Norwegian University of Life Sciences
Faculty of Veterinary Medicine and Biosciences
Department of Chemistry, Biotechnology and
Food Science

Master Thesis 2015
60 credits

Enzymatic conversion of waste paper to mono- and oligosaccharides

John-Kristian Jameson

Acknowledgments

The present work was carried out at the Department of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Science, in the Protein Engineering and Proteomics group (PEP) with professor Vincent Eijsink and dr. Anikó Várnai as supervisors.

First, I would like to thank Vincent Eijsink for the opportunity to perform my thesis in his group and Anikó Várnai, my laboratory mother, who has been my guiding light during this eventful year. I am also very grateful to Johan Larsbrink, my laboratory father, for all your help and always finding time to answer my questions. Thanks to all the members at the PEP-group for your support and encouragement.

I would like to express my deepest gratitude to my girlfriend, Ingvild and my family for supporting me and pushing me trough this process.

_____

John-Kristian Jameson

Ås, August 15th 2015

Abstract

Efficient utilization of waste products is a key step in the transition into a sustainable society. Recalcitrant lignocellulosic biomass is a major constituent of industrial and household waste, and its conversion into fuels represents a promising green alternative to fossil sources. Lignocellulosic biomass is a complex composite structure, containing mainly cellulose, hemicelluloses and lignin; this intricate matrix provides tensile strength in the plant cell wall and resistance against enzymatic degradation.

There is a high demand for paper products in the world, and consequently, it is a major part of waste products. Cellulose represents the dominant component in paper and its conversion into monosaccharides and oligosaccharides could be a first step in generating value-added products such as fuels. Enzymatic cellulose degradation involves four main types of enzymes: endoglucanases (EGs), cellobiohydrolases (CBHs), lytic polysaccharide monooxygenases (LPMOs) and β-glucosidases (BGs).

This thesis focuses on the conversion of waste paper to monosaccharides with a commercial enzyme cocktail (Cellic CTec 2), as well as conversion of oligosaccharides using three purified endoglucanases. Experiments were conducted on shredded office paper (SOP), cut cardboard (CCB) and shredded newspaper (SNP), that were all pretreated by steam explosion.

A compositional analysis of SOP, CCB and SNP was performed in order to identify the ratios of the structural components. This analysis also enabled accurate determination of the conversion yields after enzymatic hydrolysis. The cellulose-hemicellulose-lignin-ash-others ratios were (41.4%-8.4%-2.6%-21.2%-26.0%), (58.7%-10.1%-16.3%-11.2%-3.5%), (31.6%-3.5%-39.9%-6.6%-24.1%) for SOP, CCB and SNP, respectively pretreated with steam explosion at 210°C for 14 min.

An additive in paper proved a challenge in the conversion using Cellic CTec2 because of its alkaline properties; experiments where the acid loading was varied were therefore performed, in order to achieve the optimal pH for the enzymatic reaction. Following the acid-loading test, a 24-hour hydrolysis was performed, 54%, 46% and 29% of theoretical maximum yields was observed with enzyme loadings of 5 mg/g dry mass (DM) over 24 hours for SOP, CCB and SNP respectively. Furthermore, optimization experiments achieved 89.6%, 55.2% and 56.3% conversions with enzyme loadings of 25mg/g DM and 72-hour incubation.

Three endoglucanases were produced using the eukaryotic expression host *Pichia pastoris*. The purified enzymes were able to produce a range of soluble cellodextrines from both pure cellulose and the waste paper substrates. Yields between 0 to 10% of the theoretical maximum yields were observed. Avicel, which contains highly crystalline cellulose, resulted in yields below 1.3%. Steam exploded SOP, CCB and SNP, on the other hand, resulted in yields of 7.3% in SOP with MaCel45A, 4.1% in CCB with MaCel45A and 13.5% in SNP with TaCel5A

In addition to studying the release of soluble products, the effect of the enzymes on insoluble cellulose was tested to better understand how the endoglucanases work. The results demonstrate how the endoglucanases reduce the degree of polymerization by introducing cuts on the insoluble cellulose chains. It thus seem that the endoglucanases hydrolyze the accessible regions of cellulose chain but not to an extent that leads to a high degree of conversion of soluble material.

ABBREVIATIONS

OD	Optical density
dH_2O	Sterile distilled water (milli-Q)
SOP	Shredded office paper
SNP	Shredded newspaper
CCB	Cut cardboard
SRS	Sugar recovery standard
SE	steam exploded
DM	Dry mass
SDS	Sodium dodecyl sulfate
CV	Column volumes
DP	Degree of polymerization
HPLC	High-performance liquid chromatography
BG	β-glucosidase
EG	Endoglucanase
GH	Glycoside hydrolase
LPMO	Lytic polysaccharide monooxygenases
CBM	Carbohydrate binding-module
CD	Catalytic domain
IEC	Ion Exchange Chromatography
SEC	Size exclusion Chromatography
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide gel electrophoresis

Contents

1 INTRODUCTION

Our civilization has become increasingly dependent on fossil fuels, negatively affecting the environment and climate of our planet. Research in recent years has focused on finding renewable energy sources such as wind, solar, waterpower, or bioenergy. With the synergetic effect from these various alternative sources, it may be possible to replace fossil fuels to obtain a sustainable environment in the future (Hughes 2000, Abbasi, Premalatha et al. 2011). Bioenergy is generated from biomass, which is the most promising renewable source for the production of liquid fuels. Biofuels are convenient as they could be mixed with gasoline for transportation purposes. Most bioethanol is currently produced from sugar-rich food stock material such as starch and sugarcane after fermentation. This approach is not considered sustainable, as food is a limited resource already. Lignocellulose is one of the largest renewable carbohydrate sources, available in all plant cell walls. This material is, exploited in some biomass-to-sugar conversion plants around the world, but this source should definitely be exploited more as it is capable of replacing the food stock biomasses for a more sustainable future.

There is a range of applications for lignocellulose as it contains fermentable sugars that could be utilized to produce ethanol, methane or hydrogen, and longer oligosaccharides for use in foods and pharmaceuticals (Patel and Goyal 2010). Lignin is also reckoned as a valuable resource when isolated. Today it is used in a range of industries, all from concrete to emulsion stabilizers.

Non- edible biomass is available at a “low” cost as waste materials from a higher valued production such as paper and cardboard (after consumption). For example, the amount of waste biomass from forest product industries could be large enough to provide enough energy to cover 3-5% of the electrical energy generated in the US (Hughes 2000).

Biomass could be converted to sugars and other products, using both chemically and enzymatic methods. The latter process is more environmentally friendly and hence, it is preferred. However, enzymatic saccharification of lignocellulose is a challenging

task as technical and economical issues hamper the development in this area. The conversion of lignocellulosic biomass is hindered by the complexity and limited accessibility of the material itself. Optimized cellulase cocktails and economically viable pretreatments are the most commonly studied approaches for overcoming the challenges in lignocellulose conversion (Wyman 1999).

This thesis will focus on how we could exploit typical waste products such as office paper, cardboard and newspaper in a better way than we currently do, by degrading them down to monosaccharides and oligosaccharides. Monosaccharides could be fermented for ethanol production or utilized as a source of food. Cellooligomeres have a high number of hydroxyl groups that could be derivatized in several locations, whereby their properties could be selectively modified for appliances in the food and/or pharmaceutical industries. (Patel and Goyal 2011). Non-digestible oligosaccharides are functional as a dietary fiber, drug delivery systems and regulators for blood glucose in diabetics. The vast array of applications has lead to worldwide interests and intensive research.

1.1 Lignocellulosic biomass

Lignocellulose is composed of three major components: cellulose, hemicellulose and lignin (plant cell wall). These polymers are ordered in a complex structure, which gives plants tensile strength and resistance against enzymatic degradation by e.g. plant pathogens (Kim, Block et al. 2010). As cellulose in plant cell walls is a structural component, and not storage carbohydrate, it has to be resistant.

The primary cell walls of plants consist of cellulose entangled with hemicellulose and lignin in a complex and irregular network. The ratio between these structural components varies significantly between different biomasses. For example, hardwoods stems contains 40-45% cellulose, 24-40% hemicellulose and 18-25% lignin. Plant leaves on the other hand, contains 15-20% cellulose, 80-85% hemicellulose and 0% lignin (Sun and Cheng 2002). In addition, plant biomass contains some extractives and pectin.

Biomass also contain minerals (determined as ash remaining after incineration); the ash content is generally low, but it still has to be taken into consideration. S, Al, Ca, K, Si, Mg, Fe, P, Cl, Na and Mn are a summary over components found, mainly concentrated in the lignin residue because lignin contain hydroxyl groups attracting minerals (Shen, Zhu et al. 2010, Kang, Appels et al. 2014, Pandey, Bhaskar et al. 2015).

1.1.1. Cellulose

Cellulose is the most abundant carbon-based renewable material on the planet, and it accounts for about half of the organic material in the biosphere.

Cellulose consists of unbranched β-1,4-linked homopolymers of glucose (Chami Khazraji and Robert 2013) and it can make up over 50 % of the polysaccharides in plants (Reusch 2013). The polymer has a strong affinity towards itself; cellulose chains align parallel to each other and form compact cellulose micro fibrils. This structural micro fibril exhibits a high degree of three-dimensional internal hydrogen bonding, resulting in a crystalline structure (Figure 1).

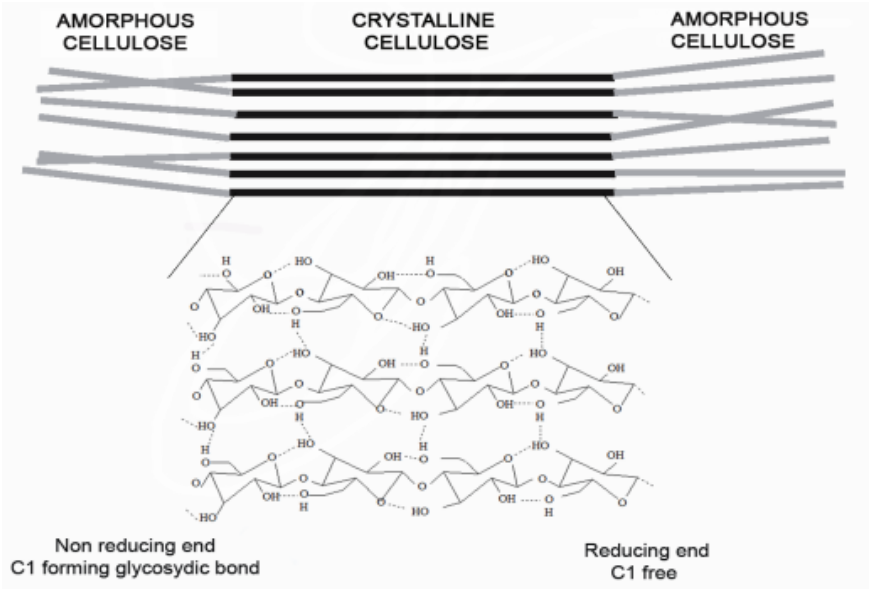


Figure 1. The structure of cellulose. *The upper figure illustrates the difference between crystalline and amorphous cellulose, whereas the lower figure shows hydrogen-bonding networks in the crystalline regions. The figure is adapted from* (Estella 2013)

This crystalline structure makes cellulose particularly resistant against enzymatic hydrolysis. There are, however, regions of the cellulose chain that have lost some of the intermolecular hydrogen bonding. These regions (Figure 1), in which cellulose chains are separated, are defined amorphous regions (Fengel and Wegener 1983). In amorphous regions, cellulose chains are accessible for specific enzymes called endoglucanases (EGs).

The strong intermolecular hydrogen bonding in cellulose renders it insoluble in water. However, if the chain length is reduced to seven or less glucose monomers it becomes water-soluble. Plant cell walls contain cellulose ranging from about 100 to 10000 glucose units per chain (DP=100-10000)(Zhang and Lynd 2005); hence, this cellulose is insoluble.

1.1.2. Hemicellulose

Hemicelluloses are heterogeneous polymers of hexaoses (glucose, mannose and galactose) and pentaoses (xylose and arabinose). Some hemicellulose are highly branched polymers, and hemicelluloses have a molecular weight lower than cellulose (Pérez, Munoz-Dorado et al. 2002). The most common hemicelluloses are arabinoxylan and galactoglucomannan.

Galactoglucomannans are the main hemicellulose in softwoods (10-25%(w/w)), while xylans make up about (10-15%(w/w)). (Willför, Sundberg et al. 2008, Varnai 2012). Glucomannan makes up about 3-5% (w/w) in hardwoods. Xylan (mostly arabinoxylan) are the main hemicellulose in hardwood, constituting between 20-30% of the dry mass (Sun, Sun et al. 2004). This hemicellulose constitutes a backbone with two free hydroxyl groups to which arabinosyl, glucuronic acid and acetyl substitutions can attach and create a complex structure that is hemicellulose (Sun, Sun et al. 2004).

The homopolymer of xylose is xylan and it consist of repeating units with two anhydrous β-D-1,4 linked xylopyranosyl residues. The homopolymer of xylan is very rare in higher plants and difficult to isolate. Xylans in higher plants have a backbone of β-1,4-linked xylopyranose sugars, some of which are substituted with acetyl groups and other sugars (see Figure 2) (Dodd and Cann 2009). Xylan normally contains 85-93% of D-xylose, with small amounts of L-arabinose and traces of

glucuronic acid residues as illustrated in Figure 2 (Bastawde 1992, Dodd and Cann 2009).

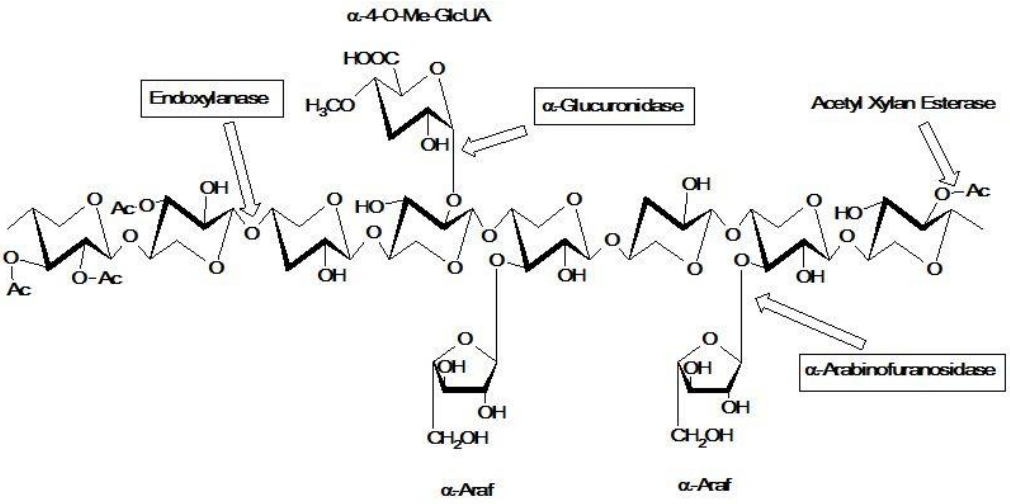


Figure 2 ***Structure of xylan.*** *Show substitutions comprising (4-O-methyl-) α-D-glucuopyranuronic acid and α-L-arabinofuranose units. The hemicellulases responsible for degradation of this hemicellulose are also illustrated. The figure is adapted from. (Motta, Andrade et al. 2013)*

Glucomannan is the most abundant hemicellulose in softwoods and the second most abundant in hardwoods (Timell 1967, Jacobs, Lundqvist et al. 2002). In woods, mannan normally contains about 5% galactose as galactomannnan (Moreira 2008, Oner 2014). Mannan is classified in four sub families: linear mannan, glucomannan, galactomannan and galactoglucomannan (see Figure 3). Each of these polysaccharides contain a β-1,4-linked backbone containing Mannose (linear mannan, top left Figure 3) or a combination of mannose and glucose (glucomannan top right Figure 3). If the backbone of mannan or glucomannan contains α-1,6-linked galactose residues, we get the two last forms, galactomannan and galactoglucomannan (Bottom left and right respectively, Figure 3)(Oner 2014).

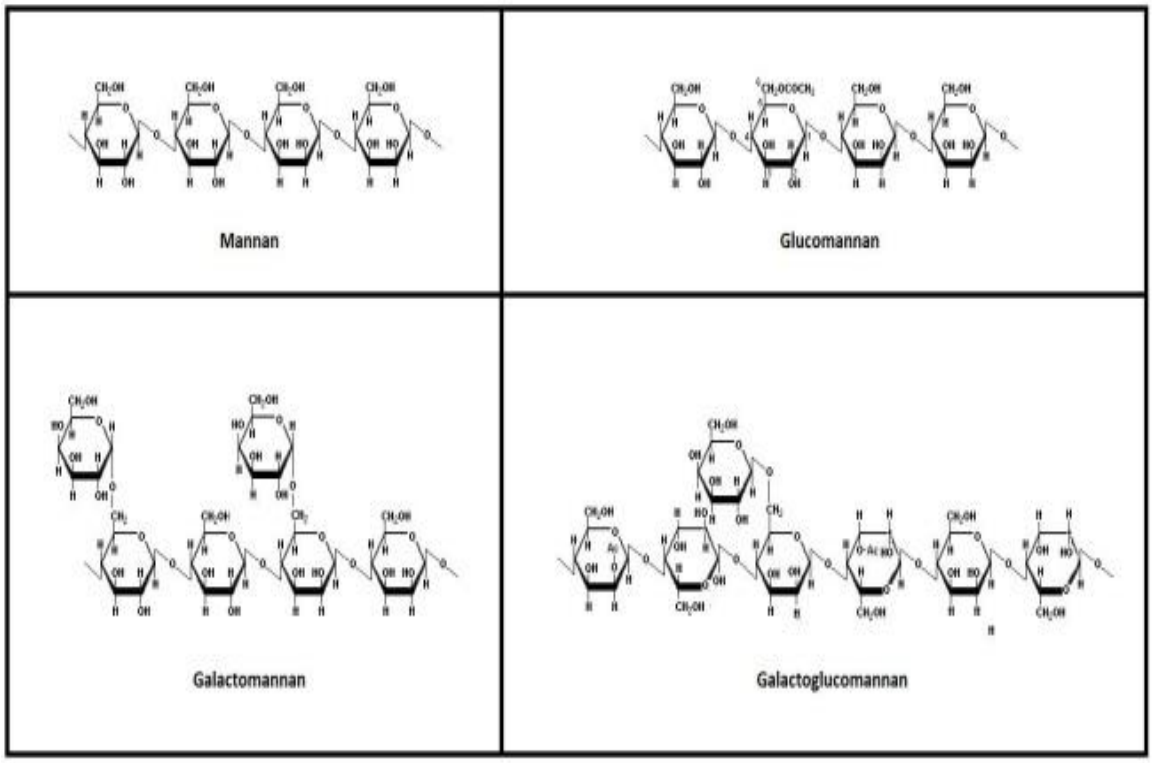


Figure 3. Four different forms of mannan*. Top left show the structure of mannan, Top right figure show the structure of glucomannan, bottom left panel show the structure of galactomannan, and the bottom right panel show the structure of glucomannan. The figure is adapted from (Oner 2014)*

1.1.3. Lignin

Lignin is a complex polymer of aromatic alcohols known as monolignols. The three common monolignol units are: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Albersheim, Darvill et al. 2010). These three monomers are methoxylated to various degrees and cross-linked with different chemical bonds (see Figure 4). Lignin makes up the complex structure that engulfs the carbohydrates (Richard 1996), which contributes strongly to making lignocellulosic biomass particularly resistant to biodegradation. It is believed that lignin constitutes a physical restriction for enzymes (Richard 1996).

Some organisms have developed the necessary enzymes to modify lignin´s structure, and to make the carbohydrates more available for degradation (Richard 1996). However, enzymes that result in complete monomerization of lignin is not identified so far.

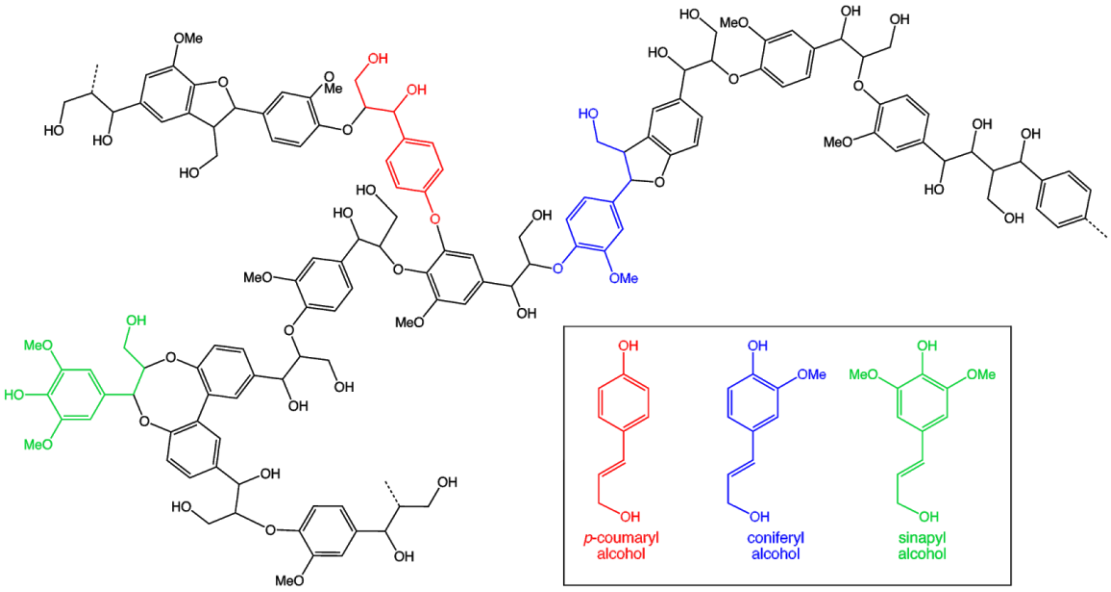


Figure 4 Structure of Lignin. *The figure illustrates the complex structure of lignin and the individual building blocs: p-coumaryl alcohol coniferyl alcohol and sinapyl alcohol* (Figure adapted from http://howaboutlignin.blogspot.no).

1.2 Classification of carbohydrate active enzymes

All enzymes are classified in the Enzyme Commission database (EC) controlled by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). The enzymes are provided with EC numbers based on the type of reaction that they catalyzes and their substrate specificity (Henrissat 1991). Enzymes for the hydrolytic degradiation of glycosyl bonds, glycosyl hydrolases (GHs), are provided with the numbers: EC. 3.2.1-X. the first three digits are generic for all GHs, i.e. the ability to hydrolyze O-glycosyl linkages. The X number indicates which substrate the enzyme targets and, in rare occasions, the molecular mechanism (Henrissat 1991). However, this method of classifying fails to identify enzymes with multiple substrate specificities and it does not account for evolutionary relationships that are visible in sequence and structure data.

In 1991 Bernard Henrissat introduced a new system for enzyme classification (Henrissat 1991). This classification system divides enzymes into families based on their structure-function relationship, instead of their substrate specificity. The enzyme families contain enzymes with similar amino acid sequence. Enzymes with similar amino acid sequences are likely to have catalytic sites with the same

structure and hence, substrate specificity. Furthermore, sequence similarity can indicate evolutionary relationships between enzymes. This formed The CAZY database (Henrissat and Davies 1997). Fungal EGs occur in the GH families 5, 6, 7, 8, 12, 44, 45 and 74. Bacterial EGs occur in families 5, 6, 8, 9, 10, 12, 26, 44, 45, 48, 51, 74 and 124. Fungal cellobiohydrolases (CBHs) occur in families 6 and 7, bacterial enzymes occur in families 5,6,9 and 48 (Henrissat 1991). β–glucosidases (BG) are in the families 1, 3, 5, 9 and 30 (Ketudat Cairns and Esen 2010). Lytic polysaccharide monoxoygenases (LPMO) is a recently discovered enzyme capable of promoting degradation in the most recalcitrant and crystalline regions of cellulose. The different types of cellulose-active enzymes are discussed in more detail, below.

1.3 Enzymatic conversion of lignocellulose

1.3.1. Cellulose degrading enzymes

Even though cellulose has a simple and linear structure, many enzymes are required to provide an efficient degradation. The classical scheme of degradation contains three classes of enzymes

- Endo-β-(1,4)-glucanases (EGs) (EC: 3.2.1.4) randomly cleave in amorphous regions on the cellulose chains. This generates reducing and non-reducing chain ends.
- Cellobiohydrolases (CBHs) are divided in two classes, one which attack the reducing ends (EC: 3.2.1.176), and one which attacks the non-reducing ends (EC: 3.2.1.91) of the cellulose chains. Cellobiose is the main end product. It is thought that some CBHs also can release longer products such as cellotetraose.
- β-glycosidase (BG) (EC: 3.2.1.21). Hydrolyses cellobiose and cellodextrines to glucose.

Recent discoveries have revealed a new class of enzymes called lytic polysaccharide monooxygenases (LPMOs)(Vaaje-Kolstad, Westereng et al. 2010). These enzymes contribute to the depolymerization of polysaccharides by introducing new chain-ends through oxidative cleavage of crystalline regions in cellulose. New chain-ends create access points for cellobiohydrolases. For efficient and total hydrolysis of polysaccharides it is necessary for all these enzymes to work synergistically. If the enzymes were to hydrolyze cellulose individually, a drastic reduction in activity would be observed (Beckham, Ståhlberg et al. 2014).

Figure 5 illustrates the process in which cellulose is degraded. EGs and LPMOs introduce new chain ends on the amorphous and crystalline regions respectively, illustrated in top left panel of Figure 5. CBHs attach to the reducing and non-reducing

chain ends and hydrolyzes processively (top right panel of Figure 5). Without the recently discovered LPMO´s, conversion is reduced as illustrated in bottom panels in of Figure 5.

BGs (Not included in Figure 5) cleave cellobiose, which is an inhibitor to CBHs. This synergy enhances the overall cellulolytic activity (Kostylev and Wilson 2012).

