchange of the cellulose and the use of organic solvents which increases process costs and has a negative impact on the environment (Missoum et al., 2013; Yoo and Youngblood, 2016). Although other methods of functionalizing the cellulose surface in aqueous media have been reported, (such as the use of oxidized cellulose materials (Zhang et al., 2013b)), these approaches have tended to result in unspecific adsorption that lead to uneven surface modification or required strong alkaline conditions (Kloser and Gray, 2010; Missoum et al., 2013).

Thus, we hoped to utilize the strong and specific interaction of CBMs with cellulose to bring new functionality to the cellulose surfaces using chemically modified CBMs to "activate" nanocellulosic materials such as cellulose nanocrystals (CNCs). Although fusion CBMs have been successfully used in the past (Shoseyov et al., 2006), bioconjugation offers several advantages such as versatility, diversity and a tailored modifying agent. In addition to Click chemistry, it is then possible to graft cellulose surfaces in water without prior modification of the cellulosic material. As a proof of concept, polyethelene glycol (PEG), a water soluble polymer often used for its biocompatibility, was grafted to the CBM, as it has been shown to have various positive effects of CNC suspension (Araki et al., 2001).

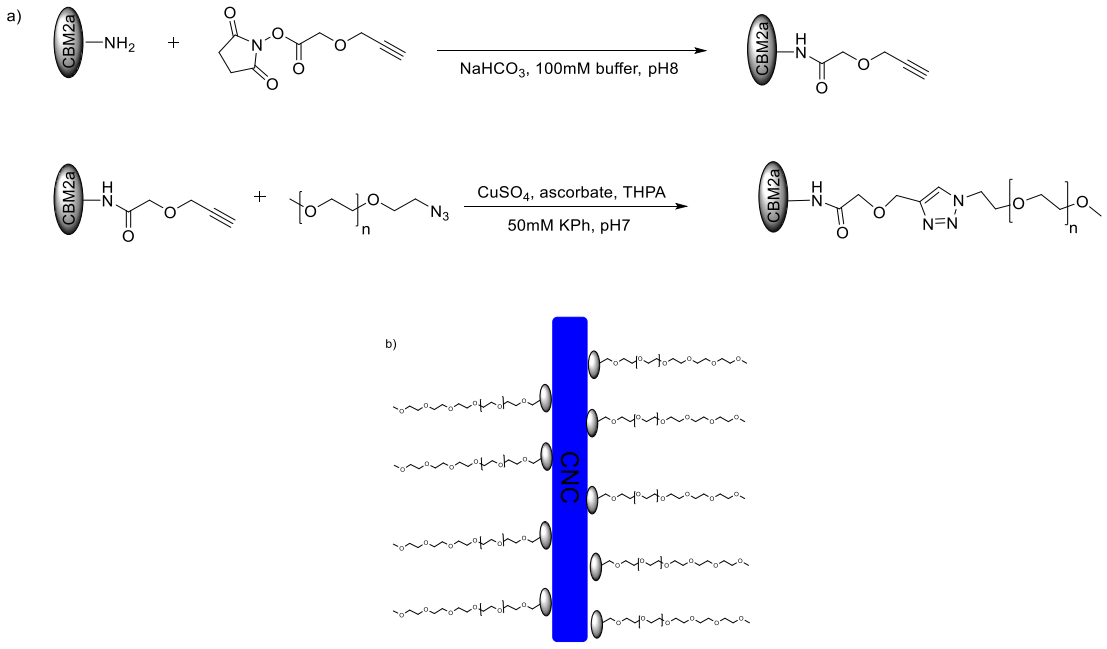


Figure 29: a) Chemical reactions involved in the two steps CBM2a functionalization. First, NHS-ester bioconjugation in sodium bicarbonate buffer followed by the Click reaction between alkyne-CBM and azide PEG in phosphate buffer b) schematic representation of a cellulose nanocrystals (CNC) covered with CBM-PEG.

5.4.3 Bioconjugation of CBM2a with NHS-ester

One of most versatile ways of functionalizing peptides and proteins involves the use of chemical groups that react with primary amines ($–NH_2$) (Hermanson, 2013b). As a first step in CBMs functionalization, alkyne groups were introduced at the terminal amine of the amino acid sequence with propargyl NHS-ester with optimized conditions resulting in yields close to 90% after purification. It was apparent that CBM concentration was a key parameter with concentrations of 10 mg CBM.mL^{-1}, resulting in highest yields and reaction rates. As well as CBM concentration a pH of 8.3 resulted in optimum bioconjugation yields **(Figure 30.b)**. At these conditions it appeared that the amine group on the CBM was fully available for the nucleophilic attack on the carbonyl of the propargyl NHS-ester **(Figure 30a)**.

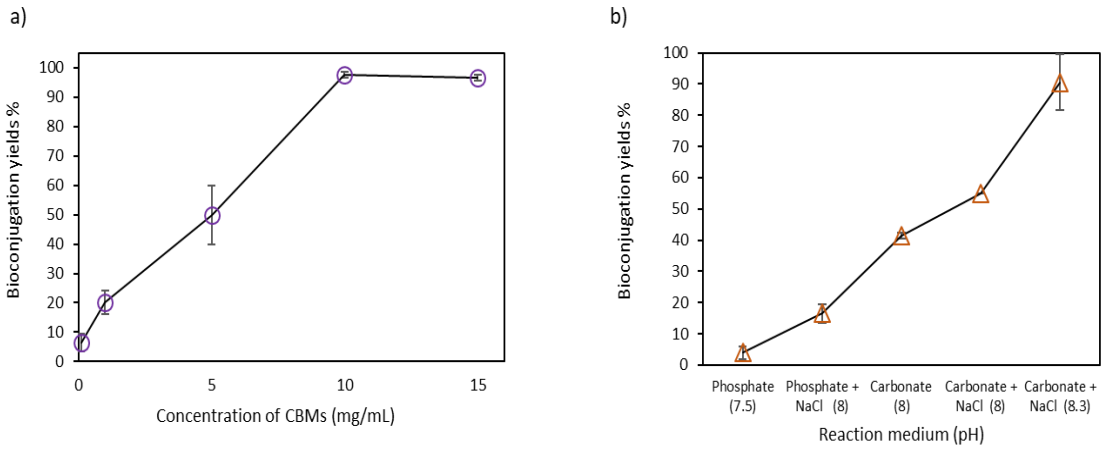


Figure 30: Optimization of CBM2a bioconjugation with NHS-propargyl (a) represent the yields increase in function of the protein concentration (b) represent the reaction optimization function of different pH and salt concentrations. Phosphate and carbonate buffer concentration used are 100mM. Sodium chloride (NaCl) concentration used is 100mM.

5.4.4 Click reaction

As mentioned previously, Click chemistry is a class of biocompatible small molecule reactions commonly used in bioconjugation, allowing the joining of specific substrates with specific biomolecules. The Click reaction has been used in a variety of areas such as pharmacological and various biomimetic applications (Kolb et al., 2009; Pahimanolis et al., 2017). Following the functionalization of CBMs through NHS-ester alkynes, reaction completion was indicated by the disappearance of the CBM-alkyne band on polyacrylamide gels (**Figure 31**) and a 95% yield, after purification. The water-soluble copper ligand, THPA was next used to catalyze the Click reaction without the addition of co-solvent using nanomolar concentrations of catalysts (Presolski et al., 2011). The CBM-PEG were purified using an FPLC column. The binding constant (Ka = 3.3 10^6 M^{-1}) was close to the unconjugated CBM2a constant (Ka = 3.6 x 10^6 M^{-1})(McLean et al., 2000). The strong affinity of the CBM2a to cellulose implied it was irreversible in water (Boraston et al., 2001), and the PEG was subsequently grafted at a 1:1 molar ratio to the CBM.

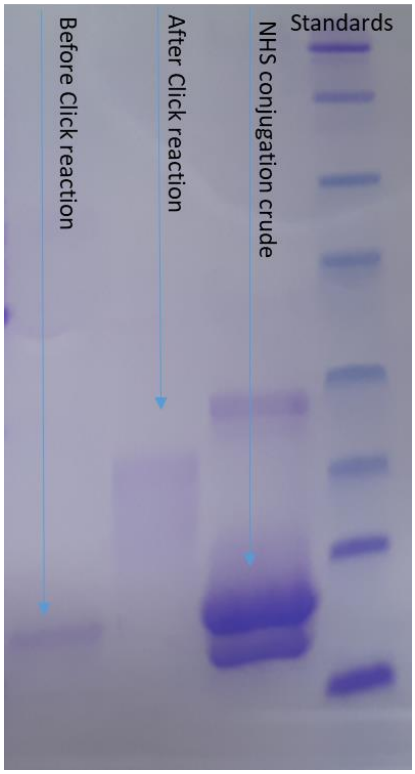
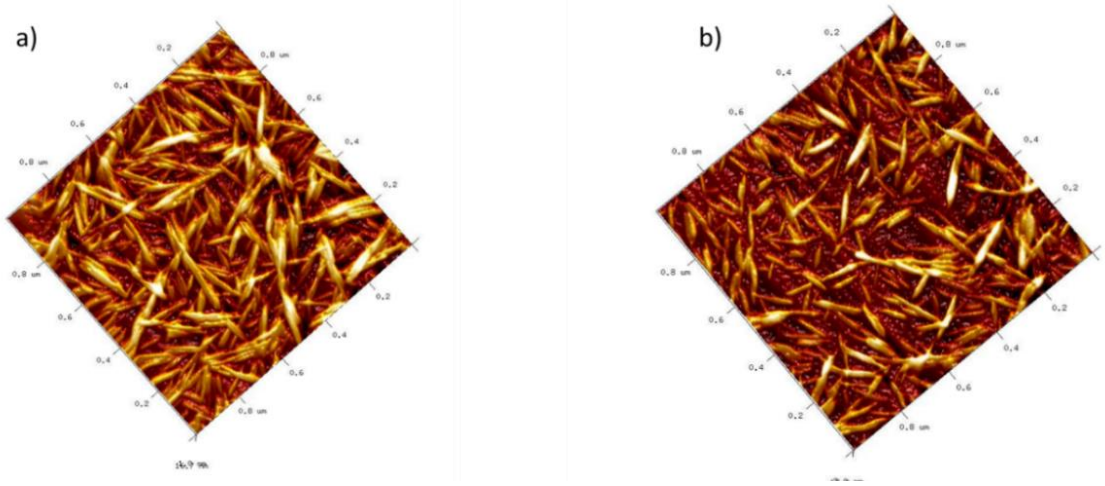


Figure 31: SDS-Page picture of carbohydrate-binding modules conjugation. The polyacrylamide gel was performed on precast gel (4–12% Criterion™ XT Bis-Tris Protein Gel, 18 well, 30 µl, Bio Rad, Hercules, CA, USA) using XT-MES buffer, XT reducing agent (Bio Rad) The gel was run at 200V, 2A for 40min. Proteins where revealed using brilliant Coomassie blue (Bio Rad) and destaining solution.