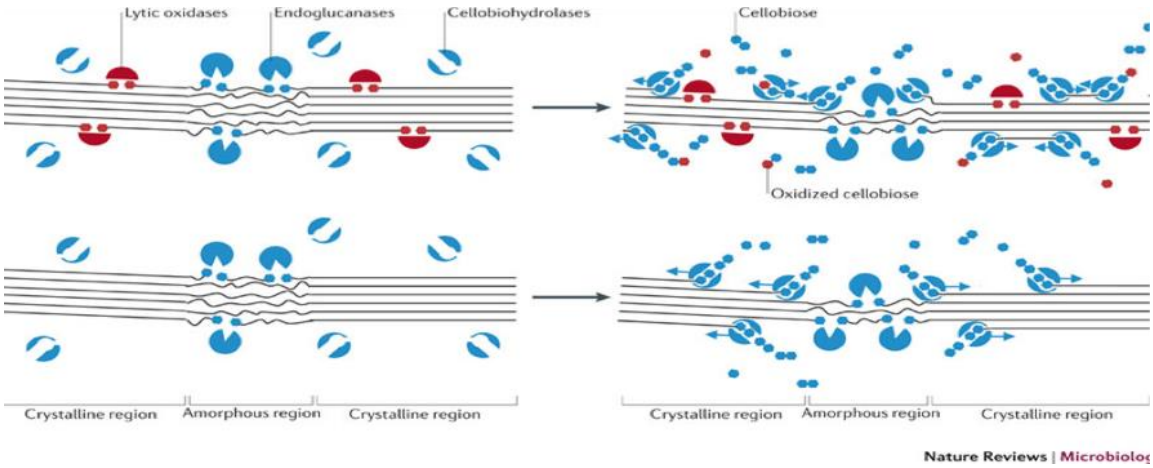


***Figure 5**. **Enzymatic degradation of cellulose by EG, LPMO and CBH**. LPMOs creates new chain-ends in the crystalline regions of cellulose and creates oxidized chain ends (red circles). EGs creates new chain ends in the amorphous regions(blue circles on the chain). Two types of CBHs hydrolyze the reduced ends and the non-reducing ends, respectively. They mainly release cellobiose, which is further hydrolyzed by BG to glucose (not illustrated here). (Medie, Davies et al. 2012)*

1.3.2. Active site topology

The shape of the catalytic domains in Glycoside hydrolases affects how many and what type of interactions the enzymes can make with the substrate. With multiple interactions the enzyme can hold on to the substrate and even slide along the chain in a processive manner, as observed in CBHs. Together with

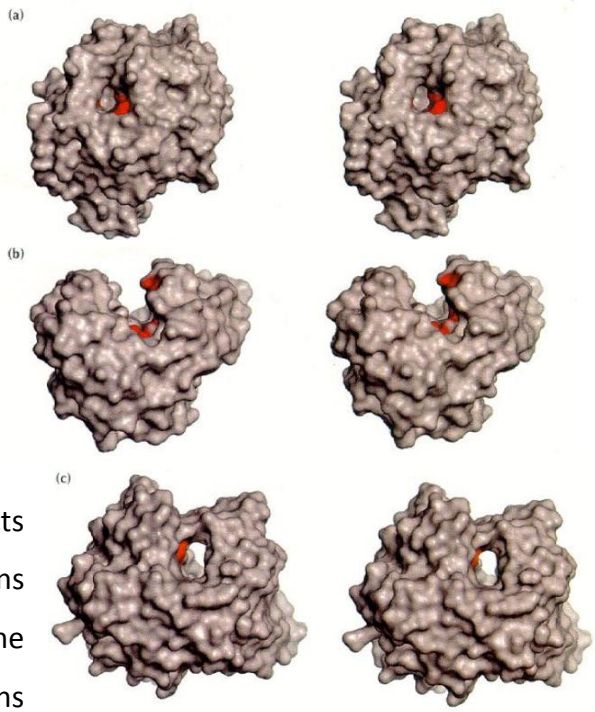


***Figure 6**. **Three types of active sites found in glycosyl hydrolases.** a) pockets typical for* β – glucosidases. b) Cleft typical for endoglucagenases. And c) the tunnel shape typical for cellobiohydrolases. *The figure is adapted from* (Davies and Henrissat 1995)

hydrogen bonding, aromatic residues are the most important interaction in protein-carbohydrate interaction (Sørlie, Zakariassen et al. 2012). Experiments have proven that mutations of certain conserved aromatic residues in the active sites of enzymes, reduces the binding affinity and activity (Li, Yan et al. 2012, Sørlie, Zakariassen et al. 2012).

The pocket topology seen in Figure 6a is adapted to short oligosaccharides substrates where there is a large number of available chain ends in the solution. The structure is typical for β – glucosidases and the effect of this protein on cellulose very low. However, when confronted with short and soluble cellodextrines, released by the CBHs. Pocket-type enzymes are indeed very effective as the substrate spends very little time in the active site (Davies and Henrissat 1995).

EGs have a substrate binding cleft or groove in their catalytic domain (Figure 6b). This open cleft enables the enzyme to lock on the cellulose chain through interactions with approximately 5-7 residues mainly through aromatic π- stacking and hydrogen bonds (Vlasenko, Schülein et al. 2010).

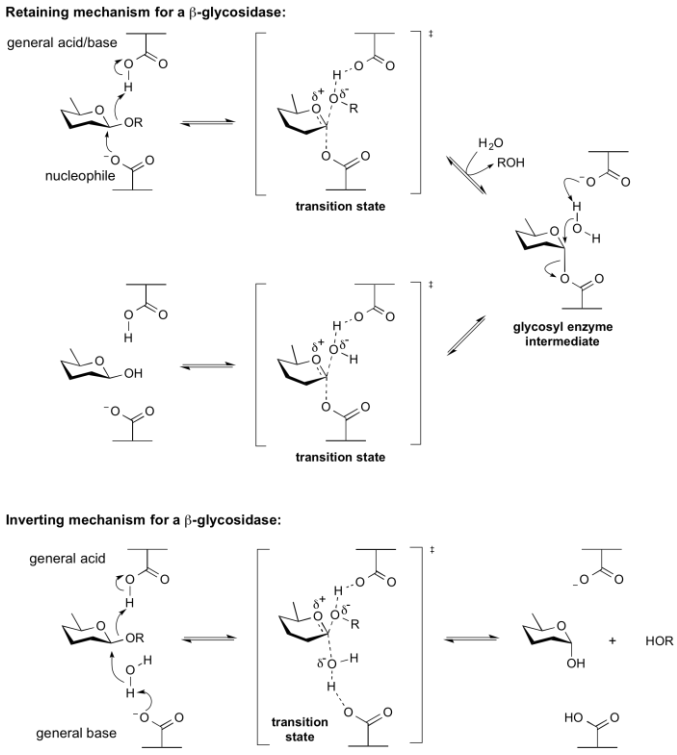
Ulike most EGs, CBHs are capable to work in a processive manner due to the shape of the active site (Figure 6c) (Teeri 1997). The enzyme could have a tunnel shape, but it could also have a deep cleft that includes more interactions to the cellulose chain than the cleft of endoglucanases. This topology enables the CBHs to stay connected to the cellulose chain and "slide" along the polymer. Through this mechanism, the detached chains are prevented from re-association with the crystalline material in between catalytic steps, which is considered favorable (Eijsink, Vaaje-Kolstad et al. 2008). Notably, it has been shown that this processivity comes with a risk that the enzymes might get "stuck" by obstacles on the cellulosic substrate. When a CBH molecule is retarded by obstacles on the chain, it has been found to reduce the processivity of other CBHs in close proximity on the chain, which enhances the "negative side" of processivity (Igarashi, Koivula et al. 2009, Varnai 2012).

EG´s increase the hydrolysis when working with CBHs, but have been observed to decrease the processivity as they contribute to the termination of the processive action if cuts are introduces in front of a processive enzyme active on the chain (Kipper, Väljamäe et al. 2005, Fox, Levine et al. 2012).

1.3.3. Mechanism involve in the degradation of polysaccharides

There are two types of mechanisms when hydrolyzing glycosidic bonds in cellulose, the retaining and the inverting mechanism. The individual endoglucanases studied in this thesis (*Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A; see below) utilizes the retaining mechanism. The main difference between the two mechanisms is whether they form a glycosyl-enzyme intermediate or not. The retaining mechanism maintains the configuration of the anomeric carbon (C1) after cleavage. Two carboxylic groups are required to interact with the polysaccharide in the first step of the hydrolysis. One of carboxylic groups donates a proton to the glycosylic bond and the other group acts as a nucleophile forming the glycosyl-enzyme intermediate (S_N2 reaction)(Vocadlo, Davies et al. 2001). Then, in the second step, the glucosyl-enzyme intermediate is hydrolysed by water when the carboxylic group that was deprotonated in the first step, acts as a base and deprotonate a water molecule(Agulia 2007)(Figure 7).

In the inverting mechanism, the enzyme changes the conformation of the anomeric compound carbon, from beta to alpha or vice versa. This mechanism is a one-step process, in which there is no glycosyl-enzyme intermediate. One of the amino acids acts as an acid, donating a proton to the glycosidic bond. The other aminiacid acts a base deprotonating a water molecule as it attacks (Agulia 2007) (Figure 7). Another key difference between the two mechanisms is that the catalytic water molecule approaches



Figur 7.Inverting and retaining mechanism. *Retaining (upper panel) and inverting mechanism of cellulases(lower panel). The figure is adapted from (Williams 2013)*

the scissile bond from a side that is similar to (retaining) or opposite of (inverting) the position of the glycosidic oxygen.

The processive mechanism of CBHs includes multiple steps. First the CBH binds to the cellulose followed by a decrystallization of the cellulose from the surface. As the cellulose gets released, the cellulose chain “slides” into the catalytic tunnel. This leads to the formation of the enzyme-substrate complex. As the cellulase slides, the cellulose gets cleaved into cellobiose (Gao, Chundawat et al. 2013).

1.3.4. Carbohydrate-binding modules

Most cellulolytic enzymes contain at least two domains, a Catalytic domain (CD) that performs the hydrolysis, and a carbohydrate-binding module (CBM). They are connected through a flexible linker and the CBM potentiates the activity of the CD. CBMs are divided in to 71 families in the CAZY database.

CMBs are thought to have a couple of putative functions to enhance the catalytic domain´s activity: 1) they effectively increases the concentration of the enzymes on the substrate by maintaining a close vicinity to the carbohydrates. 2) CBMs can target specific substrates that are compatible to the catalytic site. As a result of these functions it has been concluded that CBMs have some positive effects on conversion; they seem more vital to CBHs, because EGs sometimes lack a CBM (Boraston 2004, Fox, Levine et al. 2012). In recent research, (Igarashi, Koivula et al. 2009) have concluded that the CBM enhances the concentration of the protein on the substrate, and that it dose not seem to have any other role. For example, it was shown that the CBH travels on the cellulose chain with the same speed without the CBM.

1.4 Hemicellulases

Hemicellulases are a large and diverse group of enzymes that degrades hemicellulose. The complex structure and varying building blocks of hemicellulose requires the synergetic work of multiple enzymes for their degradation. Hemicelluloses may be highly branched polysaccharides and side groups tend to hinder the action of hemicellulases sterically, thus hindering complete hydrolysis. Xylanases and mannanases are the most important depolymerizing hemicellulases as xylan and glucomannan are the predominant hemicellulose.

Xylans are solubilized by endo-1,4-β-xylanases (EC 3.2.1.8). This conversion yields Xylo-oligosaccharides (Shallom and Shoham 2003), which could be further hydrolysed to monomers with a β-xylosidases (EC 3.2.1.37)

Endo-1,4-β-mannanase (EC3.2.1.78) cleaves glucomannan and yields oligosaccharides. Manno-oligosaccharides could be hydrolyzed with β-mannosidase or exo- β-mannanase (EC 3.2.1.25). To hydrolyze the glucose components in glucomannan, β-glucosidases is necessary (Ganter, Sabbi et al. 2001).

Both xylanases and mannanases face the problem of restricted accessibility of the main chain as the substituted groups create a physical barrier (Tenkanen, Luonteri et al. 1996, Ganter, Sabbi et al. 2001, Kolenová, Vršanská et al. 2006). However, this hindrance might be reduced if the residues in the vicinity of the cleavage point are all on the same side of the backbone(de Vries and Visser 2001). To achieve a complete hydrolysis of branched xylans, debranching enzymes are required such as α-glucuronidase (EC 3.2.1.139) and α-L- arabinofuranosidase (EC 3.2.1.73), as visualized in figure 2. These enzymes work synergistically with endo-1,4-β-xylanases. Conversion of galacto(gluco)mannans may also require enzymes such as acetyl mannan esterase (EC 3.1.1.6) and α-galactosidase (1,6-α-d-galactoside galactohydrolase, EC 3.2.1.22) co-operating in the removal of side groups that hinder hydrolysis by mannanases (Moreira and Filho 2008).

1.5 Pretreatment

1.5.1. Modification of recalcitrance and complexity in biomass

Lignocellulosic biomass consists of complex and inhomogeneous networks of polymers, as described above. The ordered structure of cellulose and the heterogeneity of plant cell walls make biomass resistant against glycoside hydrolases without any form of pretreatment. Consequently, pretreatments of lignocellulosic substrates are necessary to loosen the impediments in the biomass and obtain reasonable yields in enzymatic conversion (Alvira, Tomás-Pejó et al. 2010). The goals of pretreatments are to break down the lignin and hemicellulose structures, while disrupting the crystalline structure of cellulose (Mosier, Wyman et al. 2005). The use of high temperatures, acids, alkali, salts and oxidants has an effect on the lignocellulose structure rendering the biomass susceptible to the action of cellulases. There has been a lot of focus on pretreatment research because it seems to be one of the solutions for enzymatic conversion of biomass to be economically viable (Sun and Cheng 2002). Changes in the plant cell wall structure and the efficiency of subsequent enzymatic hydrolysis processes strongly depend on the type of pretreatment.

Table 1. Pretreatments. *List of some pretreatment options, and their effect on reducing resilience in lignocellulosic material. The table is based on (Alvira, Tomás-Pejó et al. 2010)*

	Milling	Steam explosion	Acid	Alkaline	Oxidative
Increases accessible surface area	H	H	H	H	H
Cellulose decrystallization	H	–	–	–	n.d.
Hemicellulose solubilization	–	H	H	L	–
Lignin removal	–	M	L	H	M
Generation of toxic compounds	–	H	H	L	L
Lignin structure alteration	–	H	H	H	H

H: high effect; M: moderate effect; L: low effect; n.d: not determined.

Pretreatment methods (Table 1) are commonly divided into three groups: physical (boiling, steam explosion or milling), chemical (including acid or alkali treatment), and biological (microorganisms) (Mehdi Dashtban 2009).

Physical pretreatments utilize high temperatures and physical forces to shred the biomass and make it more accessible. Physical pretreatments are cost effective and they are effective in increasing surface area to generate access for the enzymes. There are some disadvantages with physical pretreatment, such as difficulties to industrialize milling processes and inhibitor formation during steam explosion.

Chemical treatments use acids and alkaline, which primarily solubilize hemicelluloses and remove lignin (Table 1). This has proven to results in high glucose yields. However, use of acid and alkaline chemicals is connected with high costs and the processes often generate inhibitors for down stream hydrolysis steps. Microorganisms used for pretreatments produce enzymes, which selectively degrade lignin and hemicellulose. This has some advantages because they do not require any harmful products such as acids and alkaline pretreatments, but microorganisms are very inefficient in this type of processes. Since pretreatments have different strengths and weaknesses, there is no general pretreatment for all biomass; optimal pretreatment conditions may vary with the type of substrate.

In the work described in this thesis, the biomass, wastepaper was treated with steam explosion. Notably, the production of paper and cardboard entails a pretreatment of the original lignocellulosic biomass. Steam explosion and paper production will be described in detail below.

1.5.2. Steam explosion

Steam explosion is one of the most common pretreatment methods for reducing recalcitrance in lignocellulose. Steam explosion utilizes high temperature and pressure to make the biomass more accessible for subsequent processes such as fermentation or enzyme hydrolysis (Alvira, Tomás-Pejó et al. 2010). The temperatures can range from 150 to 240°C with pressures ranging from 1 to 3.4 MPa. The heating step is followed by an explosive decompression of the biomass that makes the fibers rupture. This process results in significant hydrolysis of glycosidic bonds in hemicellulose, and to a much lesser extent, in cellulose. It also leads to a cleavage of hemicellulose-lignin bonds (Li, Henriksson et al. 2007).

Furthermore, the drop in pressure under these intense conditions defibrillates cellulose and results in increased levels of solubilized polymeric hemicellulose. Lignin is redistributed and some is removed, which in turn leads to more accessible cellulose (Alvira, Tomás-Pejó et al. 2010). De Risio et al. ((Di Risio, Hu et al. 2011), states that pretreatments around at 200°C for about 8 min lead to the most efficient hydrolysis. They showed that higher severities (increased temperature, pressure and residence time) decompose carbohydrates, which in turn leads to inhibitor formation and decreased sugar production in hydrolysis. Subsequent to steam explosion, oxidative treatment could lead to complete delignification of the biomass (Li, Henriksson et al. 2007)

Advantages of steam explosion, as compared to other pretreatment technologies is that it requires no chemicals, so there is less corrosion on equipment, and that there are no extra costs related to the use of acids and alkaline chemicals.

1.6 Production of paper and cardboard

Wood (lignocellulosic material) is the primary raw material for the paper industry, and paper is often made by a variety of hardwoods (Oaks, Beeches and Eucalyptus) and softwoods (Pine and Spruce) with different characteristics to create particular types of paper and cardboards. Hardwood contains short cellulose fibers, which creates a more complex, dense and robust wood. Softwood, on the other hand, contains long cellulose fibers (Shackford 2003). The long fibers create a more robust but less smooth paper. This is practical in cardboards used for transportation purposes. Officepaper contains the long fibers from softwood but short fibers from hardwood is mixed in to fill in the areas between the long fibers and create a smoother and more dens paper (Martin, Anglani et al. 2000).

Paper production involves several steps from the lignocellulosic raw material to a finished product. Most steps in this process are considered pretreatments as they inflict the composition and structural integrity of lignocellulose. The first step in the production is to debark the wood logs, following debarking, the logs are chipped in to small fibers. The next stage is pulping and there are three main pulping processes:

1 Mechanical pulping

2 Chemical pulping

3 Semi-chemical pulping.

The primary purpose of pulping is to free the cellulose fibers from lignin and hemicellulose that holds them together (Martin, Anglani et al. 2000). Paper with low lignin content is known to be stronger and to last longer. Lignin is renowned within the papermaking industry because it discolors paper as it is easily oxidized (Manning et al. 2000). This is observed in newspapers, which turn yellow over longer periods of time. The success of paper is also dependent on the relative abundance of hemicellulose. Papers with high content of hemicellulose are known to tear easily but the pulping process removes them easily by virtue of hemicelluloses relatively labile nature (Hubbe and Lucia 2007).

Mechanical pulping involves grinding the wood chips down to individual fibers; this process leads to short fibers and lots of impurities in the pulp. This process cannot remove lignin, so it is not used for high quality products. Mechanical pulping has been replaced by chemical pulping over the years, but it is still used for low grading papers such as newspaper and recycled paper for economical reasons. When some lignin is left in the pulp in increases the total yields of paper from a given amount of wood.

Chemical pulping (kraft pulping) involves mixing highly alkaline solutions such as sodium hydroxide and sodium sulfide with the woodchips at high pressure and temperatures. This process is very effective for removing lignin, which is filtered out after alkali dissolves the lignin structure. Compared to mechanical pulping chemical pulping yields longer fibers and a purer product that is almost free of lignin and hemicellulose. The yields are low (45-55%) compared to mechanical pulping (90-95%), but the demand for the resulting high quality office paper is high (Martin, Anglani et al. 2000). Chemical pulping is often combined with addition of pigments. Office paper contains calcium carbonate as a coating pigment, giving the paper its characteristic white color. The amount of calcium carbonate may vary significantly depending on the producer, but it is estimated that a typical European fine office

paper contain 15-20% calcium carbonate (Brander and Thorn 1997). Calcium carbonate has alkaline properties when solubilized.

Semi-chemical pulping is a combination of the former two processes. First the woodchips are chemically pretreated before they are mechanically pulped.

2 Aims of the study

Efficient utilization of waste products is a key step in the transition into a sustainable society. Lignocellulosic biomass is the greatest source of renewable carbon on the planet, and consequently a major part of society's waste products. The aim of the current study was to explore possibilities for conversion of major lignocellulosic waste product, paper into value added products by enzymatically converting the cellulose into monosugars and/or cellodextrines.

The experimental work done was divided into three main parts:

- Composition analysis of lignocellulosic materials, SOP, CCB and SNP before and after pretreatment.
- Studies on the conversion of the various paper feedstocks to monosaccharides with the commercial enzyme cocktail Cellic CTec2 and optimizing experiments to get as close to a 100% of the maximum theoretical possible yield.
- Studies on the conversion of various paper feedstocks to cellodextrines with recombinantly expressed, purified endoglucanases.

Furthermore, the effect of endoglucanases on insoluble cellulose was tested to better understand how the endoglucanases work and affect the degree of polymerisation

3 Materials and methods

3.1 Chemicals and substrates

3.1.1 Chemicals

Chemicals	Supplier
Acetic acid 99.8%	Merc
Arabinose	Sigma-Aldrich
Bacto™ Pepton	Becton, Dickinson and company
Bacto™ yeast extract	Becton, Dickinson and company
Cellobiose	Sigma-Aldrich
Cellotriose	Megazyme
Celloteteraose	Megazyme
Cellopentaose	Megazyme
Cellohexaose	Megazyme
Citric acid	Sigma-Aldrich
Ethanol 96% (v/v)	Arcus
Galactose	Sigma-Aldrich
Glucose	Sigma-Aldrich
Mannose	Sigma-Aldrich
Potassium sodium tartrate tetrahydrate	Sigma-Aldrich
2-hydroxy-3,5-dinitrobenzoiz acid	Sigma-Aldrich
Sodium acetate	Sigma-Aldrich
Sodium chloride	Prolabo
Sodium hydroxide	Merc
Sodiumdodecylsulfate (SDS)	Applichem
Sulfuric acid	Sigma-Aldrich
Xylose	Sigma-Aldrich

3.1.2 Carbohydrate substrates

The following Substrates were utilized in the thesis:

- Shredded office paper (SOP)
- Cut cardboard (CCB)
- Shredded newspaper (SNP)
- Avicel

The three lignocellulosic substrates listed are waste biomasses containing various amounts of carbohydrates, lignin and other extractives. The lignin content may vary from high (SNP), to medium (CCB) to low (SOP) (see section 4.1 for more detail). In addition, microcrystalline cellulose (Avicel) was utilized as a reference material in some experiments. The substrates (except Avicel) were pretreated with steam explosion. SOP and CCB were pretreated with the following conditions: 170 ^{o}C for 7 min, 170 ^{o}C for 14min, 190 ^{o}C for 7min, 190 ^{o}C for 14min, 210 ^{o}C for 7min and 210 ^{o}C for 14min. SNP was pretreated as follows: 170 ^{o}C for 7 min 210 ^{o}C for 7min and 210 ^{o}C for 14min. Pretreatment was performed by Aniko Varnai in connection to a separate project

SOP was collected from the office printer. CCB was collected from cardboards around the office and newspaper was from Aftenposten (August, 2014). All substrates were cut down to small pieces with a scissor (≈ 0.5 x 0.5cm). Avicel is assumed to contain 100% cellulose in this study.

3.2 Compositional analysis

The carbohydrate composition, the lignin content and the ash content of the pretreated and un-pretreated substrates (SOP, SNP and CCB) were determined in a 2-step acid hydrolysis. A protocol developed by Sluiter and coworkers at the U.S Department of Energy (Sluiter, Hames et al. 2008), was used to accurately quantify structural polysaccharides such as cellulose and hemicellulose. The same protocol was also used to determine soluble and insoluble lignin

In the first step, the samples were incubated with 72% (w/v) sulfuric acid to solubilize carbohydrates. In the second step, water diluted the sulfuric acid to a concentration of 4% (w/v) followed by elevated temperatures, which hydrolyzes the released oligosaccharides from the first step, down to sugar monomers.

During the second hydrolysis step, some of the sugars may be degraded; furfural and 5-hydroxymethylfurfural are formed in the dehydration reactions of pentoses and hexoses respectively (Dunlop 1948, Ulbricht, Northup et al. 1984). To account for the loss of sugars in this procedure, the use of sugar recovery standards (SRS) was essential. SRS were sugar solutions with pre-determined concentration of all carbohydrates (glucose, xylose, galactose, mannose and arabinose), and they were autoclaved together with the samples in separate pressure tubes. These external standards were subjected to the same conditions as the substrates during the second hydrolytic step of the analysis and the sugar decomposition can be estimated to be directly comparable in the samples and in the external standards. Following autoclavation, HPLC was used to identify and quantify carbohydrates in the liquid fractions.

Lignin exists in two states after the two-step acid hydrolysis, acid soluble and non-soluble. Acid soluble is measured by UV spectroscopy, and insoluble lignin is determined gravimetrically after filtration.

The biomass determination was performed for all three substrates (SOP, CCB and SNP) with and without the different pretreatments, to observe how the composition was affected by the high pressure and temperatures in the steam explosion process.

In order to get reliable results in this method, the samples have to meet certain criteria. The dry matter content must be over 90% (Sluiter, Hames et al. 2008) because higher moisture content will lead to dilution of the strong acid and reduce the hydrolysis efficiency. This procedure is not suitable for samples with high protein content, because proteins will interfere with the lignin quantification (proteins have UV absorbance). Furthermore, it was essential that the particle size was small and uniform, so that the acid hydrolyzes all particles equally.

Apparatus:

- Scale with 0.1 mg accuracy (Sartorius, Germany)
- Water bath 30 ^{o}C
- Autoclave 120 ^{o}C (Certoclav sterilizer, Austria)
- 50 ml liquid dispenser
- Vacuum pump
- 100^{o}C dry oven
- Desiccator
- Muffle furnace (oven to burn organic samples at 550^{o}C)
- HPLC (High-performance liquid chromatography)
- UV-visible spectrophotometer Hitachi U-1900 and high purity quartz cuvettes.

Materials:

- Pressure tubes (50ml, black tops with sealing)
- Glass stir rods
- Glass filter crucibles 15ml ROBU-GLAS filters porosity 4
- Büchner flask and a funnel fitting the filters
- HPLC vials with top seals
- Sulfuric acid 72%
- dH_2O

- SRS samples, containing known amounts of all carbohydrates (glucose, xylose, galactose, mannose and arabinose) to check for loss of monosaccharides during the process.

Reagents:

- 72% w/w H_2SO_4 (specific gravity 1.6338 at 20 ^oC)
- High purity monosaccharide standards for HPLC, D(+)-glucose,D(+)-xylose, D(+)-galactose,L(+)-arabinose, and D(+)-mannose.
- 15 mM NaOH for sample dilution
- 4% H_2SO_4 (w/w)
- dH_2O

Procedure:

All tests were run in triplicates unless described otherwise. About 150mg biomass (dried to ≈95% DM (dry mass) and the sample size was as described in section 3.1.2) (all pretreated and non-pretreated samples) was weighed in to pressure tubes with 0.1 mg precision. This was measured on a milligram scale to get a precise weight. 1.5 ml of H_2SO_4 was added to each pressure tube so the entire sample was soaked in acid. Pressure tubes were incubated in a water bath at 30^oC for 60 minutes while the samples were stirred around about every 5 minutes with glass rods, this is crucial so that the all parts of the sample were wetted in acid. This incubation was followed by diluting to 4% H_2SO_4 by adding 42ml dH_2O with a liquid dispenser.

SRS samples contained glucose, xylose, galactose, mannose and arabinose within the concentrations of 0.0025 to 0.04g/L. SRSs were made in triplicates by adding 5 ml of SRS stock and 174 µl H_2SO_4 into three pressure tubes which follows the same procedure as SOP, CCB and SNP samples after the first step (see above). This gave information about any loss of carbohydrates during the process (when compared to

untreated SRS samples). All the pressure tubes (including the SRS samples) were sealed and placed in an autoclave at 121^{o}C for 60 minutes.

After cooling down the samples, they were filtered through a pre-weighed filter crucible. The vacuum setup consisted of a Büchner flask and a funnel fitting for the filters. The filters were placed on top and a falcon-tube was placed in the flask to collect the filtrate. It was important that, prior to filtration, these filters were incinerated in an oven at 550^{o}C for an hour to get the correct dry mass and to remove any contaminants.

The filters were numbered and rinsed with dH_2O after filtration to remove all soluble samples and dried over night at 100^{o}C (dry filter crucible (a)).

The liquid filtrates were prepared for monosaccharide analysis using 15mM NaOH as diluting agent. 50 and a 100-fold dilution were produced so the concentration was within the standards covering the range 0.002 too 0.4 mg/ml. HPLC was then utilized for the separation and quantification of the monosaccharides from cellulose and hemicellulose (see section 3.10.2.1 for specifications). Two vial from each liquid filtrate was analyzed on HPLC, one with 50 dilution and one with 100 dilution (could not run triplicates, it would have been to many samples). The procedure includes a polymer correction factor, as the released monosaccharides acquire an H_2O molecule (increase in mass) in the depolymerization process during acid hydrolysis. SRS samples without treatment were analyzed and compared against the treated samples to get a correction factor for loss of carbohydrates.

The filtrate also contains soluble lignin, whereas the filter holds back insoluble lignin. The insoluble lignin was measured by placing the filter crucibles back in to the muffle furnace at 550^{o}C for one hour (incinerated filter crucible (b)). The mass difference between the dry filter crucible (a) and the incinerated filter crucible (b) was used to calculate the insoluble lignin.

Soluble lignin was measured with a UV spectrometer. The filtrate was diluted 10 times in 4% w/w H_2SO_4 and the absorbance at 203nm was recorded and the amount of lignin was calculated from an adsorption factor 110L/g.cm. ASL stands for acid insoluble lignin in the following equation. a is adsorptitivity factor: in general, 110L/g.cm

$$-ASL\left(\frac{g}{L}\right) = \frac{Abs(203nm)}{1cm*a} * Dil = \frac{Abs(203nm)}{1cm*110\frac{L}{g*cm}} * Dil$$

Total ash was measured by incinerating 150 mg of DM-substrates in filter crucibles, for 1 hour. The ash, which was left after incineration, was composed of inorganic paper additives and other minerals because all the organic materials were incinerated.

3.3 Enzyme overview

Table 2 show and overview over the enzymes expressed in *pichia pastoris* in this thesis. All three endoglucanases are fungal enzymes from organisms listed.

Table 2. *Information about endoglucanases used in this thesis (Af*Cel12A*, Macel45A and Tacel5A) is listed. Pichia pastoris was utilized as an expression host for all enzymes. P. pastoris contained a pPink-HC-vector where expression of the cellulose gene is driven by a constitutive GAP-promotor. 1g/L equal Abs, is defined as the absorbance at which there is 1 g enzyme in 1 Liter solution.*

	*Ma*Cel45A	*Ta*Cel5A	*Af*Cel12A
Organism, Family	*Melanocarpus albomyces* (fungal) GH45	*Thermoascus aurantiacus* (fungal) GH5	*Aspergillus fumigatus* (fungal) GH12
EC	3.2.1.4 EG	3.2.1.4 EG	3.2.1.4EG
pI	4.97	4.47	5.69
Ext.Coefficients 1/(M.cm)	41325	86985	74750
Uniprot id	Q8JOK8	Q8TG26	Q4WGT4
1g/L equal Abs	1.804	2.478	3.211
Molecular weight (g/mol)	22913.2	35106.1	24128.4
number of amino acids	214	317	218
CBM	No	No	No
Signal peptide	Native	Native	Native
Protein expression	Extracellular	Extracellular	Extracellular

Cellic CTec2 (NS-22086, PPC 303604 (pre-prep)) is commercial multi-enzyme cocktail (Novozyme, Denmark) with a protein concentration of 57.0 mg/ml. This cocktail includes powerful cellulases, high levels of β-glucosidases, LPMOs and hemicellulases for an effective hydrolysis of all carbohydrates (Novozymes 2015).

3.4 Protein expression

3.4.1. Agars and cultivation media

Agar plates:
Premade Pichia Adenine purchased from Technova. Dropout selection agar plates Contain 2% glucose, 1% ammonium sulfate and no Adenine.

Cultivation media:

Yeast extract peptone dextrose (YPD) liquid medium was prepared as follows:

Materials:

10 g (1% w/v) yeast extract

20 g (2% w/v) peptone

100 ml 20% (w/v) glucose, filtered sterilized

Procedure:

Yeast extract and peptone were dissolved in 900ml dH_2O followed by sterilization; all media were sterilized by autoclaving at 15 psi (1 bar) and 121^{o}C for 15 minutes.

After it cooling to room temperature, 100 ml glucose solution added. As the media do not contain antibiotics, sterile working conditions were essential.

3.4.2. Long-term storage of microorganisms

For long-term storage of microorganisms, glycerol is added to ensure survival of cultures while they are kept at -80 °C. A new glycerol stock of the strains expressing endoglucanases was prepared; from such a stock the strains were spread on an agar plate when needed.

Material:

- 1 ml overnight cultures
- 300 µl glycerol, sterile (85% (w/v)
- Cryo-tubes

Apparatus:

- Sterile cabinet. 30°C with shaking tray

Procedure:

Plates containing single colonies of *P. pastoris* overexpressing the endoglucanases *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A were made from an existing glycerol stocks in the lab. Following a night of incubation at 30°C, Single cultures were transferred to an overnight culture of YPD medium with autoclaved toothpicks. The overnight culture was incubated in a sterile cabinet at 30°C with 220 rpm on a shaking tray. 1 ml from this culture was mixed with 300µl 85% (w/v) glycerol and after mixing stored at -80 °C. New cultures were made by scraping a small amount of the frozen cultures on to an agar plate, followed by incubation over night at 30 °C. Single cultures was picked with a toothpick, and transferred to the appropriate growth medium for further cultivation

3.4.3. Protein expression with *P. pastoris*

P. pastoris was the chosen expression host because it combines important properties of higher eukaryotes, such as protein processing, folding and post translation modifications, while being as easy to manipulate as *Escherichia coli*. Furthermore, *P. pastoris* has the ability to grow on simple inexpensive media; it can grow in either shake flask or fermenters, which makes it suitable for small productions in laboratory research, and large-scale industrial use. *P. pastoris* is a methylotropic yeast that can utilize methanol as its sole carbon source. The promoter AOX1, is activated by methanol, and any gene downstream of this promotor will be produced in high yields when the alcohol is added. However, the use of methanol leads to high costs and strict routines, so the glyceraldehyd-3-phosphate dehydrogenase promoter (P_{GAP}), has been increasingly used for high-end production. P_{GAP} is the promotor to a key enzyme in the glycolysis (GAPDH), which provides a strong constitutive expression on glucose at a level comparable to that seen with $P_{AOX.}$ (Qin, Qian et al. 2011)

In *Pichia*, heterologously expressed proteins are either concentrated in the intracellular space of the cell or secreted (extracellular). All enzymes in this thesis were expressed extracellularly, so there was no need to preform lysis of the cells to extract proteins. An advantage of *P. pastoris* is that it secretes very low levels of native proteins. So, secreted heterologous protein comprises the vast majority of the total protein in the medium, providing a good starting point for further purification.

Materials:

- YPD, see section 3.4.1
- Antifoam
- Filtropur S 0.45µm syringe filter (Sarsted AG & CO, Germany)

Apparatus:

- Sterile cabinet. 30°C with shaking tray
- Lex-48 bioreactor (Harbinger, Canada)
- Avanti J-25 Centrifuge (Beckman Coulter USA) with a JA-10 rotor O-ring

Procedure:

Samples of *P. pastoris* retrieved from glycerol stock cultures in storage (-80°C) (see section 3.4.2) were grown on Pichia Adenine Drop-out selective agar over night at 30°C. To start a new culture, single colonies from each strain (see section 3.3) were retrieved with a toothpick from the selective agar and incubated in 5 ml YPD-medium at 30°C (two parallels) on a shaking tray (220 rpm) in a sterile cabinet

After 24 hours, the overnight cultures were transferred to a 1 L blue-capped bottle containing 650 ml YPD medium. A Harbinger-system was used instead of shaking flasks. This system provides oxygen rich environments for optimal growth conditions over several days. Small yields were observed after 24 hours, so longer incubation time (48 hours) was tested. The YPD medium contains 20% (w/v) dextrose. If the inoculation lasts more than one night, dextrose was added to approximately 20% (w/v) to maintain the growth. In experiments trying to increase the yields, 30% (w/v) dextrose was tested.

When the culture reached the target of 24 or 48 hours (both were tested for optimal growth), the culture was transferred to 500ml centrifugal bottles and centrifuged at 8000 rpm for 8 minutes. The supernatant was subsequently filtered through a Filtropur S 0.45µm syringe filter. The supernatant was stored at 4 degrees until purification.

3.5 Protein purification

Purification was performed for endoglucanases *Ta*Cel5A, *Af*Cel12A and *Ma*Cel45A produced in *P. pastoris*, in this study. Purifications were also performed using culture filtrates donated by A. Varnai from a separate project (Várnai, Tang et al. 2014). The purification procedure was as follows:

3.5.1. Ultrafiltration

The secreted protein from *P. pastoris* was up-concentrated and rinsed with buffer to remove salts and other micro molecules. Concentration was achieved by using a tangential flow-filtration through a cassette with a 10kDa cut-off membrane (Sartorius Stedim Biotech GmbH, Germany). All the fluid from the supernatant was pumped through the cassette (illustrated in figure 8), molecules smaller than the pore size of the membrane, flow through as waste. The proteins of interest are in the retentate (do not cross the membrane) and recycled. This is a continuous process, which leads to concentration of all macromolecules from the broth. Figure 8 shows an illustration of the setup for ultrafiltration.

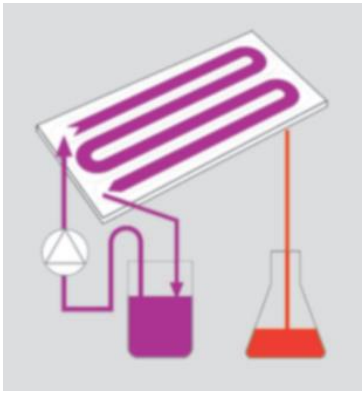


Figure 8. **Setup ultrafiltration.** Figure is adapted from *(Sartorius 2000). The purple bottle contains the broth at which the protein was located. Both the tube entering the cassette and the retentate tubing was placed here. The last tube exiting the cassette was the waste flowing through the membrane. This was collected in the red bottle.*

Materials:

- Vivaflow 200 with a 10 kDa MWCO PES membrane (Sartorius Stedim Biotech GmbG, Germany)
- dH_2O
- NaOH 1M
- 20% EtOH
- 20 mM Na-acetate buffer (pH at 5.0)
- 20 mM Na-citrate buffer (pH at 3.0)
- 20 mM Na-citrate buffer (pH at 4.0)

Apperatus:

- Tangential flow cassette with a 10kDa cut-off membrane
- pH meter (Sentron SI series Netherlands)

Procedure:

The filter was thoroughly washed with 20% ethanol, 1M sodium hydroxide, and distilled water in this respective order, to remove residual salts, proteins and other contaminates before use.

When the supernatant was filtered down to about 80 ml, from a starting total volume of 650 ml, 300ml buffer was added (see list below). This was repeated 3 times with the result that the original broth was exchanged with buffer. This removes all the salt, which was an important step before the subsequent ion exchange chromatography.

The endoglucanases were exchanged with different buffers to give the proteins a charge before ion exchange chromatography. Below is a list over buffers used for purification in IEC (ion exchange chromatography) for the respective enzymes.

- *Ta*Cel5A 20 mM Na-acetate buffer (pH at 5.0)
- *Af*Cel12A: 20mM Na-citrate buffer (pH at 3.0)
- *Ma*Cel45A: 20mM Na-citrate buffer (pH at 4.0)

After concentration, 20 ml buffer was run through the system to extract remaining proteins; they were mixed with the concentrated broth to a total volume of 100 ml. 100ml 1M NaOH washed the filter, followed by pH neutralization with dH_2O. The system was stored in 20%EtOH.

3.5.2. Ion exchange chromatography (IEC)

Ion exchange chromatography (IEC) was used to separate proteins based on their charge. Just as any form of chromatography, IEC requires a mobile phase and a stationary phase. The stationary phase is negatively charged in cation exchange and positively in anion exchange. Proteins do not have any charge at their pI; it is necessary to either increase the pH (proteins will get negatively charged and bind to the positive stationary phase in anion exchange chromatography) or decrease the pH (proteins become positively charged and bind to the negatively stationary phase in cation exchange chromatography). As a rule, the pH of the mobile phase buffer must be between the pKa (acid dissociation constant) of the protein and the pI (isoeletric point).

The charged molecules bind reversibly to the solid phase and elution is achieved by introducing a salt concentration, the molecules with the weakest ionic interactions to the solid phase will elute first. Molecules that have a stronger ionic interaction require a higher salt concentration and elute later as the gradient increases. As different proteins have various degrees of interaction to the ion exchanger (due to difference in charges and their distribution on the protein), separation is achieved as the ionic strength of the elution buffer is increased.

Materials:

- 20 mM Na-acetate buffer (pH at 5.0)
- 20mM Na-citrate buffer (pH at 3.0)
- 20mM Na-citrate buffer (pH at 4.0)
- 20 mM Na-acetate buffer 1M NaCl (pH at 5.0)
- 20mM Na-citrate buffer 1M NaCl (pH at 3.0)

- 20mM Na-citrate buffer 1M NaCl (pH at 4.0)
- 20% EtOH
- dH_2O

Apparatus:

- AKTA-prime plus purification system (GE Healthcare Bio-Science AB, Sweden)
- 5-ml DEAE sepharose fast flow purification column (GE healthcare Bio-Science AB, Sweden)
- 5-ml SP sepharose Fast flow purification column (GE Healthcare Bio-science AB, Sweden)
- pH meter (Sentron SI series, Netherlands)

Procedure:

In this procedure it was important that the pH of the buffers containing the respective enzymes was adjusted so the enzymes receive a charge. pH is adjusted with either NaOH or HCl, TaCel5A was solved in 20 mM Na-acetate buffer pH 5, this is higher than pI so anion exchange was necessary for separation., MaCel45A was solved in 20mM Na-citrate buffer with pH4, this is lower than pI, therefore cation exchange were used for separation. Afcel12A was solved in 20mM Na-citrate buffer pH3, lower than pI, so cation exchange was used for separating molecules. This was primarily achieved in ultrafiltration, but it needs to be verified by a pH meter.

Columns were washed with 20% EtOH and dH_2O before use. Buffer 2 (Table 3) for the respective enzymes was run through the system to elute any remaining enzymes from previous runs. This was followed by buffer 1, which equilibrates the column before enzyme loading.

Enzymes are loaded with a flowrate of 2ml/min (an overview of the various columns, buffers and elution procedures is provided in Table 3). When the enzyme solution has entered the column, it is washed with 5 column volumes (CV)(25ml) of equlibriumbuffer (1), to remove unbound proteins.

Proteins were eluted with a linear or stepwise gradient, from buffer 1 to buffer 2 where the NaCl concentrations increase over several column volumes (Table 3). 8ml fractions were collected in the beginning, and this was changed to 4 ml just before

proteins were expected to elute. Protein elution was observed on prime View software in real-time using UV absorbance at 280nm for protein detection. Relevant fractions were checked using SDS-PAGE (Section 3.6).

Columns were washed with 100% buffer 2 to remove all proteins followed by dH_2O and 20% EtOH. The DEAE column was stored in 20% EtOH. The SP column was stored in 20% etOH with 0.2M sodium acetate.

Table 3. Setup for ion-exchange chromatography

Enzyme	Material and chromatographic setup
MaCel45A	5-ml-SP sepharose fast flow column; cation exchange Buffer 1) 20mM Na-Citrate buffer (pH4) Buffer 2) Buffer 1 with 1M NaCl Linear gradient elution: over 20 CV, from 0 to 0.5M NaCl, 2ml/min
TaCel5A	5-ml DEAE sepharose fast flow column; anion exchange Buffer 1) 20mM Na-Acetate buffer (pH 5) Buffer 2) Buffer 1 with 1M NaCl Linear gradient elution: over 6 CV, from 0 to 0.15M NaCl, 2ml/min.
AfCel12A	5-ml-SP sepharose fast flow column; cation exchange Buffer 1) 20mM Na-Citrate buffer (pH3) Buffer 2) Buffer 1 with 1M NaCl Stepwise gradient elution: over 20 CV from 0- 0.13M NaCl, 2ml/min.

3.5.3. Vivaspin

After proteins were eluted from the IEC columns, they were very diluted and concentration was necessary before further purification. To achieve high concentrations, the collected fractions were concentrated with Vivaspin ultrafiltration spin columns. The membrane in Vivaspin tubes retains protein while the buffer goes through as waste.

Apparatus:

- Heraeus Multifuge X1R centrifuge (Thermo Scientific, Germany)
- 10 or 5 kDa MWCO Vivaspin20 ultrafiltration spin columns (Sartorius Stedim Biotech GmbH, Germany)
- pH meter (Sentron SI series Netherlands)

Procedure:

Vivaspin 20 concentration tubes were washed twice with 10ml dH_2O on 3500 rpm. After the membrane was washed, up to 15 ml enzyme solution was loaded on to the column. The enzyme solution was centrifuged at 3500 rpm for 10 minutes in each interval, concentrating all macromolecules in the solution. Between the intervals, protein solution was refilled and the filtrate was thrown away. 10 kDa Tubes were mainly used, but in some cases MaCel45A was observed in the permeate tube, in that case, tubes with a 5 kDa cut-off membrane was used.

3.5.4. SEC (size exclusion chromatography)

Size exclusion chromatography separates molecules based on their hydrodynamic radius. This property is based on both the size and the shape of the molecule. Opposed to most chromatographic methods, the molecules do not bind to a stationary phase. Instead, the molecules are separated by the speed at which they navigate through the stationary phase. The big molecules are the first to exit the column while small molecules are retained in pockets in the stationary phase (Figure 9). High protein concentrations and low loading volumes are used on SEC. This results in better separation. The smaller the volume, the less diffuse the eluted fractions will be (Ritchie 2012).

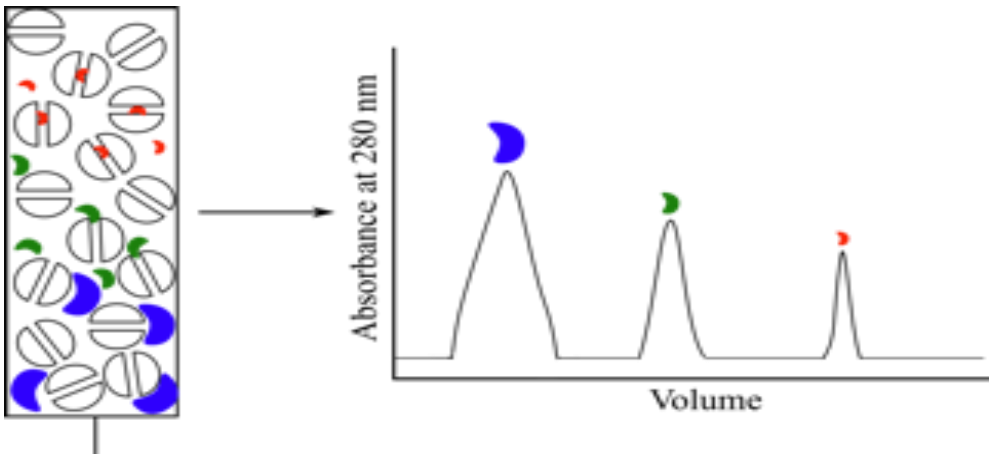


Figure 9. Size exclution chromatography. *A mixture of three proteins with different hydrodynamic radius was separated in a Size exclusion column. Large proteins elute first, and the smallest proteins elute later because they enter pores in the matrix and are retained (Ritchie 2012).*

The Column used was a Hiload 16/60 superdex 75 prepgrade. It consists of dextran covalently bound to highly cross-linked agarose resulting in a stationary phase with pockets, which retains the smaller proteins. The degree of crosslinking determines the pore sizes. Superdex 75 separates proteins in the molecular weight range between 3000 to 70000g/mol.

Materials:

- 20mM Na-Citrate pH 5.5 in 0.2M NaCl buffer
- 50mM Na-citrate + 0.2M NaCl pH 4
- 50mM Na-Citrate 0.2M NaCl pH 5
- 20%etOH
- dH_2O Filtered
- 5ml plastic syringe
- 30ml plastic syringe

Apparatus:

- AKTA P-900 pump-system (GE Healthcare Bio.science AB, Sweden)
- AKTA UPC-900 UV Lamp (GE Healthcare Bio.science AB, Sweden)
- HiLoad 16/60 Superdex 75 prep grad column (GE Healthcare Bio.science AB, Sweden)
- pH meter (Sentron SI series Netherlands)

Procedure:

Before installing the column it was important to check some parameters. The pressure needs to be kept below 0.4MPa and if there was any indication of air in the system, it was necessary to remove it to avoid getting it in to the column. This was critical because air would affect the separation.

When the column was installed, it was thoroughly washed with water to remove the ethanol which it was stored in. The column was equilibrated with buffer followed by loading of 1-2ml of protein (buffers listed in table 4). The flow rate was at 1 ml/min throughout the separation. 1 ml fractions were collected. Absorbance was measured at 280 nm and relevant fractions were checked using SDS-PAGE (se

section 3.6). Protein containing fractions were collected and concentrated on 10kDa vivaspin column (see section 3.5.3).

***Table 4**: Lists the buffers used in purifying MaCel45A and TaCel12A on SEC.*

Enzyme	Buffer
*Ma*Cel45A	50mM Na-citrate + 0.2M NaCl pH 4
*Ta*Cel12A	50mM Na-Citrate + 0.2M NaCl pH 5

3.5.5. Determination of protein concentration

Protein concentration was measured by recording absorbance at 280 nm with an Eppendorf Bio Photometer. Absorbance assays are convenient since they do not need any reagents and they do not consume any proteins in the process. However, the method has its limitations, because contaminants and other proteins in the solution may interfere and cause misleading results. Initially the protein concentration was measured with the more accurate Bradford protein assay, but the endoglucanase *Af*Cel12A did not react with the Coomassie Blue reagent. All proteins were therefore measured by absorption to obtain the same method for all proteins.

Trp, Tyr and Phe (aromatic residues) exhibit a strong UV absorbance, Cysteins forming disulfide bridges also affect the adsorption. If these cysteine residues are reduced, the coefficient decreases (Pierce 2002). Absorption is proportional to the content of these aromatic compounds and cysteine bonds. A molar absorption coefficient can be calculated based the quantity of these amino acids in the enzymes. For this study predetermined adsorbation coefficents available in the PDB database (http://www.rcsb.org/pdb/home/home.do) were used. Adsorbation coefficents for each enzyme is listed in Table 2.

Absorbance, A, is measured and can be converted to concentration using Beer-Lambert law:

$$A=\varepsilon lC$$

ε is the molar extinction coefficient (1/(M*cm), l is the path length (cm) of the cuvette and C is the protein concentration (M). in calculating the ε it is assumed that all cysteine residues form cysteine bonds.

Apparatus:

- Eppendorf Bio Photometer (VWR, Germany)

Materials:

- Eppendorf bio photometer (Eppendorf AG, Oslo)
- dH_2O
- cuvettes (UVette, Eppendorf AG, Germany)

Procedure:

For each sample, a 20 dilution was made in water. All samples were measured at 280nm in triplicates using the protein assay program on the spectrophotometer. Water was used as a zero-sample. Since this assay do not include any chemicals, the protein solution was stored at 4°C after analysis, they could be up-concentrated if necessary

3.6 SDS-PAGE (Sodium dodecyl sulfate-Poly Acrylamid Gel Electrophoresis)

SDS-PAGE separates proteins based on their chain lengths and mass to charge ratio. Reducing agents and the anionic detergent sodium dodecyl sulfate (SDS) denature the proteins by disrupting the secondary, tertiary and quaternary structure, thus creating a linear chain of amino acids. SDS Binds to the linear structure and imparts a negative charge which results in a fractionation based on size of the protein (Rath, Glibowicka et al. 2009). SDS-PAGE was used to evaluate the levels of protein expression (when growing *P.pastoris)* and the purity of the protein after each step of purification.

Material:

- NuPAGE LDS samplebuffer (4X) (Invitrogen, USA, Cat. NO. NP-0007)
- NuPAGE sample Reducing agent (10X) (Invitrogen, USA, Cat. NO. NP-0009)
- dH_2O
- Precast- Any kD TGX Stain-free gels (Bio-Rad Laboratories, USA, Cat. No 456-8123)
- BenchMark Protein Ladder (Invitrogen, USA, Cat. No, 10747012) containing 15 recombinant proteins between 10 to 220kDa.
- Tris/glycine/SDS running buffer (Bio-Rad Laboratories Inc. USA, Cat. NO. 161-0772) (25mM Tris, 192mM glycine, 0.1% (w/v) SDS, pH 8)

Apparatus:

- Powerpac 300 power supply electrophoresis (Bio-Rad Laboratories INC, USA)
- Gel Doc EZ imaging system (Bio-Rad Laboratories INC, USA, Cat No. 170-8270)

Procedure:

10µl sample was mixed with 10µl of a SDS loading buffer-stock solution containing 2.5µl sample buffer, 1µl reducing agent and 6.5µl dH_2O. This solution was boiled for 5 minutes to denature the proteins. The precasted gel was installed in a mini-

PROTEAN tetra electrophoresis system and the chamber was filled with SDS running buffer. In the first well, 3µl of Benchmark ladder was applied, followed by 12µl of each sample in the remaining wells of the precasted gel.

The electrophoresis was run at 300 V for 17 minutes. Following the electrophoresis, the samples were transferred to a Stain Free Sample Tray for imaging. The procedure consisted of a 5-minute gel activation procedure, during which the gel was radiated with UV-light, which modifies the tryptophan residues to fluorescent trihalo compounds (McDonald, Choe et al. 2008). Detection and quantification were evaluated with image lab software.

3.7 Conversion of waste paper to monosaccharides with commercial cellulase cocktail (Cellic CTec2)

In the following procedures, Cellic CTec2 was used to convert SOP, CCB and SNP to monosaccharides. First an acid loading-test was performed to find the optimal acid loading required to attain pH 5, as pH 5 is optimal for conversion with Cellic CTec2 (Novozymes 2015). Following the acid-loading test, 24-hour hydrolysis of all samples was performed with the optimal acid loadings for the respective substrates. Furthermore, optimization experiments were performed, trying to increase the yields close to 100% of the theoretical conversion. The experiments was as follows:

Materials:

- Cellic Ctec2 (Novozymes, Denmark) protein cons, 57mg/ml
- SOP, CCB and SNP with and without pretreatments
- 50mM Na-citrate buffer pH 5
- 1M citric acid

Apparatus:

- Incubator 50°C
- Waterbath 50°C
- Rotating mixer
- pH meter (Sentron SI series Netherlands)

Procedure:

General procedure for all assays:

Never-dried samples after steam explosion were used to avoid structural changes that might occur during drying. Using wet samples (with known dry matter content) was also very useful to avoid electrostatic issues during weighing. All reactions were run in triplicates, unless otherwise described. The enzyme mixture, Cellic CTec2, was pre-diluted in 50mM Na-citrate buffer pH 5 to decrease the risk of pipetting errors.

Reactions were run in volumes of 10 ml with 0.5g DM in 50ml Falcon tubes. The samples were initially suspended in 50mM Na-citrate buffer pH5 up to 7ml. According to the results from the pH control test, 1M citric acid was added to set the pH to 5, followed by additional buffer. Enzymes were added last to make up the total volume of 10 ml.

Samples were incubated in a water bath at 50^0C for 10 minutes, before being transferred to an oven, water transfers heat faster than air, this means that the samples reach 50°C faster so the enzymes start the hydrolysis. Individual tubes were incubated. After the incubation the individual tubes were taken out of the oven and the reaction was stopped by boiling (at 100°C) in a water bath for 5 min. Then the samples were cooled to room temperature and centrifuged at 3000 rpm for 10 minutes. 1 ml from the liquid fractions was taken for analyzing reducing sugars with the DNS method (as described in Section 3.10.1). The pH was measured before and after hydrolysis.

<u>pH control:</u>

pH control were performed for SOP, CCB and SNP with most severe pretreatment (210°C, 14min) and an enzyme loading of 5mg/g DM.

SOP was tested with acid loadings of 550, 600, 650, 700 and 750µL 1M citric acid, SNP and CCB were tested with acid loadings of 0, 50, 100, 150 and 200µL 1M citric acid.

<u>24-hour hydrolysis assay:</u>

24-hour hydrolysis assays were performed for SOP, CCB and SNP with varying pretreatments at optimal acid loadings, concluded from the previous tests. Enzyme concentrations of 5 mg/g DM were used.

<u>Maximizing conversion yields:</u>

In the attempt to get 100% theoretical yields, two assays were performed. In the first assay, the incubation time was prolonged to 72-hour hydrolysis whereas the enzyme concentration still was 5mg/g DM. The second assay was performed with

prolonged incubationtime-72-hours and increased enzyme concentrations (25mg/g DM).

3.8 Production of oligosaccharides

Time course experiments were performed to analyze the production of oligosaccharides over 24 hours. Each of the enzymes, Macel45A, TaCel5A and AfCel12A was incubated with each substrate, SOP, CCB, SNP with the most severe pretreatment (210°C, 14min) and Avicel, the amount of soluble cellodextrines was measured with HPLC (see section 3.10.2).

The experiment was performed over 24 hours with 6 timepoints (0.5, 1, 2, 3, 6 and 24 hours for Avicel and 5 timepoints (0.5, 1, 2, 6 and 24 hours) for SOP, CCB and SNP. The 3-hour time point was discarded for SOP, CCB and SNP because small variations were observed between 2 and 3 hours for Avicel.