5.4.5 CBM interaction with CNC

In this work, the CBMs were added at concentration of 1.4 mmol.g^{-1} of CNCs with the CBMs likely binding to the most hydrophobic faces (Lehtiö et al., 2003). Atomic force spectroscopy was next used to assess possible CBM mediated changes at cellulose surfaces and interfaces (Eaton et

al., 2000). The respective original and surface modified CNC were shown to have overall dimensions of 152 ±12 (length), 5.1 ±0.2 (width) and 146 ±13 (length), 5.6 ±0.4 (width). This slight change appeared to confirm that functionalization had not dramatically affected the size or shape of the nanocrystals (**Table 8**). However, the repulsive interactions or steric stabilization between the CBM functionalized CNCs appeared to lead to a more dispersed CNCs sample when preparing samples utilizing the same conditions (**Figure 32a** vs **32b**). It is important to note that, to obtain homogenous sample preparation, the mica was first coated with the cationic polymer APTES. Therefore, little or no aggregation was observed in either sample. However, in the absence of polymer coating, the original crystals formed bundle like aggregates. This phenomenon was greatly reduced when the CNCs was coated PEG-CBM.



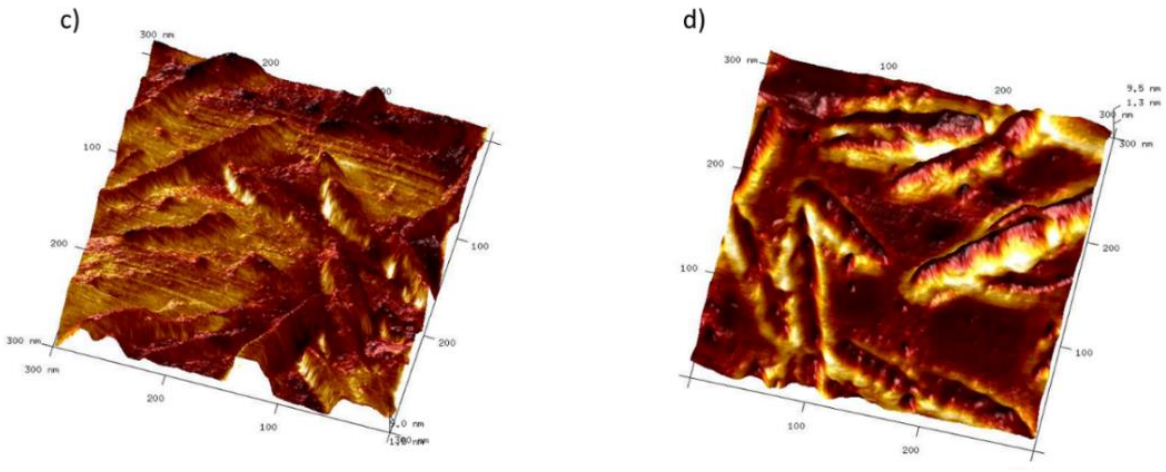


Figure 32: AFM images of CNC used to determine particle size distribution on height profile. a) control, b) modified sample with CBM-PEG, adhesion maps of individual CNC where the adhesion profiles were superimposed onto the 3D height map of the control c) and modified d) samples.

The molecular length of the PEG (2000 g.mol^{-1}) was estimated to be between 10-15 nm for its full contour length. However, the PEG "tail" was likely somewhat rearranged to a more folded, coil conformation (Araki et al., 2001). When the CBM-PEG functionalized CNC were assessed via the adhesion channel of the AFM differences in adhesion forces (nN) were observed between the samples reported in **Table 8** and visualized in the 3-D map in **Figure 32c** and **32d** (Peukert and Gotzinger, 2003). In the AFM pictures, it was apparent that the outer surfaces of the CNC had changed significantly, with a lower adhesion around the crystal, corresponding to a modified environment of protein and PEG. This contrasted with the relatively homogenous control.

Table 8. CNC dimension distribution analyzed by AFM and adhesion measurements

	Length (nm)	Width (nm)	Sample adhesion (nN)*	CNC adhesion (nN)
Conjugated CBM/CNC	146 ±13	5.6 ±0.5	2.12 ±0.3	0.5 ±0.2
Control	152 ±12	5.1 ±0.2	1.68 ±0.4	1.9±0.6

*Average adhesion of the whole sample, calculated with software. Other value were measured on individual CNC using n=20, with at least 3 different images (1*1µm).

In higher concentration of CNCs distinct patterns were observed. It is likely that the observed orientation of the CNCs was due to sample preparation (**Figure 33**), grafting or a combination of these influences. The changes to the CNC surfaces were confirmed by the FTIR and AFM results.

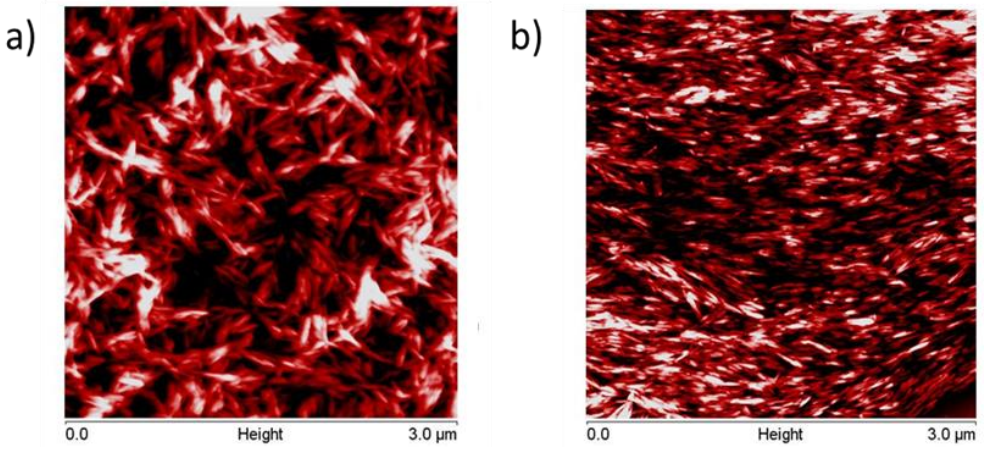


Figure 33: AFM images of CNCs CBM-PEG grafted, height only, of the control (a) and modified (b). Apparent orientation of the CNC network in high concentration during sample preparation. Such pattern is occurring sporadically in the CBM-PEG grafted CNC. It only occurs in high concentration of CNC, typically over 0.2% wt.

5.4.6 Characterization of CNC with FT-IR

The CNC was analyzed with FTIR to determine the potential surface changes to the cellulose surfaces (Li and Renneckar, 2011). The absorption peaks for the stretching vibrations at 3300, 2900 and 1000 corresponded respectively to the O-H, C-H and C-O bonds. The spectrum of CBM-PEG grafted sample (**Figure 34**, grey line) resulted in new absorptions at 1600 cm^{-1}, corresponding to amide stretching of the CBM peptide bond (**arrow 2**). Although the PEG showed characteristic peaks between 750-1500 cm^{-1} (Araki et al., 2001), corresponding to C-O bonds, only the stretching absorption was apparent under the CNC spectra (**arrow 3**). At 2850 cm^{-1} a more defined C-H stretch appeared, likely attributed to the PEG (**arrow 1**).

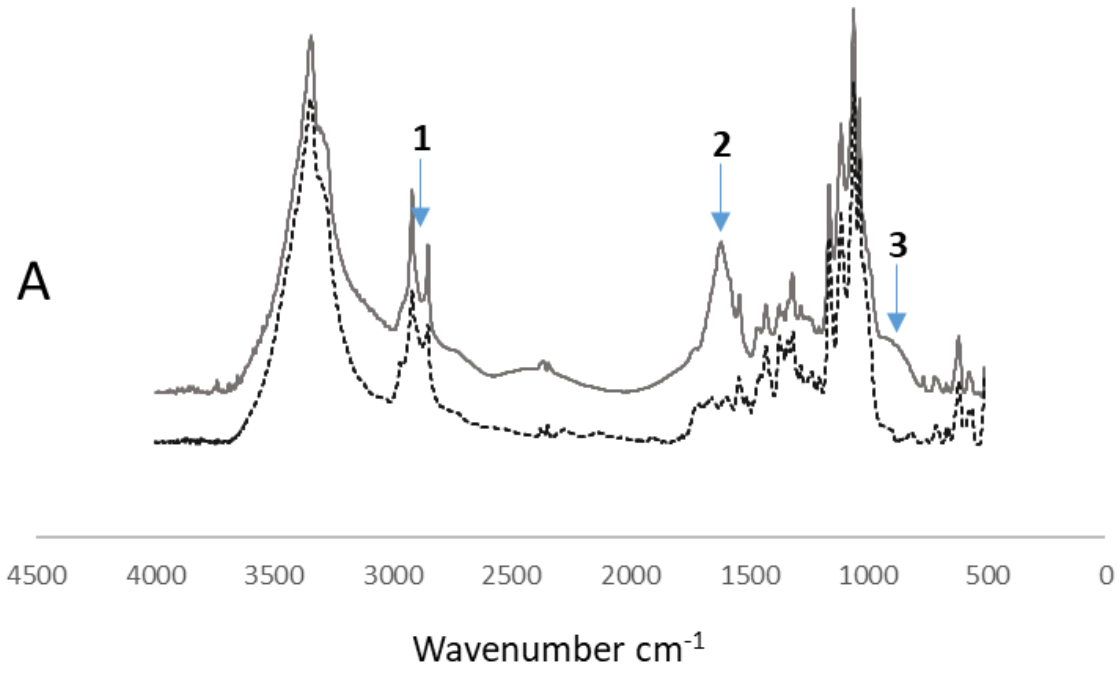


Figure 34: FTIR spectra of ungrafted CNC (black dot) and grafted with CBM-PEG (grey line)

5.4.7 Effect of the grafting on the CNC suspension

Grafted CNC suspensions were shown to be stable for several weeks and no precipitation of the colloidal nanoparticles was observed (**Figure 35, Samples 1 and 2**). When the optical density at 600 nm was measured, (OD 600 = 0.6) no precipitation was apparent for these samples. Although drying of CNCs will likely be required for material transport and storage, it is recognised that drying is problematic as it commonly leads to irreversible aggregation (**Figure 35**). However, after freeze drying, the grafted CNCs were redispersed in nanopure water without any need for sonication, which was not the case with the control (**Sample 4, Figure 35**) of the combined CNC, PEG and CBM without the Click reaction. It was apparent that, during the second freeze-drying stage, the surface modified CNC's were not redispersed while the control and samples were unchanged after two stages freeze drying (**Sample 3**). This was likely due to the refolding of CBMs in nanopure water. It should be noted that, likely because of reduced surface modification, a certain amount of CNC became aggregated and sedimented compared to the fully surface modified sample (**Sample 2**).