Materials:

- 50mM Na-Citrate buffer and pH 5
- Purified enzyme preparations (Macel45A, TaCel5A and AfCel12A)

Apparatus:

- Eppendorf Thermo mixer, with 1.5ml thermoblock.
- Vortex mixer

Procedure:

The experiment was performed in 1.5 ml eppendorf tubes with a reaction volume of 1 ml with 1% substrate (w/v) and enzyme concentration of 5 mg/g DM. The reaction was conducted in 50mM Na-Citrate buffer pH5 for all enzymes, at 50 °C and 1000 rpm shaking in an Epperdorf thermo mixer.

The reactions were run in triplicates and at each time point a 20µl sample was taken from the tubes (samples were taken from the same tubes at each time point). At sampling time points, the shaking was paused and a sample was withdrawn. Substrate was avoided when pipetting because of inhomogeneity issues (substrate chunks of different sizes would be drawn into the pipette if the solution was mixed first); as a result only the supernatant was withdrawn. 20µl

sample was mixed with 40µl 0.5M NaOH to denature the enzymes and effectively stop the reaction. This led to a final concentration of 0.33M NaOH. The samples were diluted further down to 0.1M NaOH to be similar to the equilibrium buffer (0.1 M NaOH) used in the subsequent HPLC analysis (Section 3.10.2.2).

The supernatant samples were analyzed for their content of Glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose, which were identified and quantified using HPLC (see Section 3.10.2.2). In the following calculations, a hydrolysis factor was implemented for glucose and each cellodextrine to compensate for the addition of H_2O during enzyme hydrolysis. For example: when glucose is bound to two other glucose molecules, it has a molecular weight of 162g/mol. When the chain is hydrolyzed, a proton and a hydroxyl molecule get added to the C-4 and C-1 carbon respectively, increasing the molecular weight by 18. So, the total weight of a cellotriose unit is 504 g/mol, the hydrolysis factor is calculated by the following calculation:

$\frac{504g/mol-18g/mol}{504g/mol} = 0.964$. This value was multiplied with the triose results after HPLC analysis. This calculation was performed for glucose and all cellodextrines. The total theoretical yield (100% of the theoretical maximum) was calculated and plotted against the time points. There was no pH control in this experiment and this may hamper the yields, especially for SOP, which contains calcium carbonate (see Section 1.6 for more information about calcium carbonate).

3.9 Effect of endoglucanases on the degree of polymerization

As endoglucanases introduce cuts on the cellulose chain, they may decrease the degree of polymerization. The low solubility of products longer than six glucose units prevents them from being solubilized and observed on HPLC analysis- Therefore, another approach was chosen to observe endoglucanases effect on insoluble material, experiments were done with MaCel45A acting on SOP, CCB, SNP and Avicel.

In the first test *Ma*Cel45A was used to hydrolyze SOP, CCB and SNP for 30minutes and 24 hours. In a second test, *Ma*Cel45A was used to hydrolyze Avicel with different enzyme concentrations for 24 hours. Both soluble and insoluble material were analyzed

Materials:

- 50mM Na-Citrate buffer and pH 5
- Purified enzyme preparations *Ma*Cel45A
- SOP, CCB, SNP pretreated at 210°C for 14 min and Avicel

Apparatus:

- Eppendorf Thermo mixer, with 1.5ml thermoblock.
- Vortex mixer

Procedure:

General procedure:

The experiments were performed in 1.5 ml eppendorf tubes with a reaction volume of 1 ml with 1% substrate (w/v). The reactions were conducted in 50mM Na-Citrate buffer pH5, at 50°C and 1000 rpm shaking in an Epperdorf thermo mixer. All experiments were run in triplicate. The enzyme reaction was terminated by the addition of 1.1% SDS and subsequent boiling in a water bath for 5 minutes. The SDS was added in order to remove the proteins (see section 3.11.2) before measuring the degree of polymerisation (Section 3.11) Two experiments were performed:

Effect of MaCel45A on the DP of SOP, CCB and SNP

*Ma*Cel45A hydrolyzed SOP, CCB and SNP for 30 minutes and 24 hours respectively. The enzyme concentration was 0,5 mg/g DM.

24 hour conversion of Avicel with three different concentrations of MaCel45A

Avicel was hydrolyzed with enzyme concentrations of 0,05, 0,1 and 0,5mg/g DM.

3.10 Analysis of carbohydrates

3.10.1. DNS method

The DiNitroSalicylic acid (DNS) can be used to estimates the concentration of reducing carbohydrates in a solution. Free carbonyl groups (C=O) equimolarly react with DNS molecules. The carbonyl groups reduce DNS to 3-amino-5-nitrosalicylic acids under alkaline conditions. This reduced form of DNS changes the amount of light absorbed at wavelength 540 nm. The measured change in absorbance is proportional to the amount of reducing sugars.

Materials:

- DNS reagent batch (0,04M 3,5-dinitro salicyclic acid in 1.0M K-Na-tartrate and 0.4M NaOH). New batch was produced each month.
- Standard: glucose

Apparatus:

- Water bath 100°C
- UV-visible spectrophotometer Hitachi U-1900

Procedure:

The procedure was performed in 1,5 ml eppendorf tubes. Standards of glucose were prepared with the concentration between 0.1 to 0.5 mg/ml, diluted in dH_2O.

500 µl of the samples were diluted with dH_2O to obtain values within the standard curve, preferably values corresponding to around 0,3mg/ml glucose. The samples were mixed with 750µl DNS reagent, followed by boiling in a water bath for 5 minutes; samples were cooled down to room temperature before measuring absorbance at 540nm. Blanks contained 500µl water and 750µl DNS reagent.

3.10.2. HPLC

3.10.2.1. Monosaccharides

Materials:

- Standards ranging from 0.0025g/L to 0.04g/L
- Filtered and degassed dH_2O
- Standards mixture: glucose, xylose, mannose, arabinose, galactose.

Apparatus:

- Dionex ICS3000 system (Thermo Fisher Scientific, USA)
- Carbopac PA1 column (2 x 250nm) (thermo Fisher Scientific, USA)
- Carbonpac PA1 guard column (2 x 50nm) (thermo Fisher Scientific, USA)
- Eluent generator with KOH
- Chromeleon 7.0 (thermo Fisher Scientific, USA)

Procedure:

HPAEC-PAD (High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detector) analysis for monosaccharides was performed using a Dionex ICS3000 system with an eluent generator.

Carbohydrates were separated by isocratic elution with 1mM KOH as an eluent at 0.25ml/min flow rate and 30°C column temperature. Standard (containing 1g/l of all following carbohydrates: glucose, xylose, mannose, arabinose, galactose) were prepared in dH_2O in the concentration range of 0.0025 to 0.04g/L. Samples analyzed were diluted 15 and a 100-folds to be within this concentration range. Hemicelluloses were quantified within the 15-fold dilution, and glucose was within the standard curve at 100-fold dilution.

The chromatograms were analyzed with Chromeleon 7.0 (thermo Fisher Scientific, USA). Glucose, xylose, mannose, arabinose, galactose were separated and quantified by peak area and compared to known standards.

3.10.2.2. Oligosaccharides

Materials:

- 0.1M NaOH
- 1M Na-Acetate in 0.1M NaOH
- Standard mixture containing: Glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose

Apparatus:

- Dionex ICS5000 system (thermo Fisher Scientific, USA)
- Carbopac PA1 column (2 x 250nm) (thermo Fisher Scientific, USA)
- Carbonpac PA1 guard column (2 x 50nm) (thermo Fisher Scientific, USA)
- Chromeleon 7.0 (thermo Fisher Scientific, USA)

Procedure

HPAEC-PAD (High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detector) analysis for oligosaccharides was performed using a Dionex ICS5000 system, equipped with a CarboPac PA1 column and a CarbonPac PA1 guard column. The system was operated with a flow rate of 0.25ml/min, with column a temperature at 30°C. The mobile phase (A) contained 0.1 M NaOH, and the eluent contained 1M Na-Acetate with 0.1M NaOH (B). Standards (containing: glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose) were prepared in dH_2O in the concentration range of 0.025 to 0.3g/L of each carbohydrate unit. Samples analyzed were diluted to be within this concentration range.

Elution of samples was performed using a 29min gradient mixing buffer A and B. The program starts with a 100% A, and reaching 10% B after 10 minutes, 14%B after 15 minutes, 30%B after 16 minutes and 100%B after 18 minutes. After 0.1 minutes the system was equilibrated with eluent A for 9 minutes before next injection. The chromatograms were analyzed with Chromeleon 7.0 (thermo Fisher Scientific, USA). Glucose and cellodextrines were separated and quantified by peak area and compared to know standards listen in materials.

3.11 Determination of the number-average degree of polymerization (DP)

Yong P. Zhang and Lee R. Lynd developed a method for determining the number-average degree of polymerization for insoluble cellulose and soluble cellodextrins (Zhang and Lynd 2005). The assay is based on measuring the total amount of glycosyl monomers and dividing it by the reducing end concentration. The reducing end concentration is determined with the 2,2´ bichinchoninate (BCA) method. The total amount of glucosyl monomers is determined with the phenol-sulfuric acid method, as described below.

3.11.1. Homogenization of SOP, CCB, SNP and Avicel.

SOP, CCB and SNP are inhomogeneous substrates. This did not cause a problem in the enzyme hydrolysis with Cellic CTec2 because the enzymes hydrolyzed the substrates completely and relative large sample sizes made them easy to handle in following steps of the protocol. The BCA and the phenol–sulfuric methods (section 3.11.3 and 3.11.4) require small amounts of sample, and those samples have to be divided into equal small portions. Without homogenization, high standard deviations were observed. To solve this issue an IKA T18 Basic homogenizer was used to physically homogenize SOP, CCB SNP and Avicel.

Equipment:

- IKA T 18 Basic homogenizer
- 205mm dispersing element
- 108mm dispersing element

Procedure:

Homogenizing with T18 basic homogenizer was performed on SOP, CCB, SNP and Avicel with various durations to observe how the mixer influenced the insoluble material (degree of polymerization). 5 ml reaction volumes with 1% DM were mixed

at the duration were homogenized for, 0.5, 1 or 5 minutes at 10000 1/min(rpm). Subsequently, the degree of polymerization was determined using the Phenol-sulfuric acid and BCA method, as described in section 3.11.3 and 3.11.4.

Two different dispersing elements were used in this thesis. In Section 4.5.1, a 205mm was used, in Section 4.5.2 we had to change to a 108mm dispersing element because the 205mm dispersing element got dysfunctional between the two experiments.

3.11.2. Removing proteins

If the samples have been through enzyme hydrolysis, it is essential to remove the proteins before BCA analysis (section 3.11.4). BCA reagent is not specific for sugars, other compounds, such as proteins, are also able to reduce the copper ions in the BCA reagent. The reaction produces a blue-violet color absorbing light at 562nm, further explained in section 3.11.4).

Sodium Dodecyl Sulphate (SDS), is a detergent capable of unfolding proteins and disrupt any non-covalent interactions the proteins may have to the substrates (SOP, CCB, SNP or Avicel)(Bilek, Bax et al. 2011). 75% EtOH was able to wash away the unfolded proteins, separating them from the insoluble cellulose. It may be noted that only small amounts of natural fibrous proteins remains in lignocellulosic biomass (SOP, CCB and SNP) after pulping processes during papermaking. The diminishing amount of protein will have no considerable effect on the BCA assay. If, however, the samples were treated with endoglucanases during experiments, they will have to be removed before subsequent BCA assay.

Materials:

- SDS 1.1%
- 75%EtOH
- dH_2O

Apparatus:

- Heraeus Multifuge X1R centrifuge (Thermo Scientific, Germany)

Procedure:

After enzyme hydrolysis the samples were centrifuged at 4500g for 5 minutes. The supernatant was removed and collected for analysis by HPLC for cellodextrines (section 3.10.2.2). The removed supernatant was replaced with an equal amounts of 1.1% SDS. Following 5 minutes of boiling, the samples were again centrifuged, and the SDS was removed and replaced with 75%EtOH. The pellet was washed three times with EtOH, then once with water. Subsequent to the washing steps, water re-suspended the pellet up to the original amount. This was possible by weighing the tubes before this procedure commenced.

Diluted working solutions were prepared after proteins were removed. The working solution was necessary to obtain a representative (equal) sample, and this sample will be used in both the phenol-sulfuric acid- and the BCA method. 1 ml working dilutions (1) were obtained by diluting the original sample, so that it contained approximately 1 g/L DM.

Working dilution (1): contains diluted biomass fractions after proteins was removed (Section 3.11.2)

Working dilution (2): for the BCA assay, 50:1 (v/v) BCA reagent A and B (pierce BCA protein assay kit by life technologies/thermo scientific)

3.11.3. The Phenol-sulfuric acid method for determination of total glycosyl monomer concentration

The phenol-sulfuric acid method measures the total glycosyl monomer concentration of insoluble cellulose. When cellulose reacts with concentrated sulfuric acid, they form furfural derivatives. Further reaction between the furfural and phenol develops a detectible color measured by a spectrophotometer (DuBois, Gilles et al. 1956).

Materials:

- 80% (w/w) phenol
- Concentrated H_2SO_4

- Glucose standard solution 0.0 to 0.1 g/l (0-556µM)

Apparatus:

- Water bath 30˚C
- UV-visible spectrophotometer Hitachi U-1900 and high purity quartz cuvettes.

Procedure:

Liquefaction step (for insoluble cellulosic samples)

100µl cellulosic samples from the working solution (1) (see procedure in section 3.11.2) was mixed with 100µl concentrated H_2SO_4 in triplicates. The solution was incubated for 7 minutes so the acid can hydrolyse the cellulose to glucose monomers. Following the hydrolysis, 800µl H_2O was added to dilute the samples 10-fold

Phenol-sulfuric acid assay

100µl sugar solution from the liquefaction step was diluted in 100µl H_2O to be within the ideal concentration range corresponding to around (0,00 to 0,10g/L glucose on the standardcurve. Then 10µl 80% (w/w) phenol was added to all samples, standards and blank (with water), followed by 500µl concentrated H_2SO_4. The solutions were mixed thoroughly, and incubated for 10 minutes in room temperature followed by further 20 minutes incubation in a water bath at 30˚C. Absorbance was measured at 490nm.

3.11.4. The 2,2´´-bicinchoninate (BCA) method for determination of the reducing-end concentration on insoluble cellulose.

The BCA assay measures the reducing-end concentration in insoluble cellulose.

When aldehyde groups at the reducing ends of cellulose chains react with BCA at elevated temperatures, Cu^{2+} is reduced to Cu^{1+}. The assay chemistry involves two molecules of BCA chelating each Cu^{1+} ion, forming a complex that absorbs light at 562nm (Redinbaugh and Turley 1986, Decker, Brunecky et al. 2009). The measured absorbance is proportional to the amount of reducing sugars.

The 2,2´ bichinchoninate (BCA) method and DNS both measure the amount of reducing-end sugar in a solution. Yong P. Zhang and Lee R. Lynd claimed that both these methods would give inaccuracy because they suffer from some hydrolysis during the assay (decomposition of carbohydrates). Furthermore, some chain ends could be inaccessible for the reactants as they could be "protected" by other cellulose chains. Between these methods, BCA was preferred as it was more sensitive, and because DNS seems to hydrolyze polysaccharides more extensively during the assay.

Materials:

- BCA reagent A and B (Pierce BCA protein assay kit by Life Technologies/Thermo Scientific)
- Glucose standards containing reducing ends corresponding to 0.0-0.01 g/l glucose equivalents

Apparatus:

- UV-visible spectrophotometer Hitachi U-1900 and high purity quartz cuvettes.
- Eppendorf Thermo mixer, with 1.5ml thermo block
- Eppendorf centrifuge

Procedure:

A BCA working solution (2) was prepared fresh daily by mixing 50 parts of BCA reagent A with 1 part of BCA reagent B (A/B as 50:1 v/v). The working solution contains the reagent that reacts with aldehyde-groups on the reducing ends of carbohydrates, resulting in a color change measured at 562 nm.

Insoluble biomass from the working solution (1) (section 3.11.2) was diluted until it contained reducing-ends corresponding to 0,0-0,01g/L glucose equivalents. The amount of reducing ends in different substrates varies, so initial pre-tests had to be done in order to find the correct dilution of the samples.

Following the initial tests, optimal dilutions were made. 400µl cellulose suspensions were mixed with 400µl BCA working solution. The samples were incubation for 30 minutes at 75˚C and the tubes were cooled down to room temperature before subsequent steps. Samples were centrifugation for 1 minute at 10 000g and absorbance was measured at 560nm

4 Results:

4.1 Compositional analysis

Un-pretreated and pretreated SOP, CCB and SNP were used in this thesis and their composition were analyzed according to a method developed by A.Sluiter and colleagues at the U.S Department of Energy (Sluiter, Hames et al. 2008). This method was used to find the glucose, xylose, mannose, arabinose, galactose, ash and lignin content of the biomass (Table 5). Ash contained inorganic material left after incineration amounted for an average of 21% of total DM in SOP, 11% in CCB and 7% in SNP, calculated from the ash content of untreated and pretreated samples. "Others" in Table 5 represented the gap in mass balance, including unidentified compound. "Others" value amounted to an average of 29% in SOP, 0.4% for CCB and 26% in SNP for untreated and pretreated samples. Figures 10 illustrate the compositions of the respective materials and also show how these compositions vary with increasing intensity (temperature and duration) of the pretreatment. Figure 10 also show how different the composition in SOP, CCB and SNP was; this was mainly because of the various pretreatments paper receives under production (See section 1.6 for more information).

In the following text the composition of the three types of substrates (SOP, CCB and SNP) is compared based on the average value for each component calculated from the untreated and pretreated samples of each type of substrate, unless stated otherwise.

SOP contained an average of 47% carbohydrates and 2% lignin (table 5). Glycans comprised an average of 38%, while the most abundant hemicellulose was xylose (7.3%) followed by mannose (0.9%). Only small amounts of arabinose (0.07%) were detected and galactose was under the detection limit.

Table 5: Compositional data for SOP, CCB and SNP*. Ash contained inorganic compounds. B.D.L (below detection limit) means there was no compound identified.*

Material	Pretreatment	Glucose (g/kg DM)	Xylose (g/kg DM)	Mannose (g/kg DM)	Arabinose (g/kg DM)	Galactose (g/kg DM)	Lignin (g/kg DM)	Ash (g/kg DM)	Other (g/kg DM)
SOP	untreated	486.3 ± 59.7	85.2 ± 35.2	12.6 ± 4.1	0.8 ± 0.4	B.D.L	20.8 ± 11.9	206.0 ± 0.7	187.9
	170˚C, 7min	385.3 ± 74.2	67.8 ± 19.9	8.9 ± 3.2	0.6 ± 0.0	B.D.L	20.3 ± 5.3	213.0 ± 1.5	303.8
	170˚C. 14min	336.4 ± 20.0	71.0 ± 21.6	2.1 ± 3.7	0.8 ± 0.3	B.D.L	24.1 ± 10.5	207.5 ± 1.0	357.9
	190˚C. 7min	356.1 ± 109.1	82.5 ± 24.9	10.6 ± 3.1	0.8 ± 0.4	B.D.L	21.1 ± 2.1	213.2 ± 2.4	315.4
	190˚C. 14min	348.3 ± 66.7	63.1 ± 10.9	9.0 ± 1.7	0.6 ± 0.0	B.D.L	24.4 ± 6.0	214.2 ± 1.2	340.1
	210˚C. 7min	366.9 ± 63.8	76.8 ± 3.0	9.7 ± 1.2	0.6 ± 0.0	B.D.L	20.0 ± 2.7	206.5 ± 3.2	319.1
	210˚C. 14min	415.7 ± 69.2	73.6 ± 16.4	10.3 ± 2.3	0.6 ± 0.0	B.D.L	26.3 ± 6.7	212.7 ± 6.5	260.5
CCB	untreated	747.9 ± 70.4	92.9 ± 7.7	46.8 ± 2.5	7.7 ± 0.4	6.9 ± 0.4	125.8 ± 8.6	114.2 ± 4.0	-142.1
	170˚C. 7min	672.8 ± 40.3	94.0 ± 4.5	48.3 ± 3.8	7.6 ± 0.6	7.8 ± 0.8	132.2 ± 7.3	119.1 ± 4.0	-81.8
	170˚C. 14min	494.7 ± 132.2	85.4 ± 3.4	43.6 ± 2.8	6.5 ± 0.4	7.0 ± 0.7	119.5 ± 4.1	118.7 ± 2.0	124.5
	190˚C. 7min	587.9 ± 7.1	100.9 ± 9.1	50.8 ± 7.3	7.3 ± 1.1	7.9 ± 1.6	125.3 ± 4.1	104.7 ± 1.0	15.1
	190˚C. 14min	574.6 ± 21.3	91.4 ± 3.0	45.0 ± 2.2	5.2 ± 0.4	6.8 ± 0.6	127.0 ± 7.4	107.5 ± 4.0	42.5
	210˚C. 7min	587.7 ± 30.9	72.7 ± 4.1	42.1 ± 2.1	1.6 ± 0.2	4.4 ± 0.0	137.5 ± 2.7	115 ± 12.0	39.0
	210˚C. 14min	587.4 ± 5.5	59.0 ± 7.1	40.4 ± 4.3	0.0 ± 0.0	1.9 ± 0.3	163.7 ± 5.4	112.5 ± 5.0	35.2
SNP	untreated	279.9 ± 60.5	26.1 ± 5.2	52.7 ± 11.2	6.3 ± 1.2	9.8 ± 2.4	250.5 ± 5.2	71.5 ± 4.1	303.0
	170˚C 7min	236.3 ± 22.8	22.3 ±1.8	40.5 ± 2.2	4.7 ± 0.3	7.5 ± 0.5	256.8 ± 0.0	82.2 ± 4.1	349.7
	210˚C 7min	373.6 ± 75.2	24.1 ± 7.4	56.5 ± 18.7	0.9 ± 0.5	7.2 ± 2.6	280.7 ± 2.8	80.2 ± 1.6	176.7
	210˚C 14min	316.7 ± 23.4	9.3 ± 1.3	23.7 ± 1.1	0.2 ± 0.3	2.0 ± 0.3	339.9 ± 1.2	66.4 ± 2.0	241.8

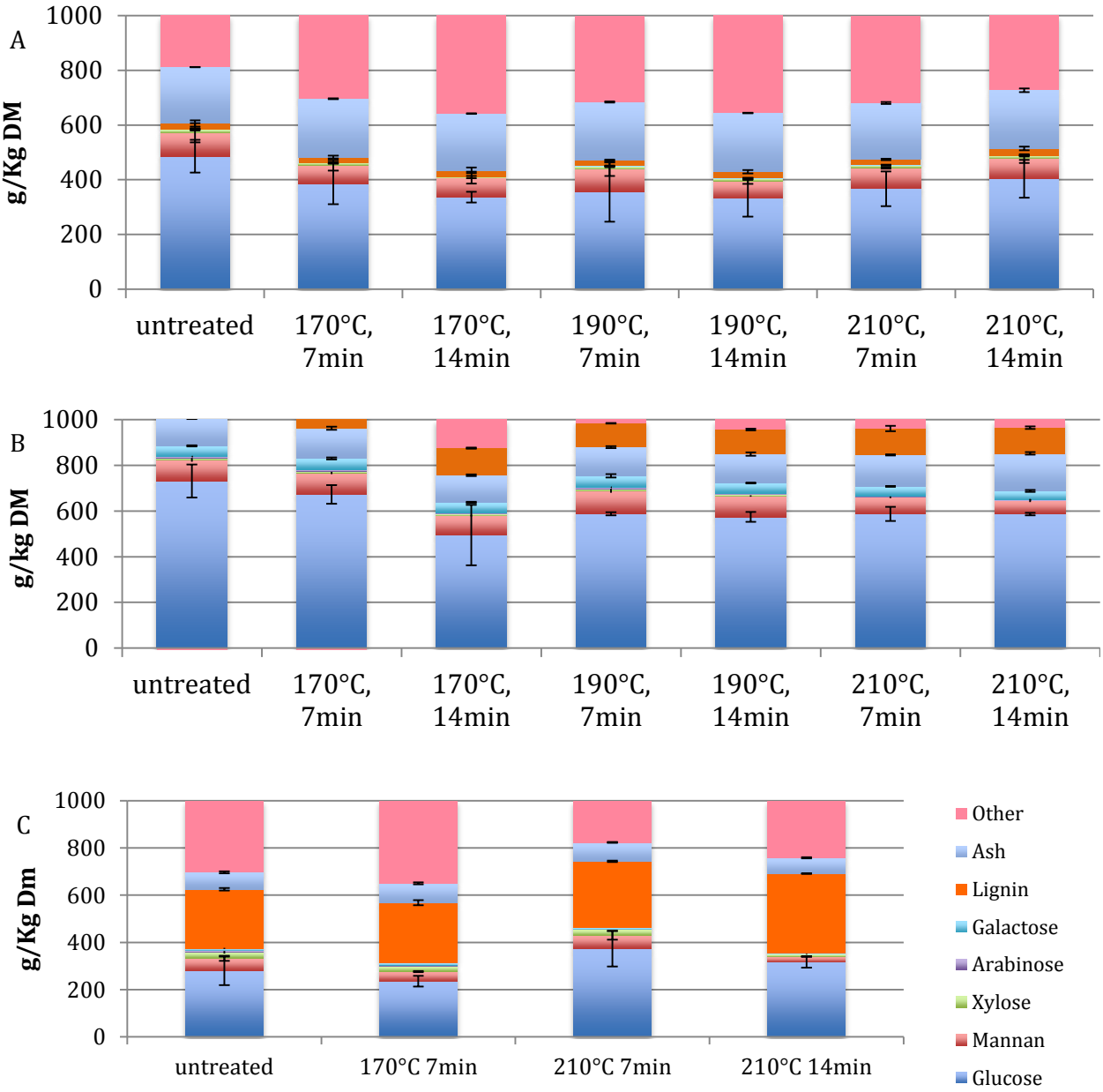


Figure 10. Compositional analysis of SOP (A) CCB (B) and SNP (C) after various pretreatments. *"Other" refers to unidentified compounds and/or mass loss, whereas the ash was the inorganic material. See section 3.2 for a detailed description of the methods used.*

CCB contained the highest values of carbohydrates of the substrates with average values of about 70% carbohydrates and 12.5% lignin. Glycans comprise an average of 60%. The most abundant hemicellulose is xylose (8.4%) followed by mannose (4.4%), small amount of arabinose (0.6%) and galactose (0.6%) was detected. CCB clearly contained more hemicellulose (15%) and lignin (12,5%) compared to SOP (9.8% and 2% respectively); CCB also contained traces of galactose. The analysis of CCB showed a good mass balance with generally low

values of “others” (Figure 10B), except from the three first samples (Table 5). The samples pretreated for 170^0C in 14min gave high standard deviations for glucose and this could account for the high values of unidentified (i.e “others”) compounds (meaning: if new compositional assay was run for 170^0C in 14min pretreated sample, higher glucose values could be observed, thereby lowering the “other/ loss value”). For untreated and 170^0C in 7min samples, there seemed to be an overestimation of the glucose content yielding negative values of “others”.

SNP contained an average of 36% total carbohydrates and 28% lignin (table 5). Glycans comprised an average of 30%. The most abundant hemicellulose was mannose (4.2%) followed by xylose (2.0%). Only small amount of galactose (0.06%) and arabinose (0.03%) and was detected.

SNP had the highest values of lignin, which comprises an average of 28%. SNP contained less xylose (2.0%) than CCB (8.4%) and SOP (7.3%) but more mannose (5.2%) than CCB (4.6%) and SOP (0.9%). SNP contained only 37,4% carbohydrates, which is the lowest of the three substrates.

Observations in Table 5 show a slight decrease in hemicelluloses for all samples during increasing intensity of pretreatments. SOP illustrated a good example; in untreated SOP, hemicelluloses make up 9.8% of the total weight. The composition of the SOP with the most severe pretreatment (210°C 14min) showed that it had decreased to 8.4%. Furthermore, the high values of ash and others in SOP were partly explained by the content of calcium carbonate, further discussed in section 5.

4.2 Enzymatic conversion to monosaccharides

The primary goal of these enzymatic conversion experiments was to convert waste paper (SOP, CCB and SNP) to monosaccharides with a commercial cellulase cocktail, and to maximize the conversion yield. From these results, we get an idea about the accessibility of the substrates. Enzyme optimization experiments were performed to evaluate which variables in enzymatic hydrolysis could result in optimized yield. Increase enzyme loadings, incubation and pH control was evaluated.

4.2.1. pH control

Enzymes are very dependent on the pH to function properly in solution. Enzymes have different pH optima, so it was important to adjust the pH with citric acid to achieve optimal conditions for the enzyme. Initial tests without pH control, resulted in low yields, especially for untreated and pretreated SOP. Depending on the alkalinity of the substrate, various amounts of acid had to be added to reach the optimal conditions (around pH 5 for Cellic CTec2). Therefore, acid loading tests were performed on the most severe pretreated substrates, in order to get an indication on the amounts of acid required for each biomass (SOP, CCB and SNP)(Table 6). The “correct” acid loading required on the most severely pretreated substrates, were used on all pretreatments in subsequent experiments.

Theoretical maximum yields were calculated as the ratio of solubilized sugars and total carbohydrate content of each substrate (Table 5). The amount of solubilized sugars was determined as the amount of reducing ends in the hydrolysate, measured with DNS reagent. Total carbohydrate content of the substrates was calculated as the sum of glucose, xylose, mannose, arabinose and galactose, as determined by the composition analysis (Section 4.1)

Table 6: Acid loading experiment. *Substrates were tested with different acid loadings to find the optimal acid loading for Cellic CTec2. Substrates pretreated at 210°C for 14min were used for all samples. The table summarizes the yields, initial pH (before the reaction starts) and final pH (after the reaction was done). The enzyme concentration of Cellic-Ctec2 was 5 mg/g dry mass. Reaction volume was 10 ml with 5% (w/v) dry mass and the samples were incubated in a oven at 50°C for 24 hours with 20 mM Na-citrate pH 5 used as a buffer. pH was measured before and after the reaction. DNS assay was used to quantify end products.*

Material	Conversion % as of Theoretical maximum	Initial pH	Final pH	1M Citric acid loading (µl)
SOP	42.9 ± 2.9	4.8 ± 0.1	6.5 ± 0.2	550
	53.6 ± 1.3	4.5 ± 0.1	5.5 ± 0.1	600
	52.0 ± 5.3	4.5 ± 0.1	5.0 ± 0.1	650
	47.1 ± 2.9	4.4 ± 0.1	4.7 ± 0.1	700
	43.1 ± 0.7	4.3 ± 0.1	4.4 ± 0.1	750
CCB	28.4 ± 1.2	5.2 ± 0.0	5.9 ± 0.1	0
	31.4 ± 0.0	5.0 ± 0.0	5.3 ± 0.1	50
	30.4 ± 0.2	4.7 ± 0.1	4.9 ± 0.0	100
	29.1 ± 0.1	4.5 ± 0.1	4.6 ± 0.1	150
	27.5 ± 0.2	4.4 ± 0.0	4.4 ± 0.1	200
SNP	29.4 ± 0.3	5.2 ± 0.0	5.4 ± 0.1	0
	52.6 ± 0.7	4.6 ± 0.0	4.5 ± 0.0	50
	48.9 ± 0.6	4.4 ± 0.0	4.3 ± 0.1	100
	45.1 ± 0.8	4.3 ± 0.0	4.2 ± 0.1	150
	42.8 ± 0.4	4.1 ± 0.0	4.0 ± 0.0	200

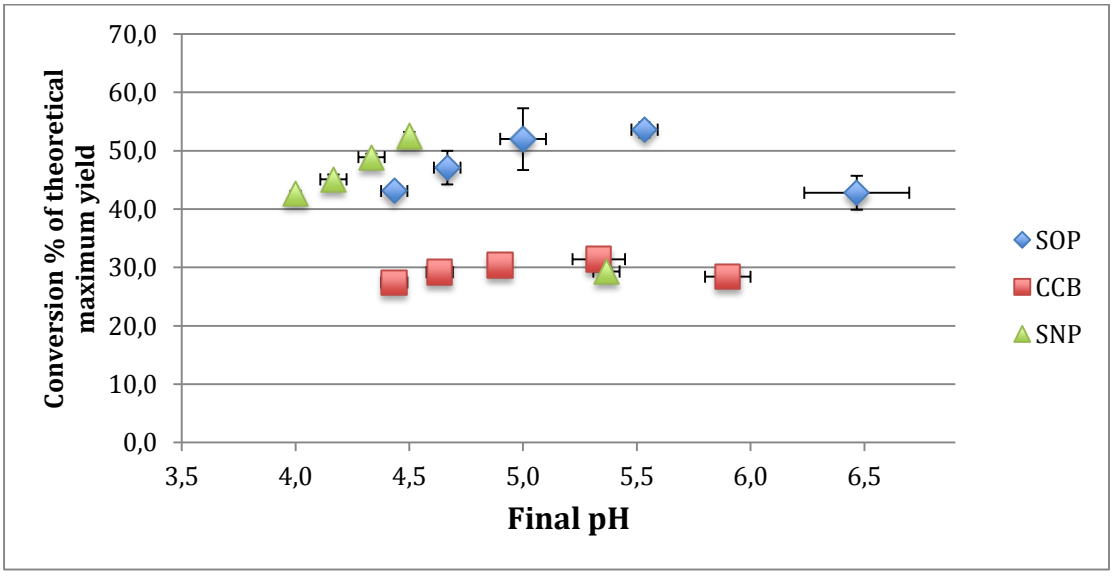


Figure 11. ***Optimal acid loading for SOP, CCB and SNP.*** *Reaction volume was 10ml with 5% DM (w/v) and an enzyme concentration of 5mg/g DM. Reaction was Incubated at 50°C in 20 mM Na-citrate pH 5 for 24 hours and DNS assay was used to quantify end products. Reactions were run in triplicates so standard deviation was illustrated.*

600µl citric acid results in the highest conversion for SOP (53.6%). Only 50µl was required to reach the highest conversion observed for CCB (31.4%) and SNP (52.6%) (Table 6). The data in Figure 11 indicates that Cellic-CTec2 had a pH optimum just over 5 for SOP and CCB. The optimal pH seemed to be relatively "wide" SOP and CCB; fluctuation in pH should not have drastic affects within deviations of pH5±1 (Figure 11). For SNP the optimum pH lies between pH 4.5 and 5.4. SNP do not seem to have the same wide range of tolerance, but was instead rather sensitive to pH fluctuation, it seems as if pH has a drastic effect on conversion of SNP when pH>5. Furthermore, it is important to notice how the pH shifted from the start to the end of the reactions in Table 6. SOP with an acid loading of 550µl had a pH shift from pH 4.8 to 6.5 during the course of the hydrolysis, and this results in a relative low conversion for SOP (42.9%). Higher acid loading reduced the pH shift; 750µl citric acid loadings in SOP resulted in a shift from pH 4.3 to 4.4 over the course of the reaction. This tendency was also observed for CCB and SNP, but the high acid loadings resulted in a low pH (close to pH 4 for all substrates) leading to low yields (optimal pH of Cellic CTec2 is pH5). From this experiment, the acid loadings, which resulted in the highest yields for SOP, CCB and SNP respectively, were used in the following experiments (600 µl for SOP, 50µl for CCB and 50µl for SNP in a total volume of 10 ml).

4.2.2. 24-hour hydrolysis yields with optimised pH-control conditions

After optimizing the acid loading, the 24-hour hydrolysis reactions with Cellic CTec2 were repeated with all substrates. In the reaction, optimal acid loadings determined in Section 4.2.1 were used, and the pH was measured before and after the reaction. The pH seems to be well controlled for each substrate, the final pH varied between 4.8 and 5.6 (Table 7). Increasing the pretreatment severity led to an increase in the conversion yield. From the SOP pretreated at 170°C for 7min, 45.3% of the theoretical maximum sugars were solubilized, whereas from SOP pretreated at 210°C for 14min, 54.0% of the theoretical maximum sugars were solubilized. Notably, in SOP pretreated for 170°C 7min, the conversion was higher than SOP pretreated for 170°C 14min (Figure 12).

Table 7. 24-hour hydrolysis with the optimised pH control. *Cellic CTec2 concentration was 5 mg/g and the reaction contains 5% dry mass. The table summarizes the yields, initial pH (before the reaction starts) and final pH (after the reaction was done). Reaction volume was 10 ml with 5% (w/v) dry mass and the samples were incubated in a thermo oven at 50°C for 24 hours with 20 mM Na-citrate pH 5 used as a buffer. pH was measured before and after the reaction . DNS assay was used to quantify end products.*

	Pretreatment	Conversion (% of theoretical maximum)	Initial pH	Final pH
SOP	170°C 7min	45.3 ± 2.6	4.6 ± 0.0	5.6 ± 1.1
	170°C 14min	43.7 ± 0.1	4.4 ± 0.1	4.8 ± 0.1
	190°C 7min	45.4 ± 1.3	4.5 ± 0.2	5.3 ± 0.2
	190°C 14min	51.9 ± 2.0	4.4 ± 0.0	5.0 ± 0.1
	210°C 7min	47.2 ± 1.7	4.5 ± 0.1	4.8 ± 0.1
	210°C 14min	54.0 ± 2.4	4.5 ± 0.0	4.9 ± 0.0
CCB	170°C 7min	36.3 ± 1.6	5.3 ± 0.1	5.7 ± 0.1
	170°C 14min	38.6 ± 0.6	5.2 ± 0.0	5.5 ± 0.1
	190°C 7min	38.3 ± 0.1	5.2 ± 0.1	5.5 ± 0.2
	190°C 14min	37.2 ± 0.2	5.2 ± 0.2	5.6 ± 0.4
	210°C 7min	44.2 ± 0.2	4.9 ± 0.0	5.2 ± 0.1
	210°C 14min	46.6 ± 2.1	4.7 ± 0.1	4.8 ± 0.1
SNP	170°C 7min	23.6 ± 1.4	4.7 ± 0.0	4.9 ± 0.1
	210°C 7min	22.1 ± 0.9	4.9 ± 0.0	5.1 ± 0.0
	210°C 14min	29.1 ± 0.6	4.8 ± 0.1	4.9 ± 0.0

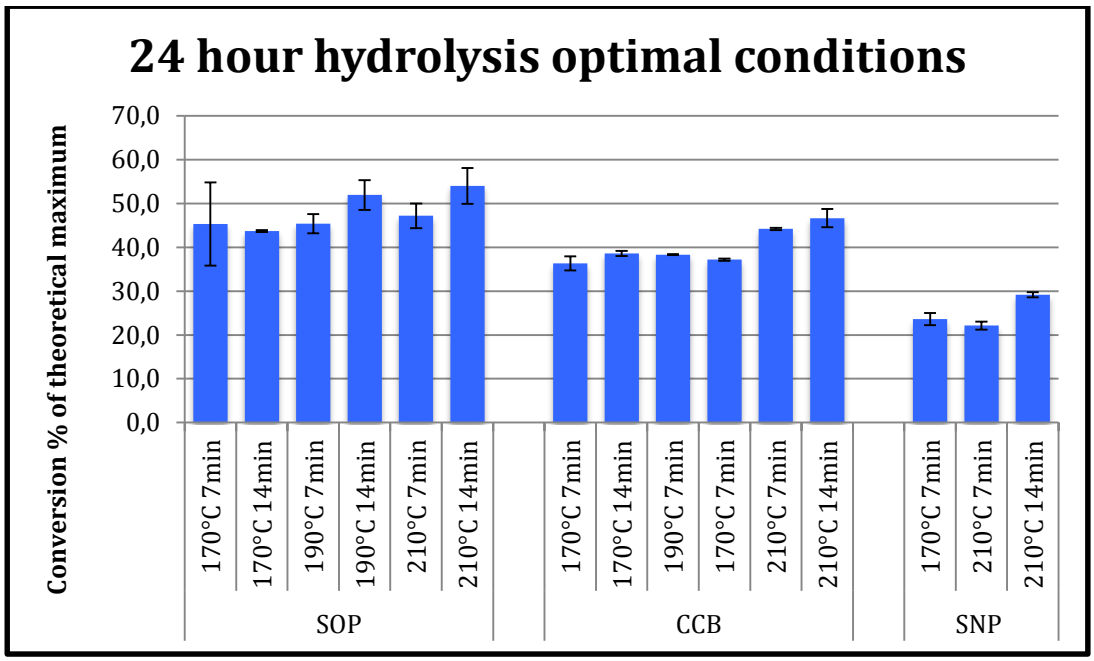


Figure 12. 24-hour hydrolysis of all substrates with Cellic CTec2 and various pretreatments. *Reaction contains a total volume of 10 ml with 5% DM and an Cellic CTec2 concentration of 5mg/g DM. The solution was incubated at 50°C in 20 mM Na-citrate pH5 for 24 hours. Reactions were run in triplicates so standard deviation was illustrated. DNS assay was used to quantify end products.*

From the CCB pretreated at 170°C for 7min, 36.3% of the theoretical maximum sugars were solubilized, whereas from CCB pretreated at 210°C for 14min, 46.6% of the theoretical maximum sugars were solubilized (Table 7). SNP pretreated at 170°C for 7min had 23.6% of the theoretical maximum sugars solubilized and SNP pretreated at 210°C for 14min had 29.1% solubilized.

Figure 12 shows a clear correlation between the amounts of lignin and the conversion yields. SOP, which contained small amounts of lignin (2%, Table 5) had the highest values of solubilized sugars. CCB had a higher carbohydrate-lignin ratio compared to SNP (observed in Section 4.1(Table 5)). CCB had about 5:1 ratios between carbohydrates and lignin. SNP, on the other hand, had about equal amounts of carbohydrates and lignin for untreated and pretreated samples. The effect of this difference was clear in Figure 12, where the conversion of SNP was lower than the conversion of CCB.

4.2.3. Maximizing saccharification yields

Cellic CTec2 contains a high diversity of cellulases and hemicellulases with CBMs (described in section 1.3.3 and 1.3.4). In the following experiments I tried to maximize the total yields by increasing the Cellic CTec2 loading and increase the duration of the incubation.

4.2.3.1. Conversion of all waste paper substrates at higher enzyme loading

In this experiment, Cellic CTec2 loadings were increased from 5 mg/g DM to 25mg/g DM. In the reaction, optimal acid loadings (determined in Section 4.2.1) were used, and the pH was measured before and after the reaction. The increased enzyme loading resulted in higher yields for SOP, CCB and SNP (Figure 13).

Table 8. Maximizing conversion of waste paper to monosaccharides*. The concentrations of Cellic CTec2 were increased 5 folds to observe how that affected the yield for SOP, CCB and SNP. The table summarizes the yields, initial pH (before the reaction starts) and final pH (after the reaction was done). Results were compared to the original 5mg/g Cellic CTec2 concentration experiment (Section 4.2.2). The respective samples contain the following acid loadings (found in Section 4.2.1):* 600µl of 1M sulfuric acid in SOP and 50µl in CCB and SNP. total volume was 10 ml and the samples was incubated at *50°C in 20 mM Na-citrate pH5 for 24 hours. DNS assay was used to quantify end products.*

	Pretreatment	Conversion (% of Dw) 5 mg/g	Initial pH	Final pH 5mg/g	Conversion (% of DW) 25 mg/g	Initial pH	Final pH
SOP	170° 7min	45.3 ± 9.5	4.6 ± 0.0	5.6 ± 1.1	55.4 ± 2.1	4.8 ± 0.1	5.3 ± 0,3
	170° 14min	43.7 ± 0.2	4.4 ± 0.1	4.8 ± 0.1	54.3 ± 4.3	4.6 ± 0.2	4.9 ± 0.1
	190° 7min	45.4 ±2.2	4.5 ± 0.2	5.3 ± 0.2	61.2 ± 0.6	4.7 ± 0.2	4.9 ± 0.3
	190° 14min	51.9 ± 3.4	4.4 ± 0.0	5.0 ± 0.1	57.1 ± 2.1	4.6 ± 0.3	5.1 ± 0.1
	210° 7min	47.1 ± 2.8	4.5 ± 0.1	4.8 ± 0.1	59.3 ± 5.0	4.7 ± 0.1	4.8 ± 0.1
	210° 14min	54.1 ± 4.1	4.5 ± 0.0	4.9 ± 0.0	68.2 ± 9.1	4.5 ± 0.0	4.8 ± 0.0
CCB	170° 7min	36.3 ± 1.6	5.3 ± 0.1	5.7 ± 0.1	41.3 ± 1.2	5.3 ± 0.1	5.8 ± 0.1
	170° 14min	38.7 ± 0.6	5.2 ± 0.0	5.5 ± 0.1	40.5 ± 1.9	5.5 ± 0.2	5.5 ± 0.0
	190° 7min	38.3 ± 0.1	5.2 ± 0.1	5.5 ± 0.2	45.7 ± 2.1	5.0 ± 0.0	5.4 ± 0.1
	190° 14min	40.0 ± 0.2	5.2 ± 0.2	5.6 ± 0.4	42.6 ± 0.2	5.2 ± 0.1	5.2 ± 0.2
	210° 7min	44.2 ± 0.2	4.9 ± 0.0	5.2 ± 0.1	44.3 ± 2.4	5.0 ± 0.2	5.3 ± 0.0
	210° 14min	46.6 ± 2.1	4.7 ± 0.1	4.8 ± 0.1	51.0 ± 3.4	4.7 ± 0.1	4.9 ± 0.1
SNP	170° 7min	23.6 ± 2.9	4.7 ± 0.0	4,9 ± 0.1	28.9 ± 3.4	4.8 ± 0.1	5.1 ± 0.2
	210° 7min	22.6 ± 1.9	4.9 ± 0.0	5.1 ± 0.0	29.8 ± 3.2	4.9 ± 0.2	5.0 ± 0.2
	210° 14min	29.2 ± 1.3	4.8 ± 0.1	4.9 ± 0.0	34.0 ± 0.9	4.7 ± 0.1	4.8 ± 0.0

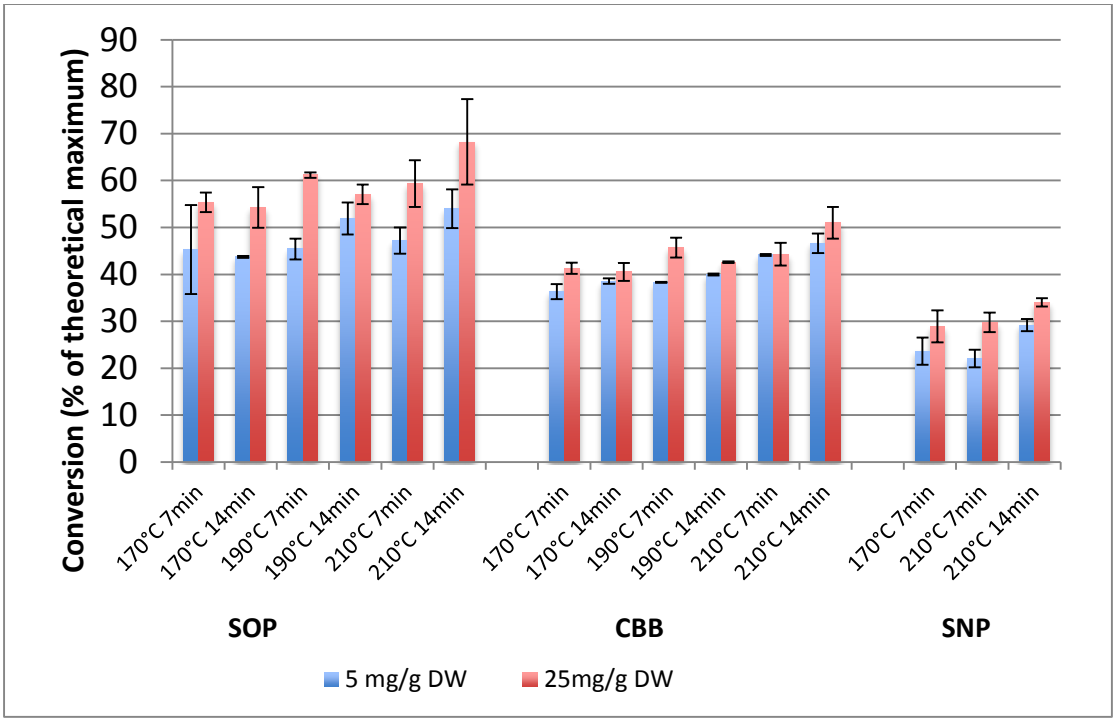


Figure 13. *Maximizing conversion of waste paper to monosaccharides. Blue bars are the 24-hour hydrolysis with 5mg/g DM Cellic CTec2 concentration (Table 9). Red bars represent increased Cellic CTec2 concentration (25mg/g DM) over 24-hours (table 10). Reaction contained 5% DM, Incubated at 50°C in 20mM Na-Citrate. Reactions were run in triplicates so standard deviation was illustrated.*

When the enzyme loadings were increased from 5 mg/g DM to 25 mg/g DM for all pretreated substrates, the average conversion for increased with 11.3% for SOP, 3.5% for CCB, and 5.7% for SNP (Table 8). In the most pretreated samples (210° 14min) SOP had the highest conversion (68.2%), CCB the second highest (51.0%) and SNP had the lowest conversion (34.0%).

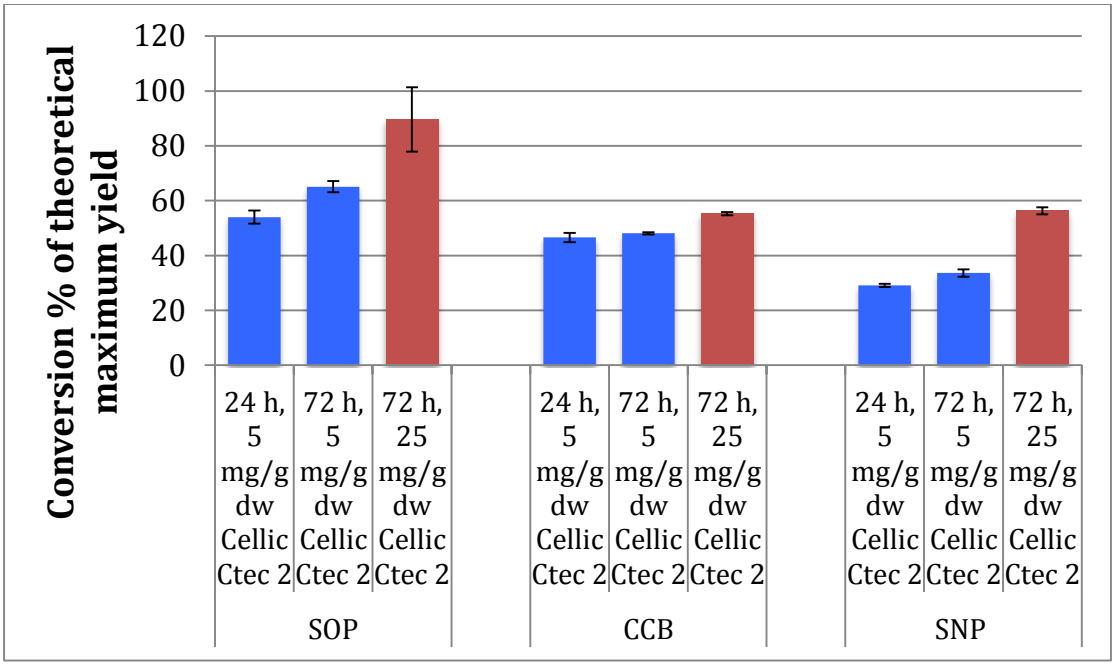
4.2.3.2. Conversion of the most severely pretreated samples at higher enzyme loading using longer incubation times

Next, increased incubation were performed to observe if this could further increase the yields. In the reaction, optimal acid loadings (determined in Section 4.2.1) were used, and the pH was measured before and after the reaction. The experiments were performed on the most severe pretreated substrates (210°C 14min).

In the first experiment, 5mg/g DM Cellic CTec2 loadings were used and the solutions were incubated for 72 hours (Table 9). From SOP, 65.1 % of the theoretical maximum sugars were solubilized. From CCB, 48.1% was solubilized and for SNP, 33.6% of the theoretical maximum sugars were solubilized (Table 9).

Table 9. Increased enzyme loading (25mg/g DM) and longer incubation times (72 hours). *Increased Cellic CTec2 concentration and incubation time was tested to observe how they affected the conversion. The table summarizes the yields, initial pH (before the reaction starts) and final pH (after the reaction was done). All substrates were from the most pretreated samples (210°C 14min). Reaction contained 5% DM and a total volume of 10ml, Incubated at 50°C in 20mM Na-Citrate. DNS assay was used to quantify end products.*

	Duration	Enzyme concentration (mg/g DM)	Conversion % of theoretical maximum yield	Initial pH	Final pH
SOP	24 hours	5	54.0 ± 2.4	4.5 ± 0.2	4.9 ± 0.2
	72 hours	5	65.1 ± 2.0	4.5 ± 0.1	5.2 ± 0.2
	72 hours	25	89.6 ± 11.7	4.6 ± 0.2	5.3 ± 0.2
CCB	24 hours	5	46.6 ± 1.7	4.7 ± 0.1	4.8 ± 0.1
	72 hours	5	48.1 ± 0.4	4.9 ± 0.0	5.4 ± 0.1
	72 hours	25	55.2 ± 0.6	4.6 ± 0.1	4.9 ± 0.1
SNP	24 hours	5	29.1 ± 0.6	4.8 ± 0,1	4.9 ± 0,0
	72 hours	5	33.6 ± 1.3	4.8 ± 0.1	5.0 ± 0.1
	72 hours	25	56.3 ± 1.3	5.0 ± 0.1	5.3 ± 0.1



Figur 14. *Longer incubation time and higher Cellic CTec2 concentrations increases the yield significantly. Tests were performed with 5% DM. Blue bar marks 5mg/g DM enzyme concentrations. And red marks 25mg/g DM. Reactions were run in triplicates so standard deviation was illustrated.*

From SOP, 65.1 % of the theoretical maximum sugars were solubilized

In the second experiment Cellic CTec2 loadings was increased to 25mg/g DM and incubation times was 72 hours. This affected the conversion considerably, SOP (Figure 14), solubilized 90% of theoretical maximum sugars, from CCB, 55% of the theoretical maximum sugars were solubilized and SNP solubilized 56% of theoretical maximum sugars.

SOP resulted in best conversion of the three substrates. SNP showed reduced yields compared to CCB in the 24-hour hydrolysis with optimal conditions (Figure 12) and the experiment with increased Cellic CTec2 concentration (Figure 13), but it shows high yields in acid loading test (Figure 11). This could be connected to the pH and it will be discussed in section 5. From these experiments, it was observed that the content of Lignin and hemicellulose seems to be the main cause for enzymatic recalcitrance. Furthermore the data displayed in Figure 14, show that the enzymatic conversion with Cellic CTec2 required more than 24 hours to obtain maximum yields.

4.3 Protein purification:

*Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A were purified in two steps. First, an ion exchange chromatography was conducted according to Section 3.5.2. The fractions collected after elution was analyzed on SDS-PAGE, as described in Section 3.6. After evaluating the results from SDS-PAGE analysis, the samples were concentrated (according to Section 3.5.3) and further purified (if necessary) by SEC, according to Section 3.5.4.

To confirm the expression of *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A from *P. pastoris,* an SDS-PAGE analysis of the supernatant in the respective broths were performed (figure 15).