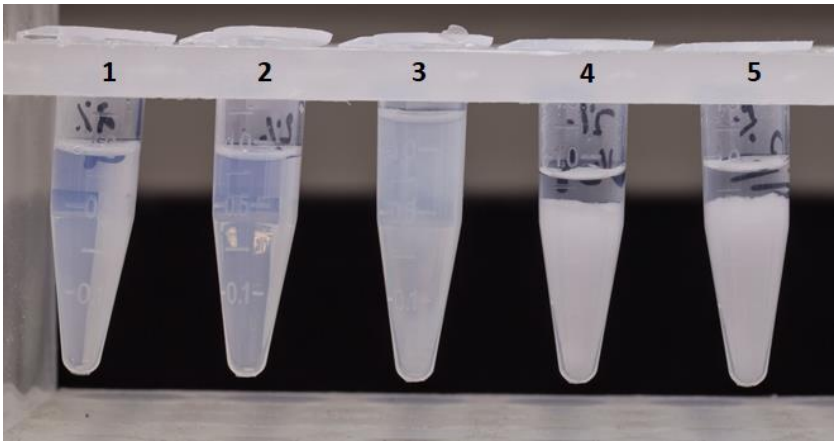


Figure 35: 1) 2% CNC 2) CNC+CBMs-PEG after first freeze drying, 3) CNC+CBMs-PEG after second freeze drying, 4) CNC after 1 freeze drying, 5) CNC after 2 freeze drying with PEG

This initial result suggested that the CNC-protein conjugated nanoparticles have considerable potential for building multifunctional nano-biomaterials via controlled assembly techniques with hybrid protein–polysaccharide nanoparticles likely leading to novel self assembly nanomaterials. The unique properties of cellulose nanomaterials in conjunction with the biorecognition abilities of CBMs offer particularly exciting opportunities in molecular imaging, biosensors, and biomolecule delivery (Kang et al., 2015; Phelps et al., 1994; Shoseyov et al., 2006; Teeri et al., 2007). For instance, nano-carriers which encapsulate doxorubicin and methotrexate and used in physiologic conditions are another class of novel nanomaterials with promise in a variety of biomedical applications including whole-blood immunoassays, cancer therapy (Rahimi et al., 2017).

It was apparent that bioconjugated alkyne CBM2a could be successfully synthesized in high yield (90%) and PEG, CBM and CNC could be combined in a one-step reaction to modify the cellulose surface. The surface functionalization of CNC with CBM-PEG resulted in aqueous suspensions that were stabilized by steric effects and prevented aggregation of nanoparticles during drying. Subsequent AFM imaging, where adhesion maps are superposed onto the height profile, showed an apparent modification of the nanocrystal surface with CBM-PEG addition improving CNC redispersion in water. It was also apparent that the PEG-grafted CNC redispersion was influenced by the CBM concentration, with a minimum loading of 25 $\mu mol.g^{-1}$ required for total redispersion. These initial results suggest that CNC-protein conjugated nanoparticles have considerable potential for building multifunctional nano-biomaterials.

5.4.8 Immobilization of liposomes on cellulose surfaces (ongoing)

Following the "proof of concept" by developing the Click-CBM, the possible immobilization of liposomes on cellulosic surfaces was next assessed as liposomes and lipid layers are of ongoing interest to the biosensor area. As biological samples (e.g. blood) are rich in physiological information, it was beneficial to be able to analyze different analytes simultaneously. The amphiphilic nature of lipids allows the incorporation of water-insoluble marker molecules in aqueous environment and, once triggered, the marker can be released from the liposomes and interact with the sample.

Alkyne-CBM, azide functionalized liposomes and cellulose were added simultaneously then vigorously stirred four 1 hour. The cellulose was centrifuged and washed 3 times to remove unbound liposomes. As fluorescein was incorporated during the making of the liposomes this allowed the imaging of the liposomes and cellulose with CLSM (Figure 36).

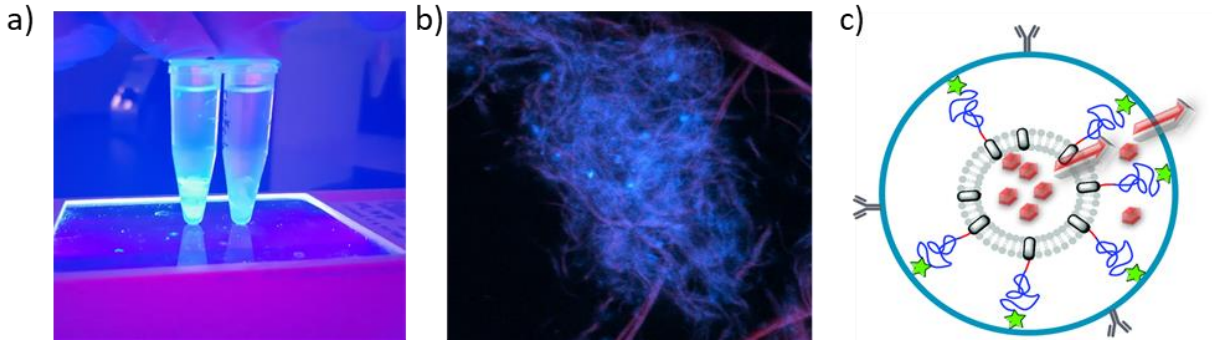


Figure 36: a) Fluorescein encapsulated in liposome, b) Imaging of liposomes immobilized on microfibrillated MFC, c) Schematic representation of an encapsulated and functionalized liposome.

Figure 36 a and **b** showed that liposomes can be effectively grafted onto cellulose filaments using CBMs. The fluorescence of the FITC can be revealed with UV excitation and under confocal microscopy and CLSM imaging showed that a high concentration of liposomes is embedded into the microfibrillated cellulose network. This provided an apparent stabilization of liposomes at 4 ºC for several weeks.

Although this work is preliminary, it showed considerable potential and opens doors for future research. For example, cellulose conjugated CBM nanoparticles have considerable potential for building multifunctional nanobiomaterials via controlled assembly techniques. The specificity and affinity of CBMs can help functionalize cellulose surfaces, immobilized structures and desired compounds. Depending on the application of interest, the choice of CBM can be tailored to meet the desired requirements (e.g. elution property, substrate affinity, etc.).

Chapter 6: Conclusions and future work

6.1 Conclusions

The focus of the thesis work was to investigate the potential of using carbohydrate-binding modules to assess accessibility and monitor structural changes during enzyme mediated hydrolysis and modification of pulps. In this way we hoped to enhance our understanding of the mechanisms involved in enzyme modification with a focus on how the cellulosic materials limited effective hydrolysis. The lack of methods of assessing "surface" changes, as well as the absence of specificity for cellulose substrate has resulted in some controversies in the bioconversion field. A major goal of the thesis work was to develop a method that would complement existing techniques to characterize glycostructures.

As a good analytical method should be readily available, we optimized the production of the CBMs probes addressing potential challenges such as oxygen depletion, long induction times, etc., to yield higher protein production. Consequently, time saving, and lower operating costs were achieved. Although the purifications of the CBMs can still be further optimized, especially with the addition of histidine-tags, the cellulose-CBM purification step was enhanced

The analysis of the differential binding of type A and B CBMs was performed on different model substrates resulting in the depletion assay method being highly reproducible and accurate across a range of substrates. This method was subsequently used to quantify cellulose accessibility while comparing it to other established methods (WRV, Simons stain, NMR). CBM adsorption was used

where type A (crystalline cellulose) and a type B (paracrystalline) CBMs were used in parallel with microscopy, fiber analysis and water retention values to evaluate accessibility of microfibrillated cellulose. Unlike other methods, the results showed that CBM adsorption followed the same trend as the hydrolysis profile, suggesting that refining results in a dispersion of the microfibrils without affecting enzyme accessibility at the microfibril level. When, depletion assays were applied to more realistic lignocellulosic substrate challenges were encountered. The heterogeneity of the substrates as well as the lignin hindrance were challenging for standard depletion assays methodology. That is why we developed a more refined method to overcome these challenges.

The quantification of accessible cellulose and cellulose substructures was further refined using fluorescent markers attached to CBM probes. Fluorescence-tagged CBMs were used to differentiate the cellulose substructures after pretreatments and during enzyme-mediated hydrolysis. The differences in CBM adsorption were imaged using confocal laser scanning microscopy and quantified via the relative fluorescence intensities of the respective probes. The quantitative image analysis was supported by other methods such as ^{13}C NMR, SEM imaging, and fiber length analysis. The results suggested that, at the initial stages of hydrolysis, enzymatic activity is more pronounced in high accessibility region (disorganised zones) leading to enhanced fiber fragmentation and an increase in cellulose surface crystallinity. The use of fluorescence tagged-CBMs with specific recognition sites provided a quantitative way to elucidate cellulose structures and their impact on enzyme accessibility, which in turn provided novel insights into the mechanisms involved in cellulose deconstruction.