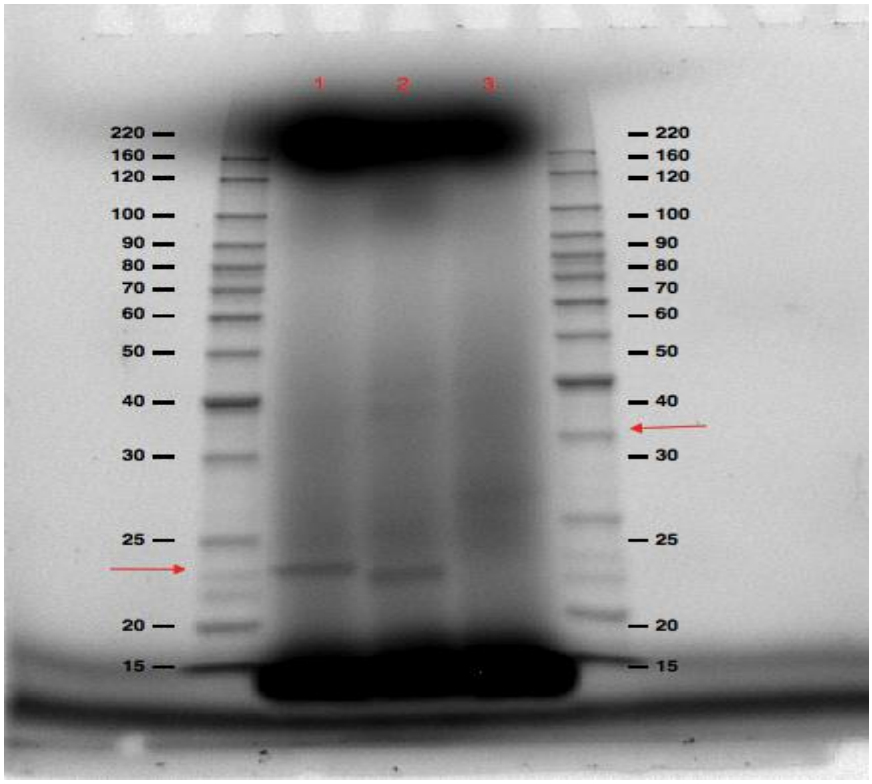


***Figure 15**. **SDS_PAGE.** Analysis of the supernatant in the YPD medium after 24 hours incubation, putatively containing AfCel12A (lane1), MaCel45A (lane2) and TaCel5A (Lane 3). 10µl supernatant were loaded to each lane 1 (according to Section 3.6). 3 µl of benchmark proteinladder Invitrogen marks the molecular weight. A red arrow on the left marks where MaCel45A (22.9kDa) and AfCel12A (24.1kDa) were located (almost the same molecular weight). The red arrow on the right marks where TaCel5A (35,1kDa not observed) should be. The broths were incubated in YPD medium at* 30°C *for 24 hours in the harbinger system.*

The faint band at approximately 23kDa (Lane 1, Figure 15) could represent *Ma*Cel45A, which have a molecular weight of 22.9kDa. The faint band at approximately 23Kda in lane 2 (Figure 15) could represent *Af*Cel12A, which have the molecular weight of 24.2kDa. Expression of *Ta*Cel5A should result in a band at 35.1kDa in lane 3 (Figure 15), but this was not observed.

The two broths potentially containing *Ma*Cel45A (Lane 1) and *Af*Cel12A (Lane 2) were purified with IEC. After IEC and subsequent determination of protein concentration (according to section 3.5.5) it was concluded that the broths contained very small amounts of proteins.

P. pastoris were grown in a harbinger system, this system provided great yields for colleagues working with *E.coli.* This was, however, not observed in the expression of *P. pastoris* in this project. Several test trying to increase the yields were performed; Incubation for longer period of time (between 1-3 days) and increasing the glucose concentration (explained in section 3.4.3), did not result in increased yields. Because of the low levels of enzymes, it was decided that broths within the Waste2Go project by A. Varnai should be used. These broths have been produced from the same *P. pastoris* production strain in a 750L fermenter. After fermentation, the cells were removed and the cell-free supernatant was buffer exchanged to 50mM Na-acetate pH5 and up-concentrated 10 times overall. The respective broths contained high concentrations of the respective enzymes (*Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A). As a result, the following purifications and experiments were performed with *Ma*Cel45A, *Ta*Cel12A and *Af*Cel12A from the donated broths to achieve comparable results.

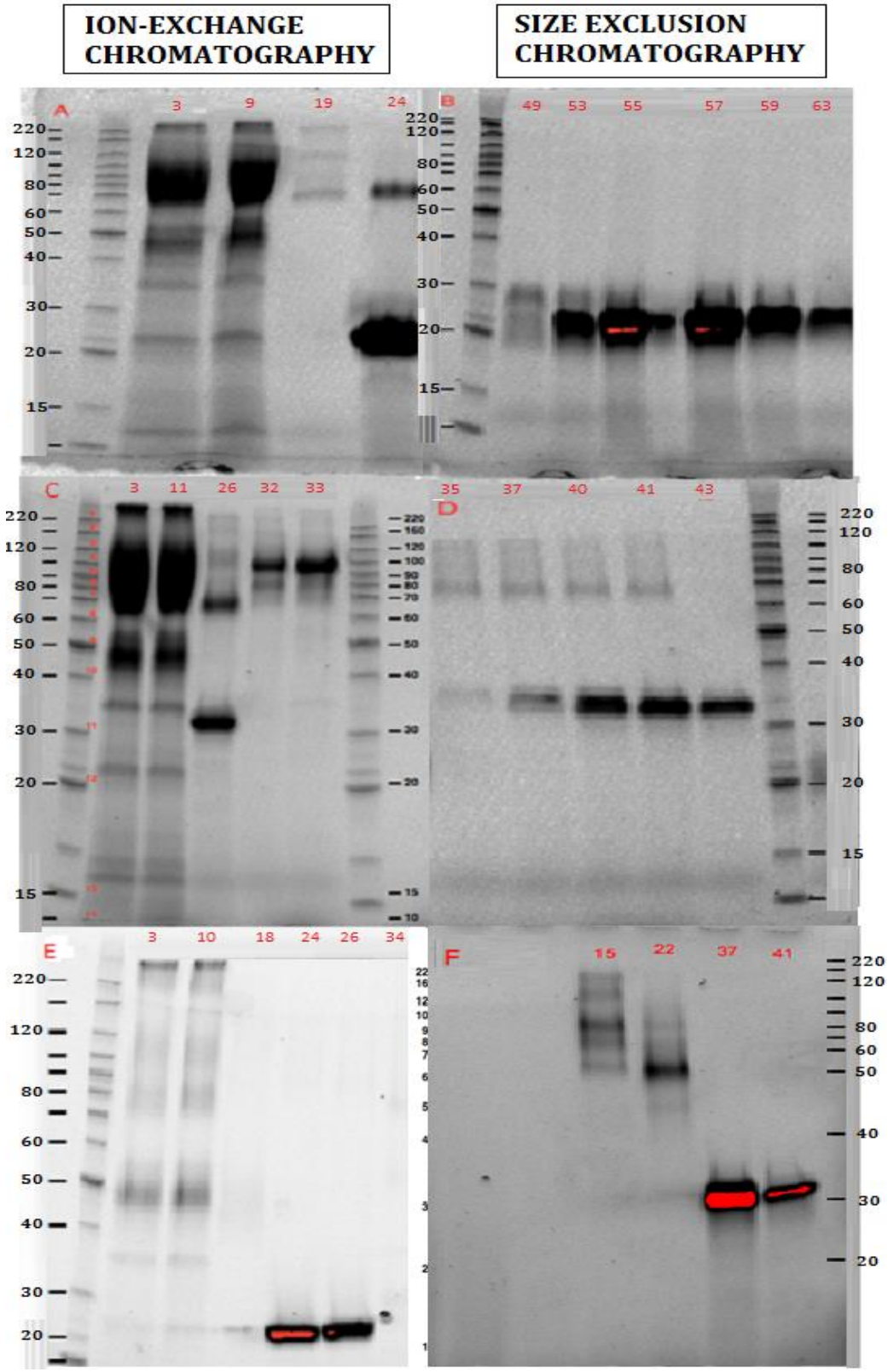


Figure 16. Purification of MaCel45A, TaCel5A and AfCel12A. *Panels to the left show SDS-PAGE analysis after purification with IEC. Panels to right show SDS-PAGE analysis after SEC. Panels A and B were the results after purification of MaCel45A. Panels C D and F show the results after purification of TaCel45A. Panel E was the result after IEC with AfCel12A (only IEC was required). 10µl supernatant were loaded to each lane. 3 µl of Benchmark protein ladder (Invitrogen) marks the molecular weight. The respective proteins are found at: MaCel45A: 22.9kDa, TaCel5A: 35,1kDa and AfCel12A 24.1kDa. Chromatograms of the respective runs are found in Appendix*

*Ma*Cel45A was not completely purified after IEC as another protein was observed on the SDS-PAGE at 70 kDa (Figure 16A). Some *Ma*Cel45A was observed in the flow through in lane 3 and 9 at 23kDa (panel A, Figure 16), but the amount was so small that they were considered negligible. Fractions containing the *Ma*Cel45A (fractions 22-31 from the chromatogram in Appendix Figure D.1) were pooled and concentrated. The size difference between *Ma*Cel45A and the unknown molecule at 70kDa (Panel A, Figure 16) enabled separation with SEC (SDS-PAGE, Figure 16B) protein at 70kDa was removed. *Ma*Cel45 A was not completely purified after SEC, (Figure 16B) show a double band (very apparent in lane 49), indicating that there might be another protein with similar mass (23kDa) as *Ma*Cel45A. This could also be non-processed *Ma*Cel45A with the signal peptid attached. It was concluded that the small concentration of the unknown protein was insignificant for the planned further studies, so no further purification was done. Fractions from 53-63 (Appendix Figure D.4) were collected and concentrated until the final volume of 12.5 ml with a protein concentration of 2.79 g/L, as measured according to Section 3.5.5.

Figure 16C show the SDS-PAGE analysis of *Ta*Cel5A after IEC. SDS-PAGE shows a protein at 70kDa in the eluted fractions together with *Ta*Cel5A. This seems similar to the protein found in *Ma*Cel45A, and the significant mass difference enabled a separation on SEC. The chromatogram of the first run on SEC can be viewed in appendix D.4. The chromatogram in D.4 shows an unsuccessful separation between the unknown protein and *Ta*Cel45A (peaks in the chromatogram merge); this was reflected in The SDS-PAGE (panel D in figure 16). The unknown protein at 70kDa was still visible in fractions 35, 37,40 and 41. Fractions from 43 to 51 contained pure TaCel5A so the fractions were pooled. Fractions from 36-42 were rerun on the SEC and the chromatogram is found in Appendix D.5. SDS-PAGE in figure 16F shows a good separation after SEC, and fractions from 32 to 46 were pooled and concentrated (according to section 3.5.3) together with the pooled fractions from the first run. The fractions were concentrated to a final volume of 2.5 ml with a protein concentration of 2.7g/l, measured according to section 3.5.5.

AfCel12A was well separated after IEC (Appendix, Figure D.3). No contaminating proteins were observed in fractions 18-34 on the SDS-PAGE (Figure 16E). Fractions 23 (where the peak starts in figure D.3) to 34 were pooled and concentrated to a total volume of 2.1 ml with a protein concentration of 4.2 g/L. SEC was not required.

4.4 Production of oligosaccharides

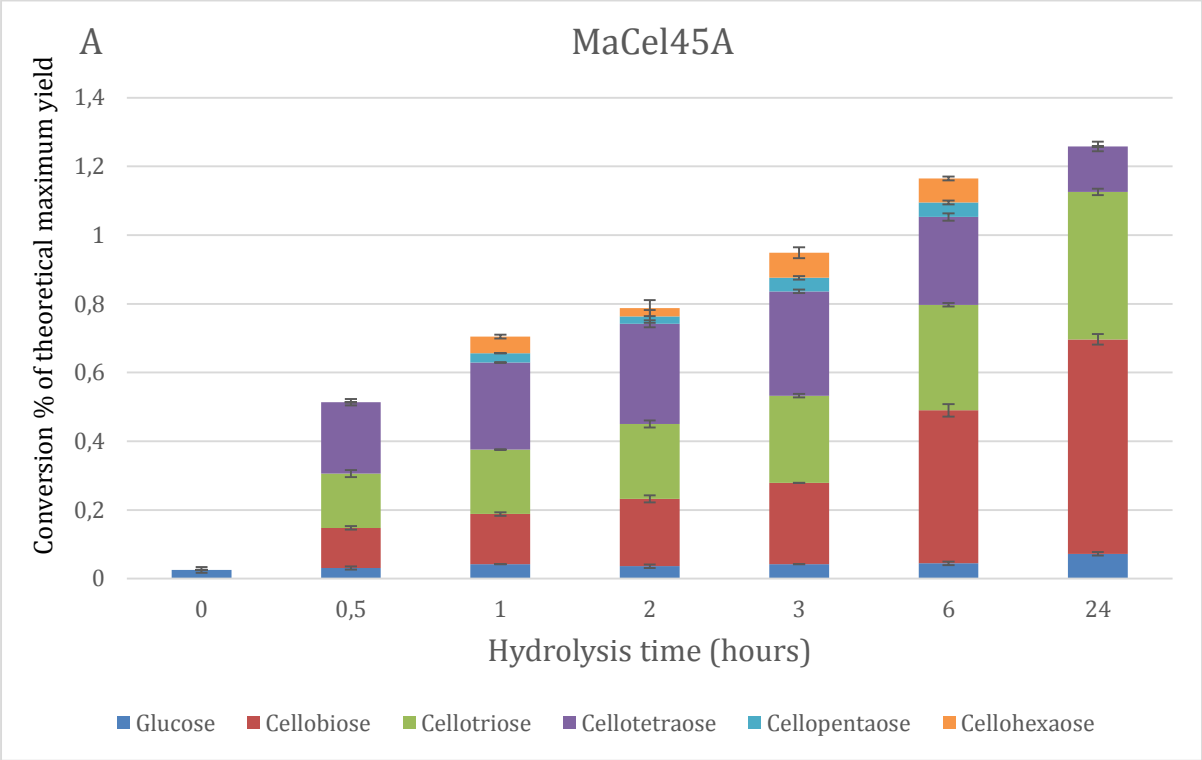
The endoglucanases *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A were produces from recombinant *P. pastoris* strains and purified from culture broths. These three enzymes are endoglucanases (section 1.3.1), which are supposed to introduce random cuts in the amorphous regions of cellulose and may release soluble cello-oligosaccharides (cellodextrines) of various lengths varying from DP1 to DP6. These cello-oligosacchLonger cello-oligosaccharide products will not be detected in the soluble fractions

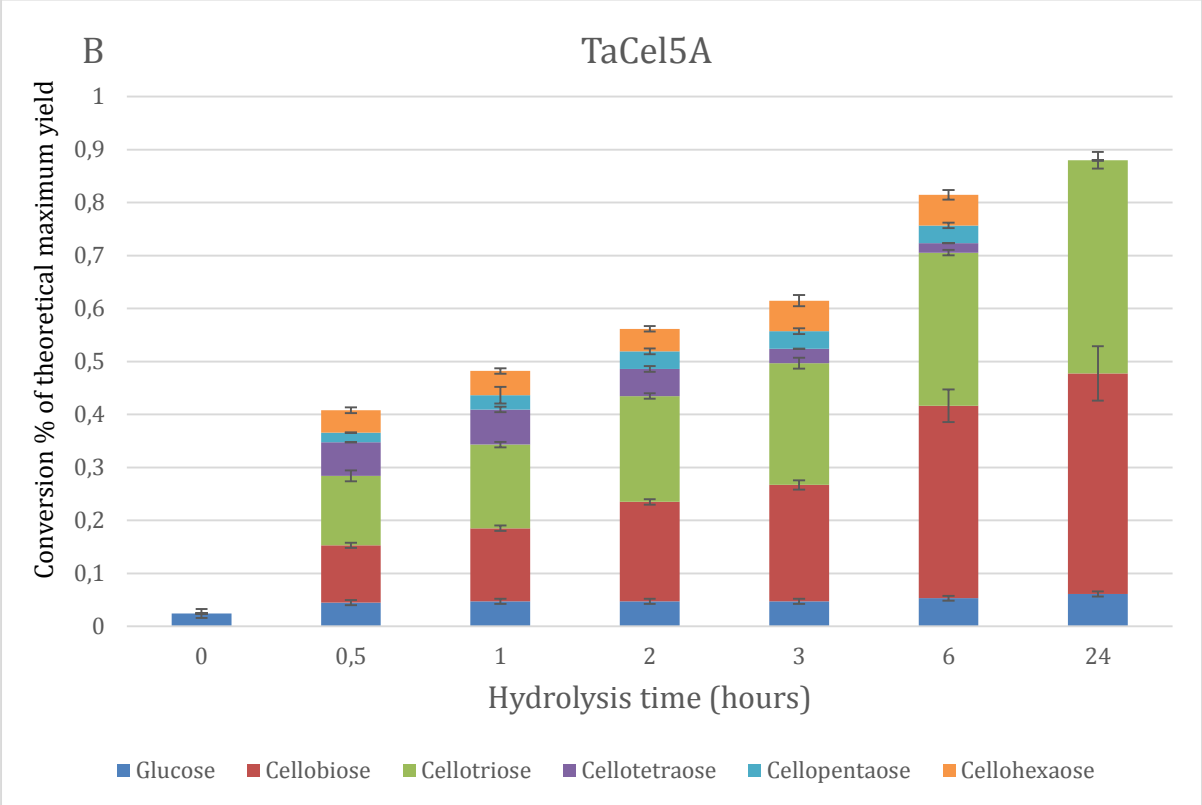
4.4.1. Time course experiments

Time course experiments were conducted to show how enzymes activity changes over time. Furthermore, it was hypothesized that longer soluble cellodextrines (DP 4-6) would be observed within the first hours of the experiment and less cellodextrines would be observed after 24 hours. This was perhaps to be expected because, as endoglucanases work on the insoluble substrates, they release soluble cellodextrines (DP 2-6). As soon as the endoglucananses struggle to find accessible insoluble cellulose (amorphous regions), they will further degrade the released soluble dextrins (DP 2-6), which would result in only short cellodextrines (DP2-3). In other words, cellodextrines with DP 4-6 would be expected within the first hours of the hydrolysis, and shorter chains with DP 2 and DP 3 would be expected after 24 hours. Conversions in the following experiments were calculated in % of theoretical maximum yield, maximum yield was calculated from the total cellulose content in the starting material (SOP, CCB and SNP) in table 5; Avicel was presumed to contain 100% cellulose. The only products measured where glucose and glucose oligomers.

4.4.1.1. Time course experiment with Avicel

In the time course experiments Avicel was included as a reference material. Avicel consists of highly crystalline material. therefore, low yields were expected compared to pretreated SOP, CCB and SNP.





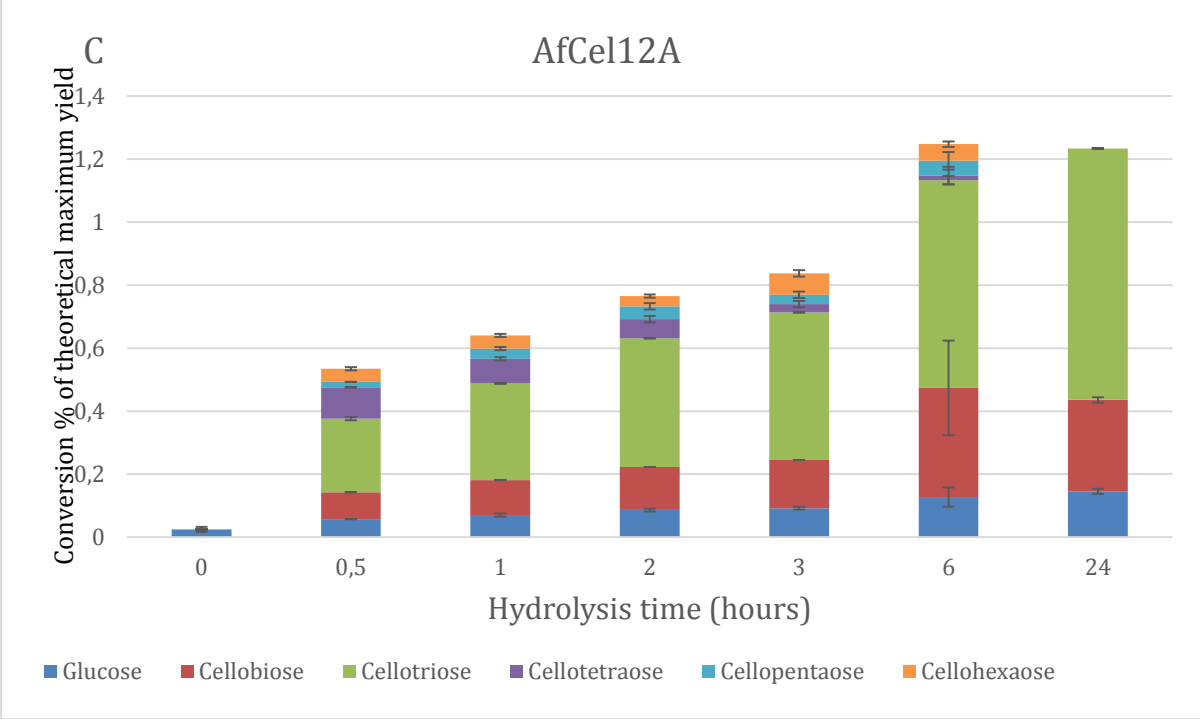


Figure 17. Conversion of Avicel with MaCel45A, Tacel5A and AfCel12A. *Formation of glucose and cellodextrines (DP 2-6) from Avicel with the three purified enzymes MaCel45A (A), TaCel5A (B), and AfCel12A (C). The reaction mixtures contained 1% DM and an enzyme concentration of 0,5mg/g DM in 20 mM Na-citrate pH5. The reaction was Incubated for 24 hours at 50°C in a thermo mixer with 1000 rpm shaking. The amount of release glucose and glucose oligomers (soluble) were analyzed over 7 time points and quantified with HPLC (according to Section 3.10.2.2). Each reaction was carried out in triplicates (error bars represent standard deviation).*

The primary products from the hydrolysis of Avicel with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A, were cellobiose and cellotriose, with relative small concentrations of celloteraoses, cellopentaose and cellohexaoses. *Ma*Cel45A showed a relative high yield of cellotetraose compared to *Ta*Cel5A and *Af*Cel12A

The panels in Figure 17 clearly show that the production formation rate decreases dramatically after the initial 30 min. However, the enzymes hydrolyzed steadily as time progressed, releasing a range of cellodextrines from glucose to cellohexaoses. The enzymes were still active after 6 hours, but as they run out of accessible chains on the amorphous regions of cellulose, they started degrading the released cello-tetraose, cellopentaose and cellohexaose, to shorter chains. Figure 17C illustrates the hydrolysis of AfCel12A; after 6 hours, the estimated yields were higher than the observed yields after 24-hours. This was obviously wrong, but the high standard deviations on cellobiose may account for this (additional experiments would most likely lower amount of cellobiose and reduce the standard deviation).

For *Ma*Cel45A, 1.23% of the theoretical maximum cellulose was solubilized in to glucose and cellodextrines after 24 hours (Figure 17 A). *Ta*Cel5A solubilized 0.88% of the theoretical maximum cellulose in Avicel (Figure 17 B), whereas *Af*Cel12A solubilized 12.2% (Figure 17 C), almost the same yield as observed for *Ma*Cel45A. However, the hydrolysis with *Ma*Cel45A results in higher concentration of longer cellodextrines (DP 4-6)(Figure 17A) when compared to *Ta*Cel5A and *Af*Cel12A (higher concentration of cellobiose and triose).

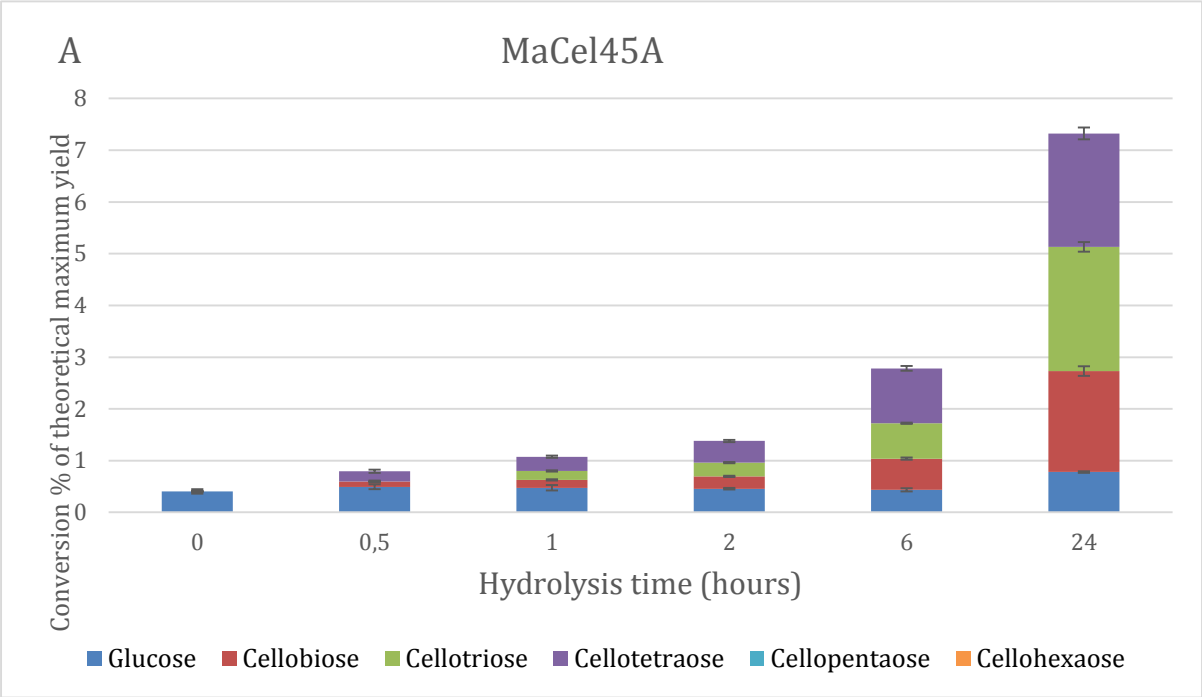
The generation of cellotetraoses from the hydrolizis with *Ma*Cel45A was high, even after 30 min. However, *Ma*Cel45A did not produce cellopentaose or cellohexaose after 30 minutes, this was only observed in the hydrolysis with TaCel5A and *Af*Cel12A. Longer cellopentaoses and cellohexaoses were first observed after 1 hour in the hydrolysis with *Ma*Cel45A.