As noted earlier, chemical functionalization can improve the overall utilization of cellulosic polymers. Although a wide variety of reactions for cellulose modification have been studied (Cunha and Gandini, 2010) the common use of environmentally problematic solvents and high costs have limited their application. In the work reported in the thesis, CBM2a (crystalline cellulose) was functionalized with NHS-alkyne making use of the terminal amine. Following this bioconjugation, Click reactions with polyethylene glycol (PEG) were conducted to modify CNC surfaces. This provided a strong and non-covalent modification of cellulose surfaces with the added advantage that this provided a "one-pot reaction" in aqueous media. The CBM-PEG modification of cellulose surfaces showed to increase CNC dispersion after drying, consequently improving suspension stability. It was apparent that polysaccharide–protein hybrid and self-assembled nanoparticles could be effectively produced with the CBM providing a versatile vector for cellulose functionalization.

6.2 Possible future work

6.2.1 CBM production and purification – fusion constructs

Although high protein concentration was obtained, new hosts as well as high performance vectors could be studied to further enhance CBM expression. This would provide a wider range of CBMs probes with an extended CBM library used to recognize, image and characterize a greater number of substrates. The CBM purification could be further optimized by using cellulose-based matrix

and liquid chromatography. Ideally a more controllable mild elution would enhance the whole process. Another way would be to add a histidine tag to all the protein.

6.2.2 CBM probes

It is possible that the diversity of CBM specificity could be used to provide detailed "maps" of the surface morphology and surface composition of lignocellulosic fibres resulting from various thermochemical or enzymatic processes. For example, hemicellulose specific CBMs could be used to track the localization and redistribution of these polymers after pretreatment processes or during hemicellulases and enzyme cocktail action. Those CBMs specific for particular hemicelluloses could be used to track the amount of each hemicellulose, in similar ways to what is described in **section 5.3**. This work could provide insights into the variable recalcitrance of particular hemicelluloses to enzymatic hydrolysis. By mapping the fibre surface, it should be possible to develop a more detailed understanding of the changes occurring to the various carbohydrate polymers within lignocellulosic biomass during thermochemical and enzymatic processes. These insights could be used to optimize enzyme cocktails towards either targeted modification (for example, increasing fibrillation while trying to maintain strength) or deconstruction (e.g. what accessory enzyme should be added for which substrate).

6.2.3 CBM used as "anchors" for cellulose functionality, immobilization and assembly

The self-assembly properties of different polysaccharides, as well as the increasing interest of the scientific community and industry to develop new biosensor and diagnostics methods could enable

the construction of 2D and 3D molecular crossroads. In this way different CBMs could be used to carry and transfer molecular functionality, polymers or nanoparticles. For example, CBMs could be used to transfer drugs in one direction and simultaneously remove toxic molecules in the other direction, as well as to store and remove molecular information in computational devices(Miller et al., 2018; Shoseyov et al., 2006). Understanding of the molecular mechanism by which CBMs move on the respective polysaccharides would be of considerable interest. However, it will take time and effort to harness the full potential of these molecules.

Bibliography

Abbott DW, Boraston AB. 2012. Quantitative Approaches to The Analysis of Carbohydrate-Binding Module Function. *Methods Enzymol.* 1st ed. Elsevier Inc. Vol. 510 211–231 p. http://dx.doi.org/10.1016/B978-0-12-415931-0.00011-2.

Abdul Khalil HPS, Bhat AH, Ireana Yusra AF. 2014. Green composites from sustainable cellulose nanofibrils: A review. *Sci. For. Sci.* **40**:345–351. http://dx.doi.org/10.1016/j.carbpol.2011.08.078.

Abitbol T, Rivkin A, Cao Y, Nevo Y, Abraham E, Ben-Shalom T, Lapidot S, Shoseyov O. 2016. Nanocellulose, a tiny fiber with huge applications. *Curr. Opin. Biotechnol.* **39**:76–88. http://dx.doi.org/10.1016/j.copbio.2016.01.002.

Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB. 2011. Biomass pretreatment : Fundamentals toward application. *Biotechnol. Adv.* **29**:675–685. http://dx.doi.org/10.1016/j.biotechadv.2011.05.005.

Aïssa K, Novy V, Nielsen F, Saddler J. 2019. Use of Carbohydrate Binding Modules To Elucidate the Relationship between Fibrillation, Hydrolyzability, and Accessibility of Cellulosic Substrates. *ACS Sustain. Chem. Eng.* **7**:1113–1119. http://pubs.acs.org/doi/10.1021/acssuschemeng.8b04780.

Ali E, Vinit S, Gilkes NR, Douglas KG, Waren AR, Saddler JN. 2001. Do cellulose binding domains Increase Substrate Accessibility ? *Appl. Biochem. biotechnolgy* **91**:575–592.

Alonso DM, Bond JQ, Dumesic J a. 2010. Catalytic conversion of biomass to biofuels. *Green Chem.* **12**:1493.

Ander P, Hildén L, Daniel G. 2008. Cleavage of softwood kraft pulp fibres by HCL and cellulases. *BioResources* **3**:477–490.

Araki J, Wada M, Kuga S. 2001. Steric Stabilization of a Cellulose Microcrystal Suspension by Poly (ethylene glycol) Grafting. *Cellulose* **17**:21–27.

Araki Y, Karita S, Tanaka A, Kondo M, Goto M. 2009. Characterization of Family 17 and Family 28 Carbohydrate-Binding Modules from Clostridium josui Cel5A. *Biosci. Biotechnol. Biochem.* **73**:1028–1032. http://www.tandfonline.com/doi/full/10.1271/bbb.80802.

Arantes V, Saddler JN. 2011. Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnol. Biofuels* **4**:3. http://www.ncbi.nlm.nih.gov/pubmed/21310050.

Araya F, Troncoso E, Mendonça RT, Freer J. 2015. Condensed lignin structures and re-localization achieved at high severities in autohydrolysis of Eucalyptus globulus wood and their relationship with cellulose accessibility. *Biotechnol. Bioeng.* **112**:1783–1791.

Arslan B, Colpan M, Ju X, Zhang X, Kostyukova A, Abu-Lail NI. 2016. The Effects of Noncellulosic Compounds on the Nanoscale Interaction Forces Measured between Carbohydrate-Binding Module and Lignocellulosic Biomass. *Biomacromolecules* **17**:1705–1715.

Babu RP, O'Connor K, Seeram R. 2013. Current progress on bio-based polymers and their future trends. *Prog. Biomater.* **2**:8. http://www.progressbiomaterials.com/content/2/1/8.

Beckham GT, Matthews JF, Peters B, Bomble YJ, Himmel ME, Crowley MF. 2011. Molecular-level origins of biomass recalcitrance: Decrystallization free energies for four common cellulose polymorphs. *J. Phys. Chem. B* **115**:4118–4127.

Bendzalova M. 1996. Accessibility of swollen cellulosic fibers. *Cellul. Chem. Technol.* **30**:19. http://gw2jh3xr2c.scholar.serialssolutions.com/?sid=google&auinit=M&aulast=Bendzalova&atitle=Accessibility+of+swollen+cellulosic+fibers&title=Cellulose+chemistry+and+technology&volume=30&issue=1-2&date=1996&spage=19&issn=0576-9787.

Blake a. W, McCartney L, Flint JE, Bolam DN, Boraston a. B, Gilbert HJ, Knox JP. 2006. Understanding the Biological Rationale for the Diversity of Cellulose-directed Carbohydrate-binding Modules in Prokaryotic Enzymes. *J. Biol. Chem.* **281**:29321–29329. http://www.jbc.org/cgi/doi/10.1074/jbc.M605903200.

Boraston AB, Chiu P, Warren RAJAJ, Kilburn DG. 2000. Specificity and affinity of substrate binding by a family 17 carbohydrate-binding module from Clostridium cellulovorans cellulase 5A. *Biochemistry* **39**:11129–11136.

Boraston AB, Kwan E, Chiu P, Warren RAJ, Kilburn DG. 2003. Recognition and hydrolysis of noncrystalline cellulose. *J. Biol. Chem.* **278**:6120–6127. http://www.ncbi.nlm.nih.gov/pubmed/12427734.

Boraston AB, McLean BW, Chen G, Li A, Warren RAJ, Kilburn DG. 2002. Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase. *Mol. Microbiol.* **43**:187–194.

Boraston AB. 2005. The interaction of carbohydrate-binding modules with insoluble non-crystalline cellulose is enthalpically driven. *Biochem. J.* **385**:479–484.

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004. Carbohydrate-binding modules : fine-tuning polysaccharide recognition. *Biochem. J.* **382**:769–781. http://www.scopus.com/inward/record.url?eid=2-s2.0-4744368323&partnerID=tZOtx3y1.

Boraston AB, McLean BW, Guarna MM, Amandaron-Akow E, Kilburn DG. 2001. A family 2a carbohydrate-binding module suitable as an affinity tag for proteins produced in Pichia pastoris. *Protein Expr. Purif.* **21**:417–423. http://www.ncbi.nlm.nih.gov/pubmed/11281716.

Cai J, He Y, Yu X, Banks SW, Yang Y, Zhang X, Yu Y, Liu R, Bridgwater A V. 2017. Review of physicochemical properties and analytical characterization of lignocellulosic biomass. *Renew. Sustain. Energy Rev.* **76**:309–322.

Caitriona M, Shawn M, Touhy M, Saddler JN. 1998. The effect of initial pore volume and lignin content on the enzymatic hydrolysis of softwoods. *Bioreseource Technol.* **64**:113–119.

Carrard G, Koivula a, Söderlund H, Béguin P. 2000. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc. Natl. Acad. Sci. U. S. A.* **97**:10342–10347.

Carvalho CC, Phan NN, Chen Y, Reilly PJ. 2014. Carbohydrate-Binding Module Tribes. *Biopolymers* **103**.

Chandra R, Ewanick S, Hsieh C, Saddler JN. 2008. The characterization of pretreated lignocellulosic substrates prior to enzymatic hydrolysis. *Biotechnol. Prog. 2008,* **24**:1178–1185.

Chandra RP, Saddler JN. 2012. Use of the Simons' Staining Technique to Assess Cellulose Accessibility in Pretreated Substrates. *Ind. Biotechnol.* **8**:230–237. http://online.liebertpub.com/doi/abs/10.1089/ind.2012.0016.

Clarke K, Li X, Li K. 2011. The mechanism of fiber cutting during enzymatic hydrolysis of wood biomass. *Biomass and Bioenergy* **35**:3943–3950. http://dx.doi.org/10.1016/j.biombioe.2011.06.007.

Creagh AL, Ongt E, Jervis E, Kilburnt DG, Haynes CA. 1996. Binding of the cellulose-binding domain of exoglucanase Cex from Cellulomonas fimi to insoluble microcrystalline cellulose is entropically driven. *Proc. Natl. Acad. Sci.* **93**:12229–12234.

Credou J, Berthelot T. 2014. Cellulose: from biocompatible to bioactive material. *J. Mater. Chem.* **2**:4767–4788.

Cunha AG, Gandini A. 2010. Turning polysaccharides into hydrophobic materials: A critical review. Part 1. Cellulose. *Cellulose* **17**:875–889.

Dam TK, Brewer CF. 2002. Thermodynamic Studies of Lectin − Carbohydrate Interactions by

Isothermal Titration Calorimetry Thermodynamic Studies of Lectin − Carbohydrate Interactions by Isothermal Titration Calorimetry. *Chem. rev.* **102**:387–430.

Daniel G. 1994. Use of electron microscopy for aiding our understanding of wood biodegradation. *FEMS Microbiol. Rev.* **13**:199–233.

Din N, Al E. 1991. non-hydolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Nat. Biochem.* **9**:1096.

Donohoe BS, Decker SR, Tucker MP, Himmel ME, Vinzant TB. 2008. Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnol. Bioeng.* **101**:913–925.

Donovan RS, Robinson CW, Click BR. 1996. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *J. Ind. Microbiol.* **16**:145–154.

Dou C, Ewanick S, Bura R, Gustafson R. 2016. Post-treatment mechanical refining as a method to improve overall sugar recovery of steam pretreated hybrid poplar. *Bioresour. Technol.* **207**:157–165. http://dx.doi.org/10.1016/j.biortech.2016.01.076.

Dufresne A. 2013. Nanocellulose: A new ageless bionanomaterial. *Mater. Today* **16**:220–227. http://dx.doi.org/10.1016/j.mattod.2013.06.004.

Eaton PJ, Graham P, Smith JR, Smart JD, Nevell TG, Tsibouklis J. 2000. Mapping the surface heterogeneity of a polymer blend: An adhesion-force-distribution study using the atomic force microscope. *Langmuir* **16**:7887–7890.

Eder M, Terziev N, Daniel G, Burgert I. 2008. The effect of (induced) dislocations on the tensile properties of individual Norway spruce fibres. *Holzforschung* **62**:77–81.

Esa F, Tasirin SM, Rahman NA. 2014. Overview of Bacterial Cellulose Production and Application. *Agric. Agric. Sci. Procedia* **2**:113–119.

http://www.sciencedirect.com/science/article/pii/S2210784314000187.

Esteghlalian AR, Bilodeau M, Mansfield SD, Saddler JN. 2001. Do enzymatic digestibility and Simons' stain reflect the differences in the available surface area of lignocellulosic substrates. *Biotechnol. Prog.* **17**:1049–1054.

Filonova L, Kallas ÅM, Greffe L, Johansson G, Teeri TT, Daniel G. 2007. Analysis of the surfaces of wood tissues and pulp fibers using carbohydrate-binding modules specific for crystalline cellulose and mannan. *Biomacromolecules* **8**:91–97.

Fisher SZ, von Schantz L, Håkansson M, Logan DT, Ohlin M. 2015. Neutron crystallographic studies reveal hydrogen bond and water mediated interactions between a carbohydrate-binding module and its bound carbohydrate ligand. *Biochemistry* **54**:6435−6438. http://pubs.acs.org/doi/10.1021/acs.biochem.5b01058.

Fox JM, Jess P, Jambusaria RB, Moo GM, Liphardt J, Clark DS, Blanch HW. 2013. A single-molecule analysis reveals morphological targets for cellulase synergy. *Nat. Chem. Biol.* **9**:356–61. http://www.ncbi.nlm.nih.gov/pubmed/23563526.

French C, Keshavarz-Moore E, Ward JM. 1996. Development of a simple method for the recovery of recombinant proteins from the Escherichia coli periplasm. *Enzyme Microb. Technol.* **19**:332–338.

Galbe M, Zacchi G. 2002. A review of the production of ethanol from softwood. *Appl. Microbiol. Biotechnol.* **59**:618–628.

Gao S, You C, Renneckar S, Bao J, Zhang Y-HP. 2014. New insights into enzymatic hydrolysis of heterogeneous cellulose by using carbohydrate-binding module 3 containing GFP and carbohydrate-binding module 17 containing CFP. *Biotechnol. Biofuels* **7**:24. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3943381&tool=pmcentrez&rendertype=abstract.

Garcia-Ochoa F, Gomez E. 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. *Biotechnol. Adv.* **27**:153–176.

Gilkes NR, Antony R, Warren J, Miller RC, Kilburn DG. 1988. Precise Excision of the Cellulose Binding Domains from Two Cellulomonas fimi Cellulases by a Homologous Protease and the Effect on Catalysis. *J. Biol. Chem.* **263**:10401–10407.

Gírio FM, Fonseca C, Carvalheiro F, Duarte LC, Marques S, Bogel-Łukasik R. 2010. Hemicelluloses for fuel ethanol: A review. *Bioresour. Technol.* **101**:4775–4800.

Gourlay K, Arantes V, Saddler JN. 2012. Use of substructure-specific carbohydrate binding modules to track changes in cellulose accessibility and surface morphology during the amorphogenesis step of enzymatic hydrolysis. *Biotechnol. Biofuels* **5**:51. ???

Gourlay K, Hu J, Arantes V, Saddler JN. 2015. The Use of Carbohydrate Binding Modules (CBMs) to Monitor Changes in Fragmentation and Cellulose Fiber Surface Morphology during Cellulase- and Swollenin-induced Deconstruction of Lignocellulosic Substrates *. *J. Biol. Chem.* **290**:2938–2945.

Grethlein HE, Converse AO. 1991. Common aspects of acid prehydrolysis and steam explosion for pretreating wood. *Bioresour. Technol.* **36**:77–82.

Gu F, Wang W, Cai Z, Xue F, Jin Y, Zhu JY. 2018. Water retention value for characterizing fibrillation degree of cellulosic fibers at micro and nanometer scales. *Cellulose* **25**:2861–2871. https://doi.org/10.1007/s10570-018-1765-8.

Gustavsson MT, Persson P V., Iversen T, Martinelle M, Hult K, Teeri TT, Brumer H. 2005. Modification of cellulose fiber surfaces by use of a lipase and a xyloglucan endotransglycosylase. *Biomacromolecules* **6**:196–203.

Hasenwinkle D, Jervis E, Kops O, Liu C, Lesnicki G, Haynes CA, Kilburn DG. 1997. Very high-level production and export in Escherichia coli of a cellulose binding domain for use in a

generic secretion-affinity fusion system. *Biotechnol. Bioeng.* **55**:854–863.

Hebert-Ouellet Y, Meddeb-Mouelhi F, Khatri V, Cui L, Janse B, MacDonald K, Beauregard M. 2017. Tracking and predicting wood fibers processing with fluorescent carbohydrate binding modules. *Green Chem.* **19**. http://dx.doi.org/10.1039/C6GC03581G.

Hemsworth GR, Johnston EM, Davies GJ, Walton PH. 2015. Lytic Polysaccharide Monooxygenases in Biomass Conversion. *Trends Biotechnol.*

Henriksson M, Henriksson G, Berglund LA, Lindström T. 2007. An environmentally friendly method for enzyme-assisted preparation of microfibrillated cellulose (MFC) nanofibers. *Eur. Polym. J.* **43**:3434–3441.

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

Hermanson GT. 2013a. Bioconjugate Techniques. *Bioconjugate Tech.* Elsevier 921–949 p. http://www.sciencedirect.com/science/article/pii/B9780123822390000212.

Hermanson GT. 2013b. Bioconjugate Techniques. *Bioconjugate Tech.* Elsevier 549–587 p. http://www.sciencedirect.com/science/article/pii/B9780123822390000145.

Herve C, Rogowski a., Blake a. W, Marcus SE, Gilbert HJ, Knox JP. 2010. Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc. Natl. Acad. Sci.* **107**:15293–15298. http://www.pnas.org/cgi/doi/10.1073/pnas.1005732107.

Hidayat BJ, Felby C, Johansen KS, Thygesen LG. 2012. Cellulose is not just cellulose: A review of dislocations as reactive sites in the enzymatic hydrolysis of cellulose microfibrils. *Cellulose* **19**:1481–1493.

Hidayat BJ, Weisskopf C, Felby C, Johansen KS, Thygesen LG. 2015. The binding of cellulase variants to dislocations: a semi-quantitative analysis based on CLSM (confocal laser

scanning microscopy) images. *AMB Express* **5**.

Hildén L, Daniel G, Johansson G. 2003. Use of a fluorescence labelled, carbohydrate-binding module from Phanerochaete chrysosporium Ce17D for studying wood cell wall ultrastructure. *Biotechnol. Lett.* **25**:553–558.

Hoeger IC, Nair SS, Ragauskas AJ, Deng Y, Rojas OJ, Zhu JY. 2013. Mechanical deconstruction of lignocellulose cell walls and their enzymatic saccharification. *Cellulose* **20**:807–818.

Hong J, Ye X, Zhang YHP. 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**:12535–12540.

Hu J, Arantes V, Saddler JN. 2011. The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnol. Biofuels* **4**:36. http://dx.doi.org/10.1186/1754-6834-4-36%5Cnhttp://www.biotechnologyforbiofuels.com/content/4/1/36%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/21974832%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3198685.

Hu J, Pribowo A, Saddler JN. 2016a. Oxidative cleavage of some cellulosic substrates by auxiliary activity (AA) family 9 enzymes influences the adsorption/desorption of hydrolytic cellulase enzymes. *Green Chem.* **18**:6329–6336. http://xlink.rsc.org/?DOI=C6GC02288J.