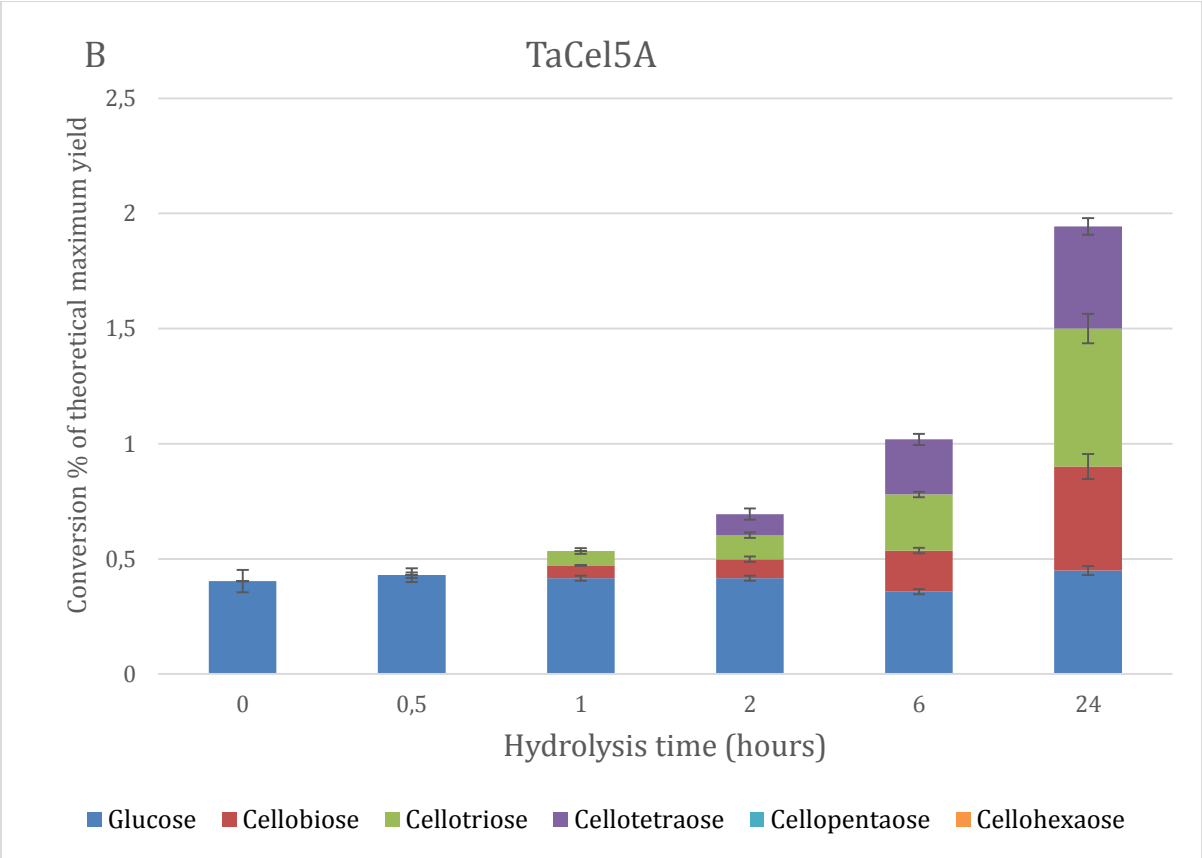
Cellopentaoses and cellohexaoses remain about the same concentration for all substrates until 6 hours. The shorter chains DP 2 and 3 increases over the same time period, indicating that the endoglucanases hydrolyze the soluble cellopentaoses and cellohexaoses at the same rate as they release new one from the insoluble cellulose. The results indicate that the concentration of cellopentaoses and cellohexaoses reached close to maximum after just 1 hour for each substrate.

The data in Figure 17 indicates that the endoglucananses had different properties and that *Ma*Cel45A was the best enzyme in the productions of longer cellodextrines (DP 4,5,6 mostly because of the high concentrations of cellotetraose). *Af*Cel12A produced less cellotetraose but about the same amount of cellohexaoses and cellopentaoses (Figure 17). *Ta*Cel5A was the least effective enzyme with a total yield below 1%. If we want to produce high yields of cellotetraose, cellopentose and cellohexaose, *Ma*Cel45A seems to be the best choice if we retrieve the samples after 1-hour. If efficient yields of cellopentaoses and cellohexaoses are desired, *Af*Cel12A was the best choice as it produced all cellopentaose and hexaose after just 30 minutes (figure 17C). Notably, the results indicate that endoglucanases are unable to further degrade cellotriose, if they do, they do so at a much slower rate compared to longer cellodextrines (DP 4, 5, and 6).

4.4.1.2. Time course experiment with SOP

SOP was hydrolyzed with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A and the product formation was analyzed over 5 time points. The product profile was analyzed with HPLC.





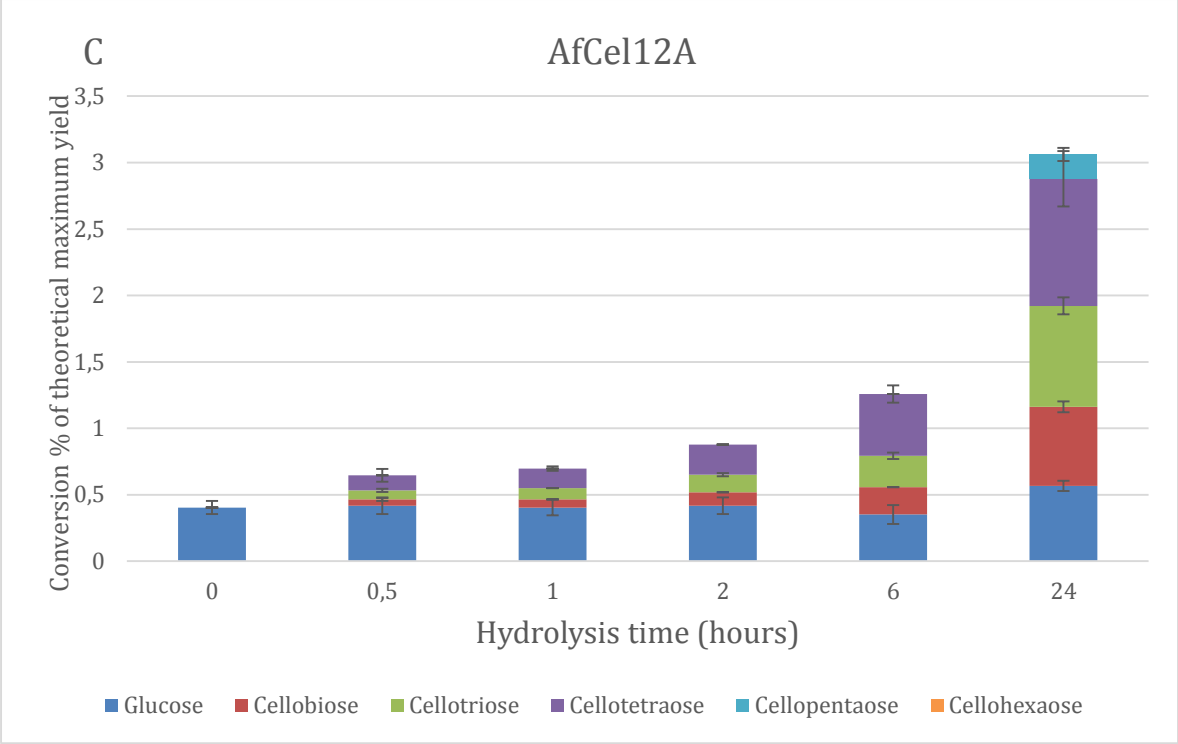
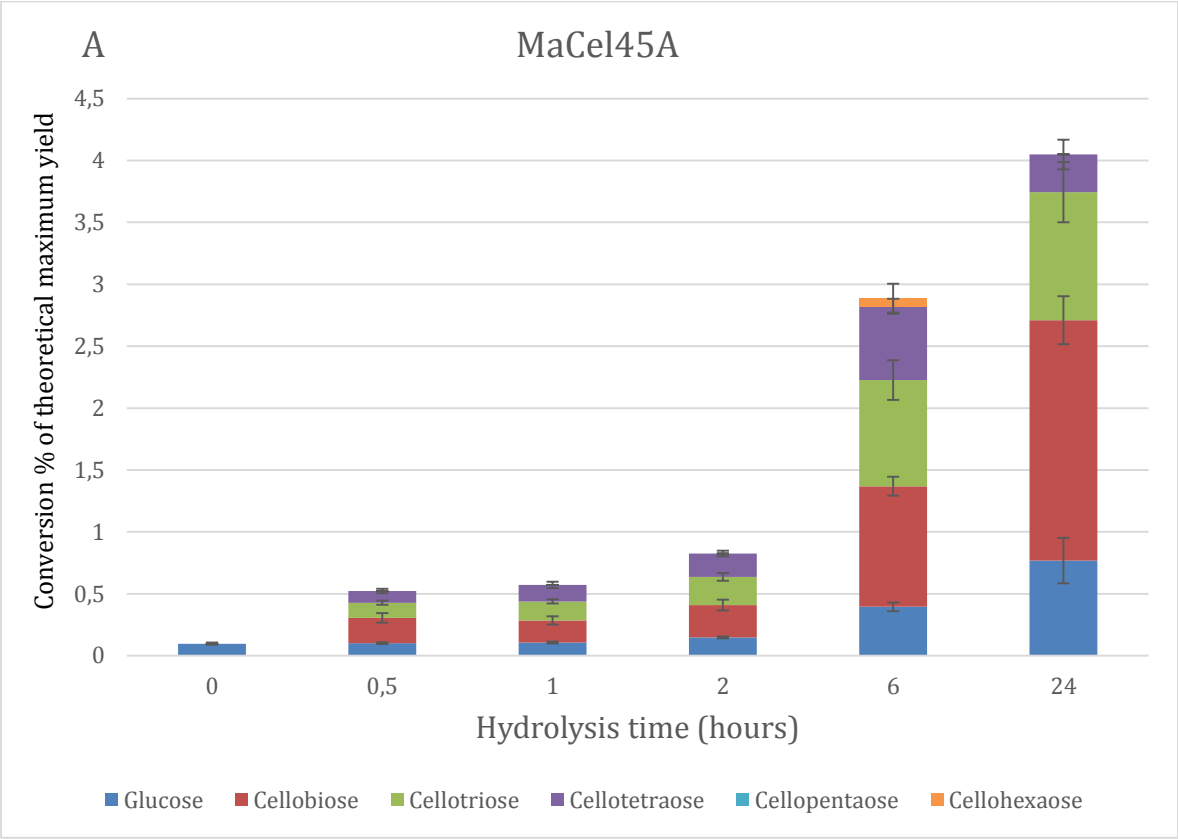


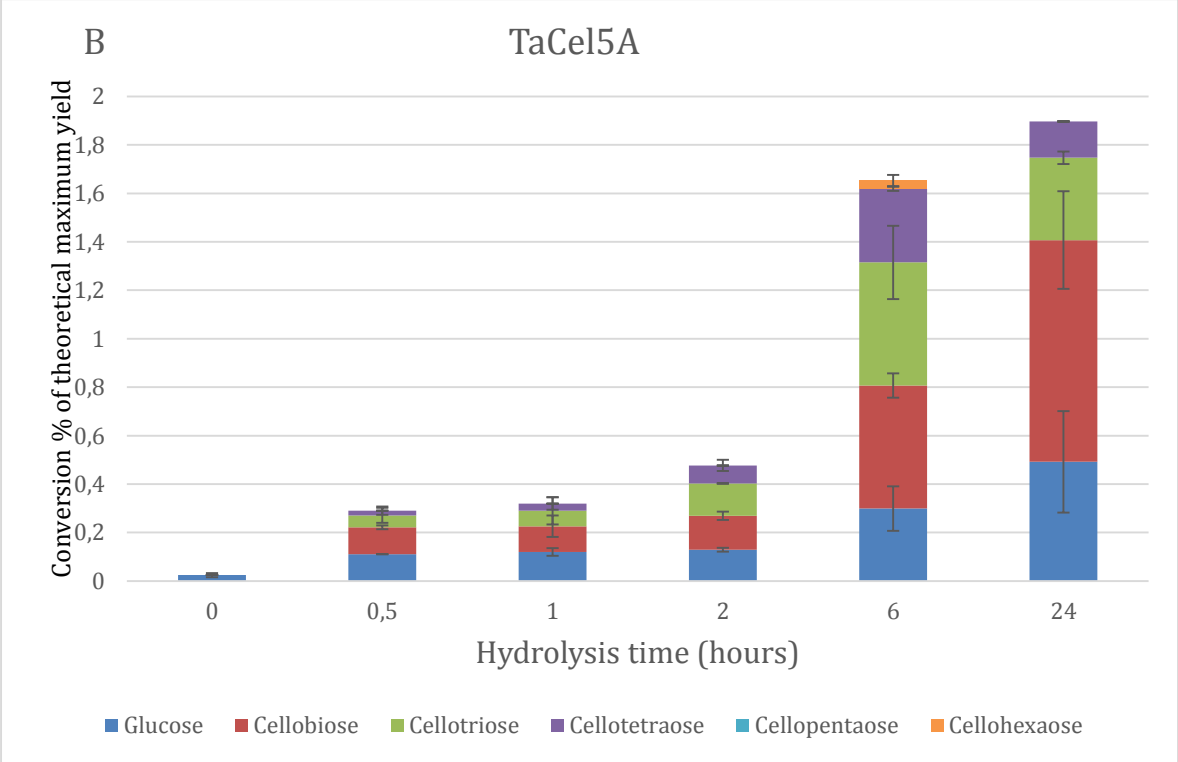
Figure 18 Conversion of pretreated SOP (210°C for 14 min) with MaCel45A, Tacel5A and AfCel12A *The three panels show the formation of Glucose and soluble Cellodextrines (DP 2-6) released from SOP after hydrolysis with the three purified enzymes MaCel45A (A), TaCel5A (B) and AfCel12A (C) Reaction mixtures contained 1% DM and an enzyme concentration of 0,5mg/g DM in 20 mM Na-citrate pH5. The reaction was incubated at 50°C in a thermo mixer with 1000 rpm shaking. The amount of release glucose and glucose oligomers (soluble) were analyzed over 6 time points and quantified with HPLC (according to Section 3.10.2.2). Each reaction was carried out in triplicates (error bars represent standard deviation*

*Ma*Cel45A was clearly the most efficient enzyme in the hydrolysis of SOP pretreated at 210°C for 14 min. Hydrolysis of SOP with *Ma*Cel45A solubilized 7.3% of the theoretical maximum cellulose (estimated from SOP pretreated at 210°C for 14 min in figure 5) after 24 hours (Figure 18 A). The yield was almost 3 times as high compared to AfCel12A (3% conversion) with the second highest yield. Interestingly, no cellopentaose or cellohexaose (DP 5 or 6) was observed at any time points with *Ma*Cel45A or *Ta*Cel5A. *Af*Cel12A released some cellopentaose (0.3%) after 24 hours (illustrated in Figure 18). All three enzymes showed significant increase in the yields between 6 and 24 hours. This was unexpected compared to the results from the hydrolysis of Avicel (Figure 17), where yields seem to be close to the maximum level after 6 hours. The significant increase between 6 and 24 hours could indicate that the yields could have come much higher if the duration of the experiments was increased.

4.4.1.3. Time course experiment with *CCB*

CCB was hydrolyzed with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A and the product formation was analyzed at 5 time points. The product profile was analyzed with HPLC.





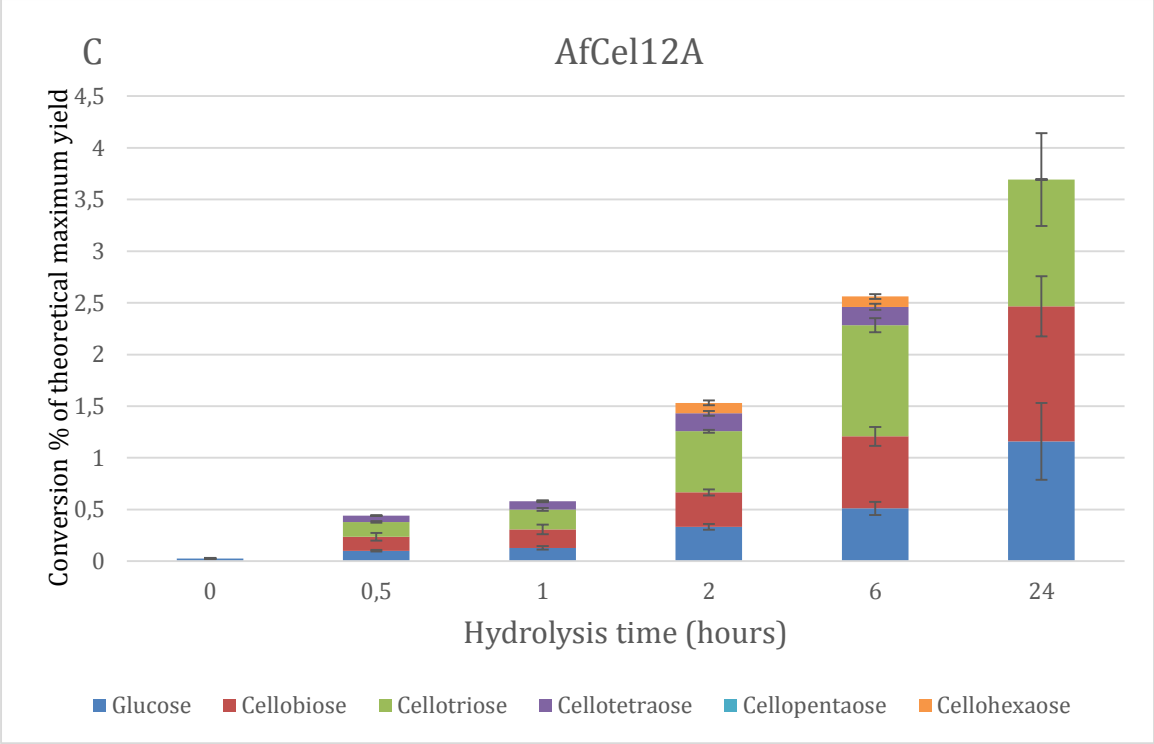


Figure 19 Conversion of pretreated CCB (210°C for 14 min) with MaCel45A, Tacel5A and AfCel12A *The three panels show the formation of glucose and soluble cellodextrines (DP 2-6 released from CCB after hydrolysis with the three purified enzymes MaCel45A (A), TaCel5A (B) and AfCel12A (C). Reaction mixtures contained 1% DM and an enzyme concentration of 0,5mg/g DM in 50°C in 20 mM Na-citrate pH5. The reaction was incubated at 50°C in a thermo mixer with 1000 rpm shaking. The amount of release glucose and glucose oligomers (soluble) were analyzed over 5 time points and quantified with HPLC (according to Section 3.10.2.2). Each reaction was carried out in triplicates (error bars represent standard deviation*

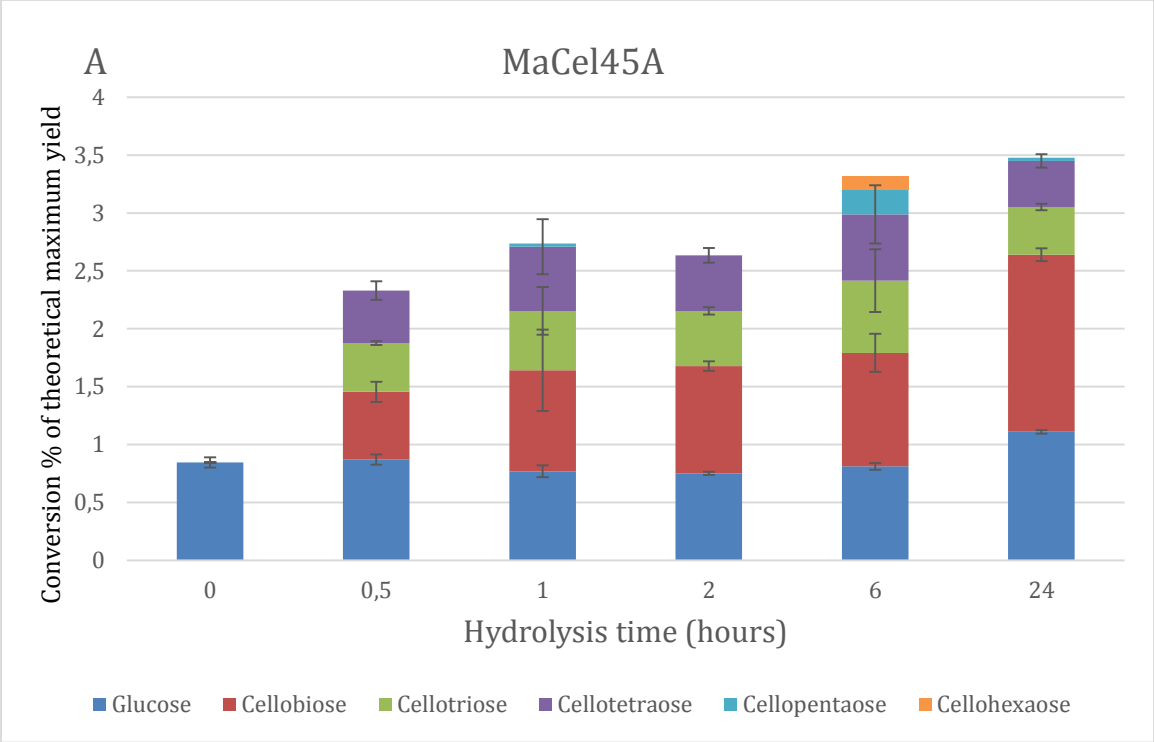
*Ma*Cel45A was the most efficient enzyme in the hydrolysis of CCB pretreated at 210°C for 14 min. *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A produced the cellodextrines: DP 2, 3, 4 and 6 within the first 6 hours (visualized in Figure 19). After 6 hours, the cellohexoases and cellopentaoses released seemed to get further hydrolyzed instantly to shorter cellodextrines because no cellohexaoses and cellopentaoses were detected after 24 hours of hydrolysis. The reason why we observe cellohexaoses and cello pentaoses before 6 hours was because of an equilibrium formed in the early phases of the hydrolysis. This equilibrium was formed when endoglucanases hydrolyzed the soluble cellopentaoses and cellohexaoses at the same rate as they release new one from the insoluble cellulose.

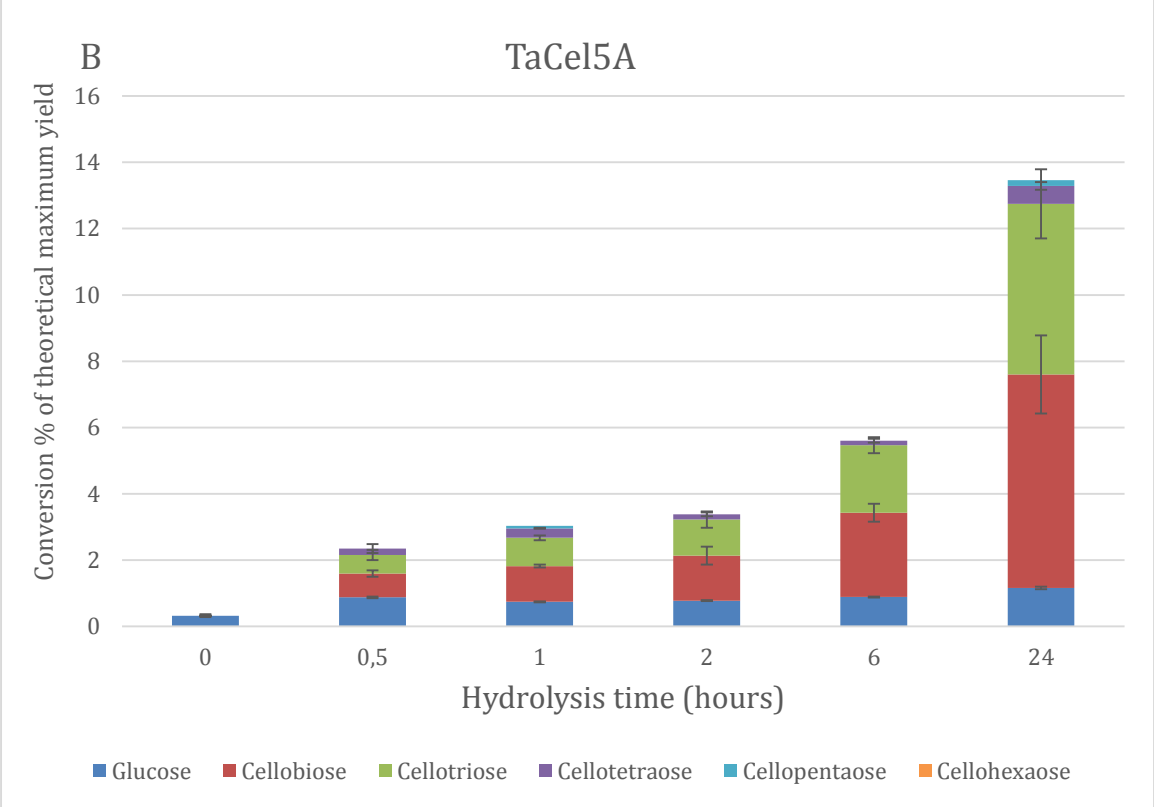
The total yield was somewhat lower for *Ma*Cel45A on CCB (4% of the theoretical maximum cellulose were solubilized) (Figure 19 A) compared to the conversion on SOP (7.3 % of the theoretical maximum cellulose were solubilized) (Figure 18 A). *Af*Cel12A, However, seem to have a better conversion after 24 hours on CCB (3.7%,

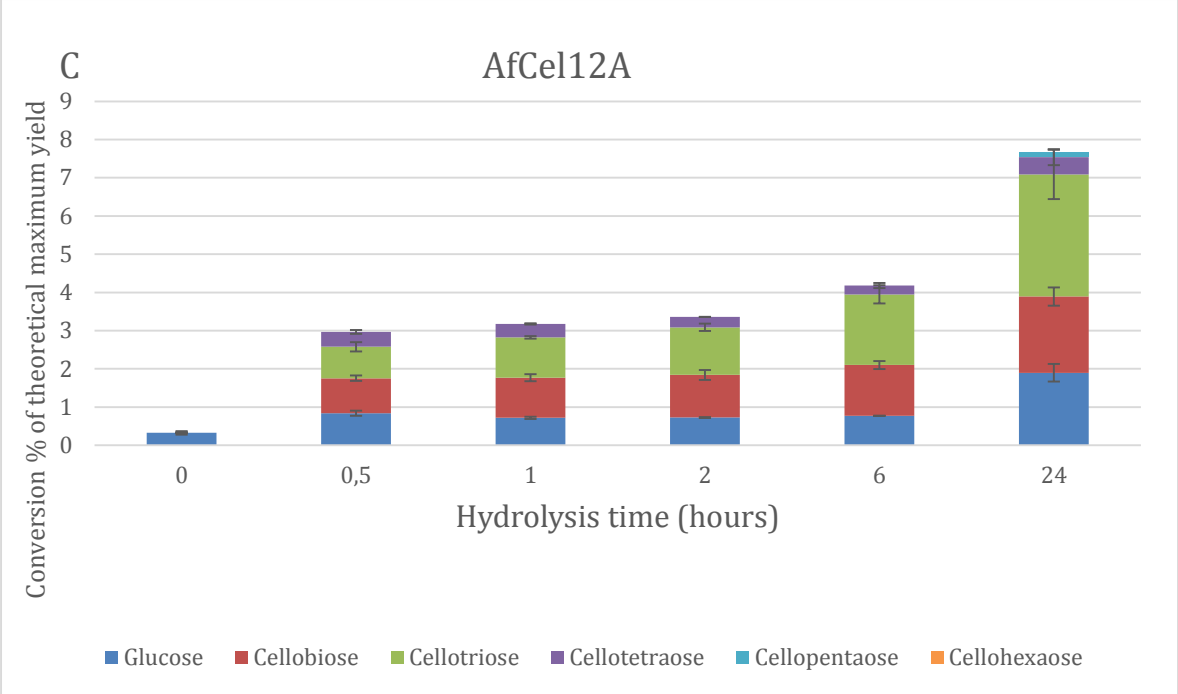
Figure 19 C) than on SOP (3%, Figure 18 C). The hydrolysis of *Ta*Cel5A on CCB (1.9%, Figure 19 B) resulted in the same yields as observed on SOP after 24 hours (1.9%, Figure 18 B). *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A were more effective initially (within 6 hours) on CCB (Figure 19) compared to the hydrolysis on SOP. If we assume that we reached 100% of the theoretical maximum yield (hydrolyzed all available carbohydrates) after 24 hours, *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A hydrolyzed between 70-90% of the final product amount after 6 hours in CCB. In SOP, *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A only managed to release between 30-50% of the final product amount after 6 hours. This means that the enzymes work slower on SOP or something may retard them. This will be discussed in section 5.

4.4.1.4. SNP

SNP was hydrolyzed with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A and the product formation was analyzed at 5 time points. The product profile was analyzed with HPLC.