Hu Z, Marway HS, Kasem H, Pelton R, Cranston ED. 2016b. Dried and Redispersible Cellulose Nanocrystal Pickering Emulsions. *ACS Macro Lett.* **5**:185–189.

Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttilä M, Ando T, Samejima M, Penttila M, Ando T, Samejima M. 2011. Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science (80-.).* **187**:1279–1282. http://www.sciencemag.org/cgi/doi/10.1126/science.1208386.

Jamal-Talabani S, Boraston AB, Turkenburg JP, Tarbouriech N, Ducros VMA, Davies GJ. 2004. Ab initio structure determination and functional characterization of CBM36: A new family of calcium-dependent carbohydrate binding modules. *Structure* **12**:1177–1187.

Johnston W, Cord-Ruwisch R, Cooney MJ. 2002. Industrial control of recombinant E. coli fed-batch culture: New perspectives on traditional controlled variables. *Bioprocess Biosyst. Eng.* **25**:111–120.

Jones BW, Venditti R, Park S, Jameel H, Koo B. 2013. Enhancement in enzymatic hydrolysis by mechanical refining for pretreated hardwood lignocellulosics. *Bioresour. Technol.* **147**:353–360. http://dx.doi.org/10.1016/j.biortech.2013.08.030.

Josefsson P, Henriksson G, Wågberg L. 2008. The physical action of cellulases revealed by a quartz crystal microbalance study using ultrathin cellulose films and pure cellulases. *Biomacromolecules* **9**:249–254.

Kang B, Opatz T, Landfester K, Wurm FR. 2015. Carbohydrate nanocarriers in biomedical applications: Functionalization and construction. *Chem. Soc. Rev.* **44**:8301–8325.

Karaaslan MA, Gao G, Kadla JF. 2013. Nanocrystalline cellulose/β-casein conjugated nanoparticles prepared by click chemistry. *Cellulose* **20**:2655–2665.

Kataeva IA, Iii RDS, Shah A, West LT, Li X, Ljungdahl LG. 2002. The Fibronectin Type 3-Like Repeat from the Clostridium thermocellum Cellobiohydrolase CbhA Promotes Hydrolysis of Cellulose by Modifying Its Surface. *Appl. Environ. Microbiol.* **68**:4292–4300.

Kawakubo T, Karita S, Araki Y, Watanabe S, Oyadomari M, Takada R, Tanaka F, Abe K, Watanabe T, Honda Y, Watanabe T. 2010. Analysis of exposed cellulose surfaces in pretreated wood biomass using Carbohydrate-Binding Module (CBM)-Cyan Fluorescent Protein (CFP). *Biotechnol. Bioeng.* **105**:499–508.

Khoshkava V, Kamal MR. 2014. Effect of cellulose nanocrystals (CNC) particle morphology on

dispersion and rheological and mechanical properties of polypropylene/CNC nanocomposites. *ACS Appl. Mater. Interfaces* **6**:8146–8157.

King JR, Bowers CM, Toone EJ. 2015. Specific Binding at the Cellulose Binding Module–Cellulose Interface Observed by Force Spectroscopy. *Langmuir* **31**:3431–3440. http://pubs.acs.org/doi/abs/10.1021/la504836u.

Kitaoka T, Tanaka H. 2001. Novel paper strength additive containing cellulose-binding domain of cellulase. *J. Wood. Sci.* **47**:322–324.

Kleinman HK. 1987. Use of extracellular matrix components for cell culture. *Anal. Biochem.*

Klemm D, Cranston ED, Fischer D, Gama M, Kedzior SA, Kralisch D, Kramer F, Kondo T, Lindström T, Nietzsche S, Petzold-Welcke K, Rauchfuß F. 2018. Nanocellulose as a natural source for groundbreaking applications in materials science: Today's state. *Mater. Today* **21**:720–748.

Kljun A, Benians TAS, Goubet F, Meulewaeter F, Knox JP, Blackburn RS. 2011. Comparative analysis of crystallinity changes in cellulose i polymers using ATR-FTIR, X-ray diffraction, and carbohydrate-binding module probes. *Biomacromolecules* **12**:4121–4126.

Kloser E, Gray DG. 2010. Surface grafting of cellulose nanocrystals with poly(ethylene oxide) in aqueous media. *Langmuir* **26**:13450–13456.

Kolb HC, Finn MG, B. SK. 2009. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie - Int. Ed.* **65**:323–340.

Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer WD. 1995. Simple fed-batch technique for high cell density cultivation of Escherichia coli. *J. Biotechnol.* **39**:59–65.

Krause M, Ukkonen K, Haataja T, Ruottinen M, Glumoff T, Neubauer A, Neubauer P, Vasala A. 2010. A novel fed-batch based cultivation method provides high cell-density and improves yield of soluble recombinant proteins in shaken cultures. *Microb. Cell Fact.* **9**:11.

Kumar P, Barrett DM, Delwiche MJ, Stroeve P. 2009. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production:3713–3729.

Laaksonen P, Walther A, Malho JM, Kainlauri M, Ikkala O, Linder MB. 2011. Genetic engineering of biomimetic nanocomposites: Diblock proteins, graphene, and nanofibrillated cellulose. *Angew. Chemie - Int. Ed.* **50**:8688–8691.

Lavoine N, Desloges I, Dufresne A, Bras J. 2012. Microfibrillated cellulose - Its barrier properties and applications in cellulosic materials: A review. *Carbohydr. Polym.* **90**:735–764. http://dx.doi.org/10.1016/j.carbpol.2012.05.026.

Lee I, Evans BR, Woodward J. 2000. The mechanism of cellulase action on cotton fibers: Evidence from atomic force microscopy. *Ultramicroscopy* **82**:213–221.

Lee SY. 1996. High cell-density culture of Escherichia coli. *Trends Biotechnol.* **14**:98–105.

Lee SB, Shin HS, Ryu DDY, Mandels M. 1982. Adsorption of cellulase on cellulose: Effect of physicochemical properties of cellulose on adsorption and rate of hydrolysis. *Biotechnol. Bioeng.* **24**:2137–2153.

Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. 2003. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc. Natl. Acad. Sci. U. S. A.* **100**:484–489.

Levy I, Nussinovitch A, Shpigel E, Shoseyov O. 2002. Recombinant cellulose crosslinking protein: A novel paper-modification biomaterial. *Cellulose* **9**:91–98.

Levy I, Shoseyov O. 2002. Cellulose-binding domains: Biotechnological applications. *Biotechnol. Adv.* **20**:191–213.

Li Q, Renneckar S. 2011. Supramolecular structure characterization of molecularly thin cellulose i nanoparticles. *Biomacromolecules* **12**:650–659.

Lin JK, Ladisch MR, Patterson JA, Noller CH. 1987. Determining pore size distribution in wet cellulose by measuring solute exclusion using a differential refractometer. *Biotechnol. Bioeng.* **29**:976–981. http://doi.wiley.com/10.1002/bit.260290809.

Liu W, Wang B, Hou Q, Chen W, Wu M. 2016. Effects of fibrillation on the wood fibers' enzymatic hydrolysis enhanced by mechanical refining. *Bioresour. Technol.* **206**:99–103. http://dx.doi.org/10.1016/j.biortech.2016.01.074.

Losen M, Fro B, Pohl M, Bu J. 2004. Effect of oxygen limitation and medium composition on Escherichia coli fermentation in shake-flask cultures. *biotechnol. prog.* **20**:1062–1068.

Lowe CR. 2001. Combinatorial approaches to affinity chromatography. *Curr. Opin. Chem. Biol.* **5**:248–256.

Lowry H, Carter J, Ward JB, Glaser L. 1971. The Effect of Carbon Metabolic Intermediates and Nitrogen Sources in *Escherichia coli*. *J. bio* **246**:6511.

Lu X, Zheng X, Li X, Zhao J. 2016. Adsorption and mechanism of cellulase enzymes onto lignin isolated from corn stover pretreated with liquid hot water. *Biotechnol. Biofuels* **9**:1–12.

Luli GW, Strohl WR. 1990. Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations. *Appl. Environ. Microbiol.* **56**:1004–11. http://www.ncbi.nlm.nih.gov/pubmed/2187400.

Luterbacher JS, Parlange JY, Walker LP. 2013a. A pore-hindered diffusion and reaction model can help explain the importance of pore size distribution in enzymatic hydrolysis of biomass. *Biotechnol. Bioeng.* **110**:127–136.

Luterbacher JS, Walker LP, Moran-Mirabal JM. 2013b. Observing and modeling BMCC degradation by commercial cellulase cocktails with fluorescently labeled Trichoderma reseii Cel7A through confocal microscopy. *Biotechnol. Bioeng.* **110**:108–117.

Ma L, Wang T, Liu Q, Zhang X, Ma W, Zhang Q. 2012. A review of thermal-chemical

conversion of lignocellulosic biomass in China. *Biotechnol. Adv.* **30**:859–73. http://www.scopus.com/inward/record.url?eid=2-s2.0-84862806037&partnerID=tZOtx3y1.

Mansfield SD, De Jong E, Stephens RS, Saddler JN. 1997. Physical characterization of enzymatically modified kraft pulp fibers. *J. Biotechnol.* **57**:205–216.

Mansfield SD, Mooney C, Saddler JN. 1999. Substrate and enzymatic characteristics that limit cellulose hydrolysis. *Biotechnol. Prog.* **15**:804–816.

McCartney L, Gilbert HJ, Bolam DN, Boraston AB, Knox JP. 2004. Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. *Anal. Biochem.* **326**:49–54.

Mcguffee SR, Elcock AH. 2010. Diffusion , Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm. *PLoS Comput. Biol.* **6**.

McLean BW. 2002. Carbohydrate-binding Modules Recognize Fine Substructures of Cellulose. *J. Biol. Chem.* **277**:50245–50254. http://www.jbc.org/cgi/doi/10.1074/jbc.M204433200.

McLean BW, Bray MR, Boraston AB, Gilkes NR, Haynes CA, Kilburn DG. 2000. Analysis of binding of the family 2a carbohydrate-binding module from Cellulomonas fimi xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. *Protein Eng.* **13**:801–809.

Meng X, Foston M, Leisen J, DeMartini J, Wyman CE, Ragauskas AJ. 2013. Determination of porosity of lignocellulosic biomass before and after pretreatment by using Simons' stain and NMR techniques. *Bioresour. Technol.* **144**:467–476. http://dx.doi.org/10.1016/j.biortech.2013.06.091.

Meng X, Ragauskas AJ. 2014. ScienceDirect Recent advances in understanding the role of cellulose accessibility in enzymatic hydrolysis of lignocellulosic substrates. *Curr. Opin. Biotechnol.* **27**:150–158. http://dx.doi.org/10.1016/j.copbio.2014.01.014.

Miller EA, Baniya S, Osorio D, Al Maalouf YJ, Sikes HD. 2018. Paper-based diagnostics in the antigen-depletion regime: High-density immobilization of rcSso7d-cellulose-binding domain fusion proteins for efficient target capture. *Biosens. Bioelectron.* **102**:456–463.

Missoum K, Belgacem M, Bras J. 2013. Nanofibrillated Cellulose Surface Modification: A Review. *Materials (Basel).* **6**:1745–1766. http://www.mdpi.com/1996-1944/6/5/1745/.

Moon RJ, Martini A, Nairn J, Simonsen J, Youngblood J. 2011. Cellulose nanomaterials review: structure, properties and nanocomposites. *Chem. Soc. Rev.* Vol. 40 3941–3994 p.

Morais JPS, Rosa MDF, De Souza Filho MDSM, Nascimento LD, Do Nascimento DM, Cassales AR. 2013. Extraction and characterization of nanocellulose structures from raw cotton linter. *Carbohydr. Polym.* **91**:229–235.

Mordocco A, Kuek C, Jenkins R. 1999. Continuous degradation of phenol at low concentration using immobilized Pseudomonas putida. *Enzyme Microb. Technol.* **25**:530–536.

Nakagaito AN, Yano H. 2004. The effect of morphological changes from pulp fiber towards nano-scale fibrillated cellulose on the mechanical properties of high-strength plant fiber based composites. *Appl. Phys. A Mater. Sci. Process.* **78**:547–552.

Nascimento DM, Nunes YL, Figueirêdo MCB, De Azeredo HMC, Aouada FA, Feitosa JPA, Rosa MF, Dufresne A. 2018. Nanocellulose nanocomposite hydrogels: Technological and environmental issues. *Green Chem.* **20**:2428–2448.

Nieves RA, Ellis RP, Todd RJ, Timothy JA, Grohmann K, Himmel ME. 1991. Visualization of Trichoderma reesei Cellobiohydrolase I and Endoglucanase I on Aspen Cellulose by Using Monoclonal Visualization of Trichoderma reesei Cellobiohydrolase I and Endoglucanase I on Aspen Cellulose by Using Monoclonal Antibody-Colloidal Gold C. *Appl. Environ. Microbiol.* **57**:3163–3170.

Nigmatullin R, Lovitt R, Wright C, Linder M, Nakari-Setälä T, Gama M. 2004. Atomic force

microscopy study of cellulose surface interaction controlled by cellulose binding domains. *Colloids Surfaces B Biointerfaces* **35**.

Nishiyama Y, Sugiyama J, Chanzy H, Langan P. 2003. Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **125**:14300–14306.

Notenboom V, Boraston AB, Chiu P, Freelove AC. J, Kilburn DG, Rose DR. 2001. Recognition of cello-oligosaccharides by a family 17 carbohydrate-binding module: an X-ray crystallographic, thermodynamic and mutagenic study 1 1Edited by R. Huber. *J. Mol. Biol.* **314**:797–806. http://linkinghub.elsevier.com/retrieve/pii/S0022283601951538.

Oehme DP, Downton MT, Doblin MS, Wagner J, Gidley MJ, Bacic A. 2015. Unique aspects of the structure and dynamics of elementary iβ cellulose microfibrils revealed by computational simulations. *Plant Physiol.* **168**:3–17. http://www.plantphysiol.org/content/168/1/3.abstract.

Ofir K, Berdichevsky Y, Benhar I, Azriel-Rosenfeld R, Lamed R, Barak Y, Bayer EA, Morag E. 2005. Versatile protein microarray based on carbohydrate-binding modules. *Proteomics* **5**:1806–1814.

Oliveira C, Carvalho V, Domingues L, Gama FM. 2015. Recombinant CBM-fusion technology — Applications overview. *Biotechnol. Adv.* **33**:358–369. http://dx.doi.org/10.1016/j.biotechadv.2015.02.006.

Östlund Å, Idström A, Olsson C, Larsson PT, Nordstierna L. 2013. Modification of crystallinity and pore size distribution in coagulated cellulose films. *Cellulose* **20**:1657–1667.

Pahimanolis N, Hippi U, Johansson LS, Saarinen T, Houbenov N, Ruokolainen J, Seppälä J. 2017. Surface functionalization of nanofibrillated cellulose using click-chemistry approach in aqueous media. *Cellulose* **24**:5707.

Pala H, Lemos MA, Mota M, Gama FM. 2001. Enzymatic upgrade of old paperboard containers. *Enzyme Microb. Technol.* **29**:274–279.

Palonen H, Tjerneld F, Zacchi G, Tenkanen M. 2004. Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *J. Biotechnol.* **107**:65–72.

Park J, Jones B, Koo B, Chen X, Tucker M, Yu JH, Pschorn T, Venditti R, Park S. 2016. Use of mechanical refining to improve the production of low-cost sugars from lignocellulosic biomass. *Bioresour. Technol.* **199**:59–67. http://dx.doi.org/10.1016/j.biortech.2015.08.059.

Park J, Meng J, Lim KH, Rojas OJ, Park S. 2013. Transformation of lignocellulosic biomass during torrefaction. *J. Anal. Appl. Pyrolysis* **100**:199–206. http://dx.doi.org/10.1016/j.jaap.2012.12.024.

Park S, Baker JO, Himmel ME, Parilla P a, Johnson DK. 2010. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol. Biofuels* **3**:10. http://www.biotechnologyforbiofuels.com/content/3/1/10.

Pelton R. 2009. Bioactive paper provides a low-cost platform for diagnostics. *TrAC - Trends Anal. Chem.* **28**:925–942.

Peukert W, Gotzinger M. 2003. Dispersive forces of particle – surface interactions : direct AFM measurements and modelling Martin Go. *Powder Technol.* **130**:102–109.

Phelps MR, Hobbs JB, Kilburn DG, Turner RFB. 1994. Technology for Regenerable Biosensor Probes Based on Enzyme-Cellulose Binding Domain Conjugates. *Biotechnol. Prog.* **10**:433–440.

Ponnusamy VK, Nguyen DD, Dharmaraja J, Shobana S, Banu JR, Saratale RG, Chang SW, Kumar G. 2019. A review on lignin structure, pretreatments, fermentation reactions and biorefinery potential. *Bioresour. Technol.* **271**:462–472.

Presolski SL, Hong VP, Finn MG. 2011. Copper-Catalyzed Azide–Alkyne Click Chemistry for Bioconjugation. *Curr. Protoc. Chem. Biol.* **3**:153–162.

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon R a, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan G a, Wyman CE. 2014. Lignin valorization: improving lignin processing in the biorefinery. *Science* **344**:1246843. http://www.ncbi.nlm.nih.gov/pubmed/24833396.

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick Jr. WJ, Hallett JP, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, Tschaplinski T. 2006. The Path Forward for Biofuels and Biomaterials\r10.1126/science.1114736. *Science (80-.).* **311**:484–489. http://www.sciencemag.org/cgi/content/abstract/311/5760/484.

Rahikainen JL, Martin-Sampedro R, Heikkinen H, Rovio S, Marjamaa K, Tamminen T, Rojas OJ, Kruus K. 2013a. Inhibitory effect of lignin during cellulose bioconversion: The effect of lignin chemistry on non-productive enzyme adsorption. *Bioresour. Technol.* **133**.

Rahikainen JL, Evans JD, Mikander S, Kalliola A, Puranen T, Tamminen T, Marjamaa K, Kruus K. 2013b. Cellulase-lignin interactions-The role of carbohydrate-binding module and pH in non-productive binding. *Enzyme Microb. Technol.* **53**:315–321. http://dx.doi.org/10.1016/j.enzmictec.2013.07.003.

Rahimi M, Safa KD, Salehi R. 2017. Co-delivery of doxorubicin and methotrexate by dendritic chitosan-g-mPEG as a magnetic nanocarrier for multi-drug delivery in combination chemotherapy. *Polym. Chem.* **8**:7333–7350.

Ramamoorthy SK, Skrifvars M, Persson A. 2015. A review of natural fibers used in biocomposites: Plant, animal and regenerated cellulose fibers. *Polym. Rev.* **55**:107–162.

Riesenberg D, Schulz V, Knorre WAA, Pohl H-DD, Korz D, Sanders EAA, Ross A, Deckwer W-DD, Roß A, Deckwer W-DD. 1991. High cell density cultivation of Escherichia coli at controlled specific growth rate. *J. Biotechnol.* **20**:17–27.

http://www.ncbi.nlm.nih.gov/pubmed/1367313.

Rosa AMM, Louro AF, Martins SAM, Inácio J, Azevedo AM, Prazeres DMF. 2014. Capture and detection of DNA hybrids on paper via the anchoring of antibodies with fusions of carbohydrate binding modules and ZZ-domains. *Anal. Chem.* **86**:4340–4347.

Rubin EM. 2008. Genomics of cellulosic biofuels. *Nat. Rev.* **454**:841–845.

Shevchenko SM, Beatson RP, Saddler JN. 1999. The Nature of Lignin from Steam Explosion/ Enzymatic Hydrolysis of Softwood: Structural Features and Possible Uses(Scientific Note). *Appl. Biochem. Biotechnol.* **79**:867–876.

Shoseyov O. 1999. Methods of detection using a CBD fusion product. *United States Pat.*

Shoseyov O, Shani Z, Levy I. 2006. Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* **70**:283–295. http://www.scopus.com/inward/record.url?eid=2-s2.0-33745161547&partnerID=tZOtx3y1.

Široký J, Benians T a S, Russell SJ, Bechtold T, Paul Knox J, Blackburn RS. 2012. Analysis of crystallinity changes in cellulose II polymers using carbohydrate-binding modules. *Carbohydr. Polym.* **89**:213–221.

Structural Genomics Consortium et al. 2008. Protein production and purification Structural. *Nat Methods.* **5**:135–146.

Suchy M, Hakala T, Kangas H, Kontturi E, Tammelin T, Pursula T, Vuorinen T. 2009. Effects of commercial cellobiohydrolase treatment on fiber strength and morphology of bleached hardwood pulp. *Holzforschung* **63**:731–736.

Sugiyama J, Vuong R, Chanzy H. 1991. Electron Diffraction Study on the Two Crystalline Phases Occurring in Native Cellulose from an Algal Cell Wall. *Macromolecules* **24**:4168–4175.

Tang J, Sisler J, Grishkewich N, Tam KC. 2017. Functionalization of cellulose nanocrystals for advanced applications. *J. Colloid Interface Sci.* **494**:397–409. http://dx.doi.org/10.1016/j.jcis.2017.01.077.

Tang Z, Chen H, Chen L, Liu S, Han X, Wu Q. 2014. Improving endoglucanase activity by adding the carbohydrate-binding module from Corticium rolfsii. *J. Microbiol. Biotechnol.* **24**:440–446.

Teeri TT, Brumer H, Daniel G, Gatenholm P. 2007. Biomimetic engineering of cellulose-based materials. *Trends Biotechnol.* **25**:299–306.

Thygesen LG, Bilde-Sørensen JB, Hoffmeyer P. 2006. Visualisation of dislocations in hemp fibres: A comparison between scanning electron microscopy (SEM) and polarized light microscopy (PLM). *Ind. Crops Prod.* **24**:181–185.

Thygesen LG, Hidayat BJ, Johansen KS, Felby C. 2011. Role of supramolecular cellulose structures in enzymatic hydrolysis of plant cell walls. *J. Ind. Microbiol. Biotechnol.* **38**:975–983.

Tomme P, Driver DP, Amandoron EA, Miller RC, Antony R, Warren J, Kilburn DG. 1995. Comparison of a fungal (family I) and bacterial (family II) cellulose-binding domain. *J. Bacteriol.* **177**:4356–63. http://www.ncbi.nlm.nih.gov/pubmed/7635821.

Tomme P, Tilbeurgh HVAN, Pettersson G, Damme JVAN, Vandekerckhove J, Knowles J, Teer T, Claeyssens M. 1988. Studies of the cellulolytic system of Trichoderma reesei QM 9414. *Eur. J. Biochem.* **581**:575–581.

Tormo J, Lamed R, Chirinol AJ, Morag E, Bayer EA, Shoham Y, Steitz TA. 1996. Crystal structure of a bacterial family-lIl cellulose-binding domain : a general mechanism for attachment to cellulose. *EMBO J.* **15**:5739–5751.

Trache D, Hussin MH, Hui Chuin CT, Sabar S, Fazita MRN, Taiwo OFA, Hassan TM, Haafiz

MKM. 2016. Microcrystalline cellulose: Isolation, characterization and bio-composites application—A review. *Int. J. Biol. Macromol.* **93**:789–804.

Valo H, Kovalainen M, Laaksonen P, Häkkinen M, Auriola S, Peltonen L, Linder M, Järvinen K, Hirvonen J, Laaksonen T. 2011. Immobilization of protein-coated drug nanoparticles in nanofibrillar cellulose matrices-Enhanced stability and release. *J. Control. Release* **156**:390–397. http://dx.doi.org/10.1016/j.jconrel.2011.07.016.

Varjonen S, Laaksonen P, Paananen A, Valo H, Hähl H, Laaksonen T, Linder M Ben. 2011. Self-assembly of cellulose nanofibrils by genetically engineered fusion proteins. *Soft Matter* **7**:2402. http://xlink.rsc.org/?DOI=c0sm01114b%5Cnhttp://pubs.rsc.org/en/content/articlehtml/2011/sm/c0sm01114b.

Várnai A, Siika-Aho M, Viikari L. 2013. Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnol. Biofuels* **6**:30. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3599012&tool=pmcentrez&rendertype=abstract.

Varshney VK, Naithani S. 2011. Chemical Functionalization of Cellulose Derived from Nonconventional Sources. *Cellul. Fibers Bio- Nano-Polymer Compos.* 43–61 p. http://link.springer.com/10.1007/978-3-642-17370-7.

Wang L, Zhang Y, Gao P. 2008. A novel function for the cellulose binding module of cellobiohydrolase I. *Sci. China Ser. C Life Sci.* **51**:620–629. http://link.springer.com/10.1007/s11427-008-0088-3.

Wang QQ, He Z, Zhu Z, Zhang YHP, Ni Y, Luo XL, Zhu JY. 2012. Evaluations of cellulose accessibilities of lignocelluloses by solute exclusion and protein adsorption techniques. *Biotechnol. Bioeng.* **109**:381–389.

Way AE, Hsu L, Shanmuganathan K, Weder C, Rowan SJ. 2012. PH-responsive cellulose

nanocrystal gels and nanocomposites. *ACS Macro Lett.* **1**:1001–1006.

Welf ES, Venditti RA. 2005. The effects of heating without water removal and drying on the swelling as measured by water retention value and degradation as measured by. *Prog. Pap. Recycl.* **14**:1–9. http://repository.lib.ncsu.edu/publications/handle/1840.2/35.

Westermarck S. 2000. Use of Mercury Porosimetry and Nitrogen Adsorption in Characterisation of the Pore Structure of Mannitol and Microcrystalline Cellulose Powders, Granules and Tablets p.

Wiman M, Dienes D, Hansen MAT, Van Der Meulen T, Zacchi G, Lidén G. 2012. Cellulose accessibility determines the rate of enzymatic hydrolysis of steam-pretreated spruce. *Bioresour. Technol.* **126**:208–215. http://dx.doi.org/10.1016/j.biortech.2012.08.082.

Wong KKY, Richardson JD, Mansfield SD. 2000. Enzymatic treatment of mechanical pulp fibers for improving papermaking properties. *Biotechnol. Prog.* **16**:1025–1029.

Xiao LP, Sun ZJ, Shi ZJ, Xu F, Sun RC. 2011. Impact of hot compressed water pretreatment on the structural changes of woody biomass for bioethanol production. *BioResources* **6**:1576–1598.

Yee L, Blanch HW. 1992. Recombinant Protein Expression in High Cell Density Fed-Batch Cultures of Escherichia Coli. *Nat. Biotechnol.* **10**:1550–1556. http://www.nature.com/doifinder/10.1038/nbt1292-1550.

Yeh AI, Huang YC, Chen SH. 2010. Effect of particle size on the rate of enzymatic hydrolysis of cellulose. *Carbohydr. Polym.* **79**:192–199.

Yoo Y, Youngblood JP. 2016. Green one-pot synthesis of surface hydrophobized cellulose nanocrystals in aqueous medium. *ACS Sustain. Chem. Eng.* **4**:3927–3938.

Yu X, Minor JL, Atalla RH. Mechanism of action of Simons ’ stain:175–180.

Zhang M, Wang B, Xu B. 2013a. Measurements of single molecular affinity interactions between carbohydrate-binding modules and crystalline cellulose fibrils. *Phys. Chem. Chem. Phys.* **15**:6508–6515. http://pubs.rsc.org.ezproxy.lib.uh.edu/en/content/articlelanding/2013/cp/c3cp51072g%5Cnhttp://pubs.rsc.org.ezproxy.lib.uh.edu/en/Content/ArticleLanding/2013/CP/c3cp51072g#!divAbstract%5Cnhttp://pubs.rsc.org.ezproxy.lib.uh.edu/en/content/articlepdf/2013/cp/c3c.

Zhang M, Wang B, Xu B. 2014. Mapping Single Molecular Binding Kinetics of Carbohydrate-Binding Module with Crystalline Cellulose by Atomic Force Microscopy Recognition Imaging. *J. Phys. Chem.* **118**:6714−6720.

Zhang W, Johnson RK, Lin Z, Chandoha-Lee C, Zink-Sharp A, Renneckar S. 2013b. In situ generated cellulose nanoparticles to enhance the hydrophobicity of paper. *Cellulose* **20**:2935–2945.

Zhang YP. 2008. Reviving the carbohydrate economy via multi-product lignocellulose biore W neries:367–375.

Zhang YP, Lynd LR. 2004. Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose : Noncomplexed Cellulase Systems. *Biotechnol. Bioeng.* **88**:797–821.

Zhang Y. 2010. Biofuels from Agricultural Wastes and Byproducts. Ed. Hans P. Blaschek, Thaddeus C. Ezeji, Jürgen Scheffran. *Biofuels from Agric. Wastes Byprod.* Oxford, UK: Wiley-Blackwell 201–232 p. http://www.scopus.com/inward/record.url?eid=2-s2.0-84871373412&partnerID=tZOtx3y1.

Zhao L, Pang Q, Xie J, Pei J, Wang F, Fan S. 2013. Enzymatic properties of Thermoanaerobacterium thermosaccharolyticum β-glucosidase fused to Clostridium cellulovorans cellulose binding domain and its application in hydrolysis of microcrystalline cellulose. *BMC Biotechnol.* **13**:101. http://www.ncbi.nlm.nih.gov/pubmed/24228818%5Cnhttp://www.pubmedcentral.nih.gov/ar

ticlerender.fcgi?artid=PMC3840712.

Zheng Y, Pan Z, Zhang R. 2009. Overview of biomass pretreatment for cellulosic ethanol production. *Int J Agric Biol Eng* **2**:51–69.

Zhu P, Moran-Mirabal JM, Luterbacher JS, Walker LP, Craighead HG. 2011. Observing Thermobifida fusca cellulase binding to pretreated wood particles using time-lapse confocal laser scanning microscopy. *Cellulose* **18**:749–758.

Zoppe JO, Venditti RA, Rojas OJ. 2012. Pickering emulsions stabilized by cellulose nanocrystals grafted with thermo-responsive polymer brushes. *J. Colloid Interface Sci.* **369**:202–209. http://dx.doi.org/10.1016/j.jcis.2011.12.011.