Figur 20 Conversion of pretreated SNP (210°C for 14 min) with MaCel45A, Tacel5A and AfCel12A *The three panels show the formation of glucose and soluble cellodextrines (DP 2-6) released from SNP after hydrolysis with the three purified enzymes MaCel45A (A), TaCel5A (B) and AfCel12A (C). Reaction mixtures contained 1% DM and an enzyme concentration of 0,5mg/g DM in 50°C in 20 mM Na-citrate pH5. The reaction was incubated at 50°C in a thermo mixer with 1000 rpm shaking. The amount of release glucose and glucose oligomers (soluble) were analyzed over 5 time points and quantified with HPLC (according to Section 3.10.2.2). Each reaction was carried out in triplicates (error bars represent standard deviation*

A closer look at product profile in Figure 20 after hydrolysis with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A, revealed (rather surprisingly) high yields for SNP. It was hypothesized that the high content of lignin would create a barrier around the cellulose, reducing the total yield (as observed in section 4.2, production of monosaccharides. This was, however, not observed when hydrolyzing with endoglucanases. *Ma*Cel45A solubilized 3.4% of the theoretical maximum cellulose, *Ta*Cel5A solubilized 13.5%, and *Af*Cel12A solubilized 7.6% of the theoretical maximum cellulose after 24-hour hydrolysis. Interestingly, TaCel5A had "low" yields compared to *Ma*Cel45A and *Af*Cel12A on the substrates SOP (Section 4.4.1.2) and CCB (Section 4.4.1.3). When hydrolyzing SNP, on the other hand, it resulted in the highest yield observed in these time course experiments.

Higher yields was also observed for *Af*Cel12A when hydrolyzing SNP (solubilized 7.6% of the theoretical maximum cellulose) compared to hydrolyzing CCB (3.7% solubilized, Section 4.4.1.3) and SOP (3% solubilized, Section 4.4.1.2)

4.5 Analysis of solid residues

In the time course experiments described in section 4.4, I measured the amounts of soluble cellodextrines released after hydrolysis with endoglucanases. In the following experiments I wanted to study how the endoglucanases affects the insoluble cellulose.

As an example if endoglucanases introduce random cuts in a cellulose chain with DP 200 (example: untreated Avicel in Figure 21), it could be cut into shorter fragments with DP>7. None of these are water-soluble, however, during the course of the hydrolysis the DP 20 and 30 fragments are degraded into shorter (DP<6) oligomers. In the time course experiments only the latter was observe, but that does not tell us much about the insoluble residue. For that reason, measurements of the DP of the solids residues were performed to observe the change in the insoluble material. Notably, this is an experimental challenging procedure, with expected inaccuracies.

4.5.1. Homogenizing biomass

After hydrolysis with endoglucanases, the biomass remains as fiber bundles in the solution. This causes problems in the phenol-sulfuric acid and BCA assays (further explained in section 3.11). High standard deviations were observed in initial tests because of pipetting issues. When pipetting solutions with fiber bundles, it was difficult to get representative samples as fibers of different sizes were withdrawn. To solve this problem, an IKA T18 Basic homogenizer was used to homogenize the remaining substrates (i.e. after hydrolysis). First, a test was done to determine if the homogenizer affects the degree of polymerization.

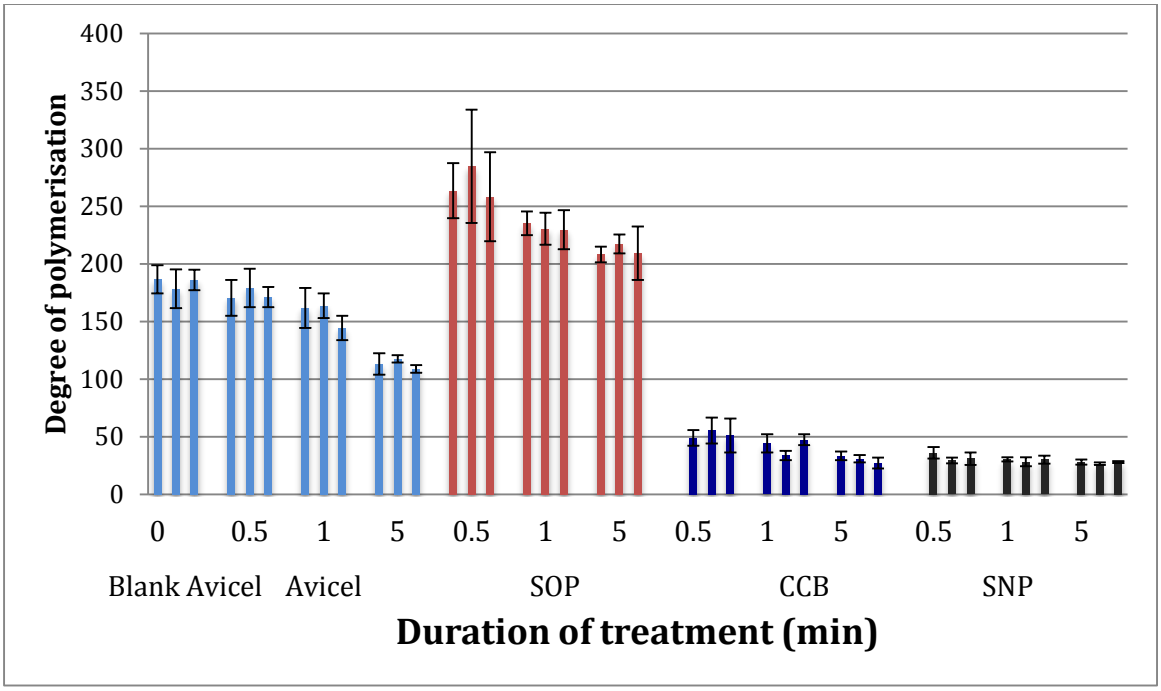


Figure 21. Effect of T18 basic homogenizer on the Degree of polymerization*. 5 ml reactions mixtures containing 1% DM of Avicel, SOP, CCB and SNP resectivly, were homogenized for 0.5, 1 and 5 minutes on 10 000 1/min rpm. No blanks are used for SOP, CCB and SNP, as they had high standard deviations. For this experiment, it was not necessary to remove proteins as described in section 3.11.2. Most of the protein from the biomass was removed during the production of paper, and no enzymes were used in this procedure.*

SOP, CCB, SNP (pretreated for 210°C for 14 min) and Avicel were homogenized for 0.5, 1 and 5 minutes (Figure 21) according to Section 3.11.1. Results show a reduction in DP, and the duration of the treatment seemed to be an important factor for how much the homogenizer affected the DP. In Avicel, a relative small reduction in the DP was observed from the untreated to 0.5 and 1 minute samples. After 5 minutes, however, a drastic reduction in DP was observed. Avicel was easy to handle, even without treatment and that was reflected in the standard deviations remained unchanged between the time points. SOP, CCB and SNP, on the other hand, show decreased standard deviations from 0.5 to 1 and 5 minutes. This was confirmed visually as there still were fibrous bundles of biomass left after 0.5 minutes of homogenizing. After 1 minute, the bundles were gone. From these observations, it was concluded that 1- minute homogenization was the best option. This procedure gave low standard deviations and the DP was not affected as severely as after 5 minutes treatment. After 1 minute, the DP (determined by the method described in section 3.11.3 and 3.11.4) was at an average of 230 for SOP, 40 for CCB, 30 for SNP and 160 in Avicel (Figure 21).

4.5.2. Time course conversion of wastepaper with MaCel45A

As endoglucanases introduce new cuts on the cellulose chain, they decrease the degree of polymerization. To observe this effect, *Ma*Cel45A was used to hydrolyze SOP, CCB and SNP over 24 hours and the DP was measured in the insoluble residues.

*Ma*Cel45A had the best performance between the three enzymes in the time course experiments (Section 4.4.1). As a result, it was used in the following experiments.

It was hypothesized that endoglucanases could introduce enough cuts to solubilize all amorphous cellulose after 24 hours, resulting in only crystalline cellulose remaining. This would result in a low DP because there are few reducing ends vs. the total cellulose content (measured according to Section 3.11). However, it was considered possible that the endoglucanases generate short insoluble oligosaccharides, especially in the initial phase of the reaction (30 min), which still were attached to the insoluble cellulose; these reducing ends would be measured in the BCA method (Section 3.11.4) and thereby reduce the DP (increased number of reducing ends compared to the total cellulose content). To confirm this thought, a 30 min test were run.

In the assay determining DP (Section 3.11) the enzymes was removed after hydrolysis and all the soluble material was washed away. Only the reducing ends of insoluble biomass were measured.

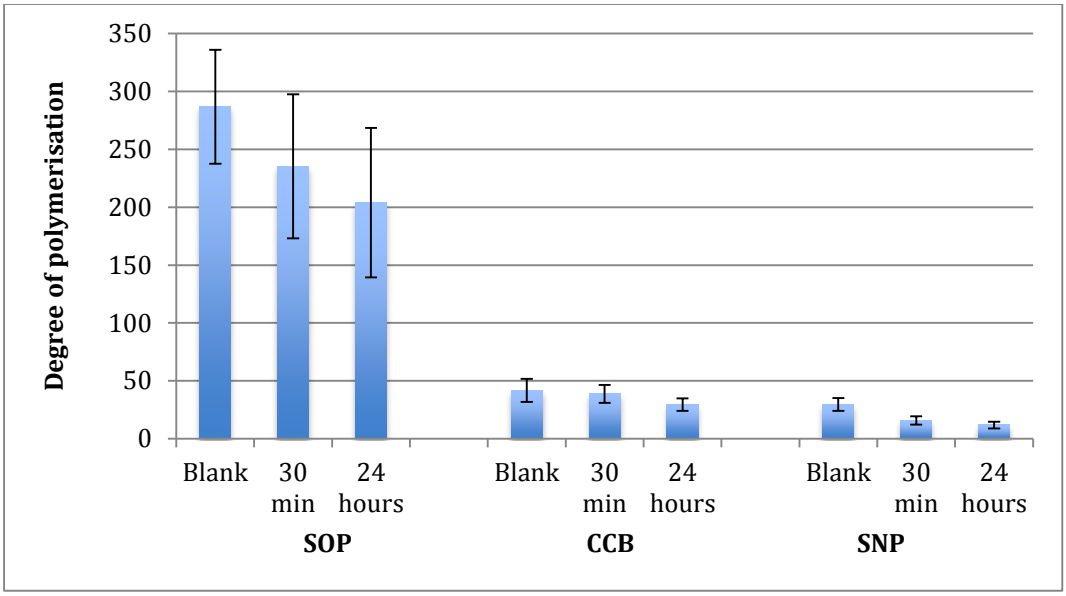


Figure 22. Effect of MaCel45A on DP for SOP, CCB and SNP*. 5 ml reaction mixtures contained 1% DM and an enzyme concentration of 0,5mg/g DM were incubated in a oven at 50°C in 20 mM Na-citrate pH5 buffer. After hydrolysis the mixture was homogenized for 1 min and the degree of polymerization was measured according to section 3.11. The figure illustrates how the action of MaCel45A effectively reduces the degree of polymerization in the substrates by introducing new chain ends.*

Even though the standard deviations were high, we can observe a trend: the endoglucanases reduce the DP of the substrates by introducing insertions in the amorphous regions of cellulose (as described in Section 1.3.1). Figure 22 illustrates that the DP was lower after 24 hours, compared to 30 minutes. This was not as hypothesized, and it will be discussed is section 5.

SOP was homogenized with and other dispersing element on the homogenizer (205mm instead of 108mm) between the experiments in Section 4.5.1 and Section 4.5.2. This could explain why the DP was slightly different between the experiments. In figure 21 SOP had a DP of 230 after homogenization. In Figure 22 SOP had a DP of 280 after homogenization.

4.5.3. Different concentrations of MaCel45A hydrolyzing Avicel

4.5.3.1. Insoluble sample

To examine whether changing enzyme concentrations have an effect on degree of polymerization of the insoluble material, the following test was performed.

This test builds on the hypothesis made in the previous experiment (Section 4.5.2 where 30min analysis was performed), where it was hypothesized that less activity of the endoglucanases would introduce cuts on the cellulose chain but not to such a degree that it would solubilize all the available cellulose. Instead the low activity would leave shorter (DP>6) insoluble chains that could be measured in the assay determining the degree of polymerization (see Section 3.11 for method description for DP measurements).

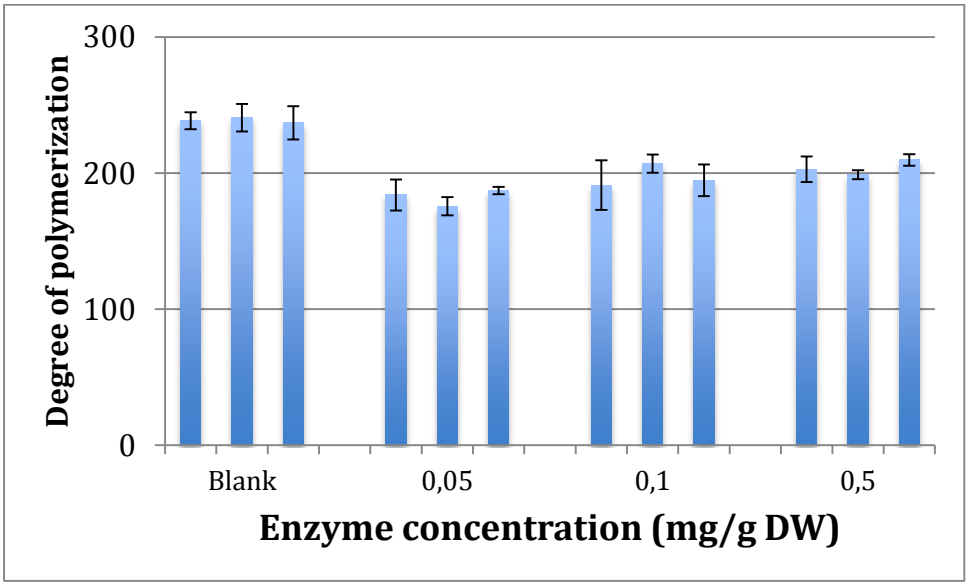


Figure 23. Degree of polymerization on insouble Avicel after hydrolysis with MaCel45A. *5 ml reaction mixtures contained 5% DM incubated in an oven at 50°C in 20 mM Na-citrate pH5 buffer. Avicel was hydrolyzed with MaCel45A concentrations of 0,05, 0,1 and 0,5 mg/g DM over 24 hours. The different MaCel45A loadings have different impacts on the degree of polymerization. The experiment was performed to se if lower enzyme loadings resulted in reduced degree of polymerization compared to higher enzymeloadings. Each reaction was carried out in triplicates (error bars represent standard deviation). Measured according to Section 3.11*

In Figure 23 we see that all MaCel45A loadings reduce the degree of polymerization. MaCel45A introduces cuts on the amorphous parts of cellulose, which shortens the

hydrolyzed chains (more reducing ends), thereby reducing the DP. The high enzyme loading were to "effective" when hydrolyzing, resulting in more soluble cellodextrines (see Figure 24) When the enzyme concentration was reduced, less soluble cellodextrines were produced and the chains still remain on the insoluble material with a DP>6; in other words, the low enzyme concentration makes cuts on the cellulose chain, but it was not effective enough over 24 hours to soluble all the available chains, thereby reducing the degree of polymerization. DP was measured according to Section 3.11. This fits the initial hypothesis and will be discussed in Section 5.

4.5.3.2. Soluble sample

Soluble cellodextrins were analyzed in order to observe if decreased enzyme concentrations results in less soluble carbohydrates. It was also observed if lower enzyme concentrations yielded longer cellodextrins (DP 4,5 and 6).

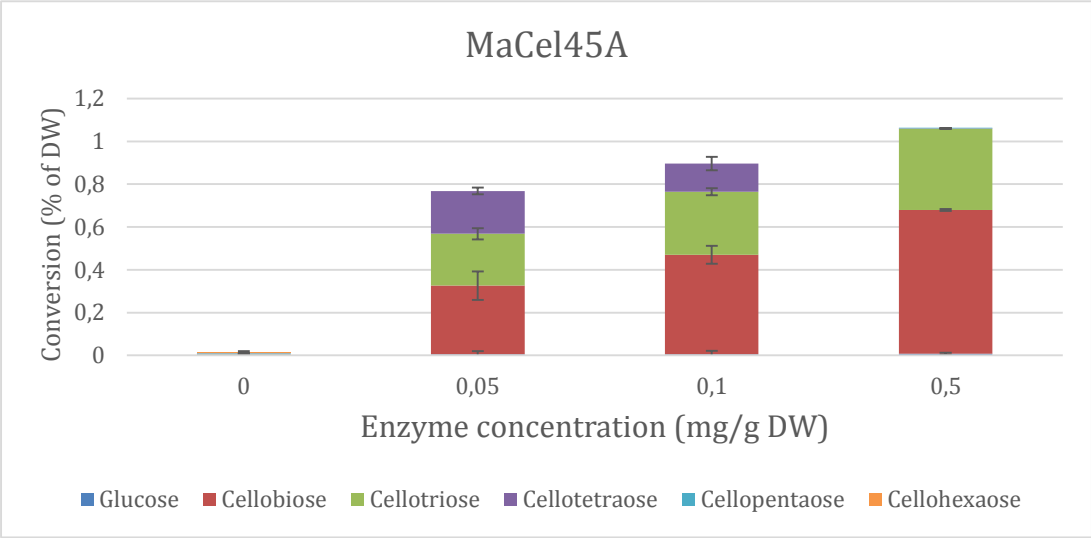


Figure 24. Soluble cellodextrins released from enzymatic hydrolysis of Avicel with MaCel45A*. 5 ml reaction mixtures contained 5% DM incubated in an oven at 50°C in 20 mM Na-citrate pH5 buffer. Avicel was hydrolyzed with MaCel45A concentrations of 0,05, 0,1 and 0,5 mg/g DM over 24 hours. Each reaction was carried out in triplicates (error bars represent standard deviation) Hydrolysis yield were analyzed with HPLC.*

Observation of the soluble products in Figure 24 revealed how the endoglucanases were “effective” even at low concentrations. MaCel45A concentration of 0.05 mg/g DM solubilized 0.78% of the theoretical maximum cellulose in Avicel after 24 hours. MaCel45A concentrations of 0.1 mg/g DM solubilized 0.89%, and MaCel45A concentrations of 0.5 mg/g DM solubilized 1.08% of the theoretical maximum cellulose. Notably, as the enzyme concentration decreases, it was observed increasing values of cellotetraose. The cellulose that did not solubilized at the low enzyme loading (0.05 mg/g DM, Figure 24) remained as insoluble cellulose chains and they were the reason why we observed a decreasing DP at low enzyme concentrations in Figure 23. DP was measured according to Section 3.11.

5 Discussion

The objective of this study was to hydrolyse SOP, CCB and SNP with a pre-produced enzyme mixture (Cellic CTec2) and three purified endoglucanases (MaCel45A, TaCel5A and AfCel12A). Product formations from both the Cellic CTec2 and the three endoglucanases (*Ma*Cel45A, *Ta*Cel5A and *Af*Cel112A) were evaluated. It was also studied which affect the endoglucanases had on the DP of the remaining insoluble Avicel, SOP, CCB and SNP after hydrolysis. In order to correctly evaluate the yields after the hydrolysis, it was necessary to obtain the composition of the starting biomass.

Compositional analysis

It was important to know the composition of carbohydrates and lignin in biomasses, as they can give a clear indication towards enzymatic conversion, and what yields to expect. High amounts of lignin in the substrates were expected to result in low conversion after enzyme hydrolysis because cellulose is protected by lignin (as described In Section 1.1.3). Compositional analysis was useful because it could identify these compounds and explain results post hydrolysis.

As described in Section 1.5.2, higher temperatures during steam explosion result in partially hemicellulose degradation (Alvira, Tomás-Pejó et al. 2010). This was observed in Table 7, showing that hemicelluloses sugars (xylose, mannose, arabinose and galactose) decomposed, as the pretreatment got more intensive.

The quality of paper (described in Section 1.5.3) depends in part on the ratio between lignin and hemicellulose. Low lignin and hemicellulose fractions are essential for high quality office paper, and this is achieved by pulping procedures designed to remove these impurities.

As described in Section 3.2, ash is inorganic material, which do not incinerate at high temperatures. “Others” was the mass loss in the procedure and unidentified compounds, but it could also indicate errors in the methods. In the pulping process used to produce SOP, calcium carbonate is added as cheap filler and pigment. This

additive composes about 20%(w/w) of the total mass. This material appears under others and ash in Figure 10. When calcium carbonate is heated up to 550˚C it turns into $CaO_{(S)}$ and $CO_{2(g)}$. The ratios of these compounds are 56%(w/w) to 44%(w/w) respectively. $CaO_{(S)}$ was registered together with all non-organic compounds in the total ash after incineration of the sample (described in Section 3.2). When calcium carbonate reacts with acids, it is dissolved and $CO_{2(g)}$ is released; $CO_{2(g)}$ was registered as mass loss in the analysis (others).

Calcium carbonate comprise about 15-20%(w/w) (see Section 1.6) of the total weight in SOP, and 56%(w/w) of this was in the ash (Equation 1, $CaO_{(S)}$) and 44% was registered in others as loss (Equation 2, CO_2 released). Calculations below estimate how much calcium carbonate represents of ash and others (Table 5). The average ash in the pretreated SOP substrate (calculated from the values in Table 5) was 210g in 1000 g dry mass. If we assume that 20%(w/w) of the total mass was calcium carbonate; Calculations estimate that Cao composes about 53%(w/w) of the total ash.

Equation 1 $$\left(\frac{\left(\left(1000g\left(\frac{20\%(w/w)}{100}\right)\right)*\left(\frac{56\%(w/w)}{100}\right)\right)}{average\ ash\ 210g}\right)*100 = 53{,}3\%$$

The same calculation was performed for CO_2 and it was estimated that it composes about 30%(w/w) of "Others"

Equation 2 $$\left(\frac{\left(\left(1000g\left(\frac{20\%(w/w)}{100}\right)\right)*\left(\frac{44v(w/w)}{100}\right)\right)}{average\ "others"\ 297g}\right)*100 = 29.6\%$$

The high values of "Others" in SOP were partly explained by the additive calcium carbonate. The remaining 70% are either non-quantified material (mentioned in Section 1.1), errors in the procedure or mass loss during the assay. For SNP and SOP, Ink printed on the paper could be the reason for the high values of "others". SNP do not contain any identified additives such as calcium carbonate and it is unlikely that ink represents all others; the high values are probably a result of errors in the procedure. From the general low values of unidentified compounds in CCB, we could conclude that there was little to no additives in the production of cardboards.

Production of monosaccharides

In the process of enzymatically converting waste biomass to monomeric carbohydrates, it was observed how critical pH- control was for optimal yields. When converting waste biomass. It was important to identify the ideal acid loadings to achieve the best conditions for the enzymes throughout the incubation. pH shifts (initial pH was different for final pH) was observed in all substrates (SOP, CCB and SNP), and this could affect the conversion yield. Optimal conditions were achieved when the pH shift was close to pH 5 (from pH 4.5 to 5.5) because this is the optimal pH for hydrolysis with Cellic CTec2.

SOP contains calcium carbonate as a pigment and the presence of this alkaline substance required considerable amounts of acid to reach optimal conditions (pH5)(Table 8). When loading citric acid to SOP, a reaction with calcium carbonate was observed. The reaction solubilized calcium carbonate and released CO_2. CCB and SNP do not contain the same additives affecting pH. The low acid loading required confirmed that there was no calcium carbonate added in the production.

This has major economical advantages because the amount of acid required were significantly lower for CCB and SNP compared to SOP. Even though SOP result in higher yields, it may not be the best choice from an economical point of view. Acid is very expensive and the corrosive material can damage reactors in a large-scale production. CCB was probably the best option of the three substrates as it required low amounts of acid and contained high amounts of polysaccharides, and relative small amounts of lignin, compared to SNP.

T. W. Lan and colleagues (Lan, Lou et al. 2013), claim that enzymatic saccharification of lignocellulose should be conducted at pH around 5.2-6.2 when using Cellic CTec2. This was in line with the results for SOP and CCB in the pH control experiments. SNP, on the other hand, seems to have optimal conditions under pH 5.2 in the pH control experiment and this was confirmed by the following experiment in Section 4.2.2, where the yields were lower when the final was pH 5.1 compared to pH4.9. The yield of SNP was actually quite high in the pH control experiment with conversions of 50% of the theoretical yield (pH4.5), but the yields were reduced in the following

experiment with 24 hours hydrolysis and optimal acid loadings, but notice how the pH was closer to 5.

The 24-hour hydrolysis at optimal conditions show a clear increase in conversion as pretreatments gets more intensive, confirming the effect of steam explosion on reducing the resilience in biomass. The more intensive treatment distorts the lignocellulose-structure and makes the carbohydrates more accessible as explained in Section 1.5.2. Furthermore, the 24 hour hydrolysis with Cellic CTec 2, shows a connection between lignin content and total conversion as commonly observed (Richard 1996, Jørgensen, Kristensen et al. 2007). SOP pretreated at 210°C for 14min, contain small amounts of lignin resulting in a conversion of 54% of theoretical maximum conversion in the 24 hour hydrolysis with optimized pH control, CCB had a conversion of 46.6% and SNP 29.1% when the substrates were pretreated at 210°C for 14min.

SOP pretreated at 170°C 7min had a higher conversion than SOP pretreated for 170°C 14min when hydrolyzed with Cellic CTec2 in the 24 hour hydrolysis experiment with optimized pH control. This was not expected because a more severe pretreatment should make the carbohydrates more accessible. These unexpected results could be explained by the pH, which shifts from 4.6 to 5.6 (over the optimal pH for Cellic CTec2 (pH5)) in SOP pretreated at 170°C for 7min, while SOP pretreated for 170°C 14min shifts from 4.4 (initial pH) to 4.8 in the final pH. This was below the optimal pH conditions for Cellic CTec2 and could explain why the yield was lower.

Steam explosion seems to have the best effect on CCB in the 24-hour hydrolysis with optimized pH control (Section 4.2.2, Table7), which increased the conversion with 10% from the least pretreated substrate (170°C 7min) to the most severely pretreated substrate (210°C 14min). SOP increased with 8.7% and SNP increased yields with 5.5%.

It was expected that lignin-rich substrates such as SNP, would have a low yield compared to SOP containing minimal amounts of lignin. The 24-hour hydrolysis with optimized pH control shows a clear correlation between the amount of lignin in the biomass and the conversion yields. Untreated and pretreated SOP contains low amounts of lignin (about 2%, Table 5), SNP, on the other hand, contains up to 25%

lignin and CCB contains 12% in average for untreated and pretreated samples. Corresponding to this, the yields for SNP were 50% lower than the yields for SOP. The yields should be somewhat higher for SNP if we compare with the results in the pH control experiment where reactions with SNP gave yields close to 50%. pH was the main difference between results in the pH control experiment and the 24 hour hydrolysis with optimized pH control (pH 4.5 in Table 6, pH4.9 in Table 7).

Tests with Cellic CTec2 concentrations 5 mg/g DM did not result in conversions exceeding 60%. Inhibitors or the fact that Cellobiohydrolases might get "stuck" on the cellulose chain (explained in Section 1.3.2) could explain this. We cannot state that the enzymatic hydrolysis has reached maximum yields with 5mg/g enzyme concentration after 72 hours (Figure 14). So it might be that the Cellic CTec2 are slowly making progress, and the conversion might be higher with increased incubation time.

Longer incubation, 72-hours with an enzyme concentration of 5mg/g DM seems result in about the same conversion (SOP 65% of maximum theoretical yield) as increasing the Cellic CTec2 concentration from 5mg/g DM to 25mg/g DM over 24 hours (SOP 68% of maximum theoretical yield). The same was observed with CCB and SNP. This would confirm the hypothesis that the enzymes at 5mg/g DM concentration had insufficient time to hydrolyze the substrate in 24 hours. When the enzyme concentration was increased to 25mg/g DM, they were able to hydrolyze the substrates faster, so they managed to solubilize the same amount of sugars in 24 hours as 5mg/g DM Cellic CTec2 managed to hydrolyze in 72 hours. When the Cellic CTec2 concentration was increased to 25mg/g DM and incubated for 72-hours, there was a drastic increase in total yield, illustrating that there are available carbohydrates, and maximum yield was not accomplished with 5mg/g DM Cellic Ctec 2 over 72 hours or with increased enzyme concentration (25mg/g DM) over 24 hours. It may be possible that inhibitors in the solution render the enzymes incapable of passing 70% conversion with the low Cellic CTec2 loading (5mg/g DM) (but it might also be that 72 hours was insufficient time)

As for the high Cellic CTec2 concentration (25mg/g), 24 hours was not enough time for the enzymes to reach maximum yields, even 72 hours might be too limit to reach maximal yields. (Cannella, Hsieh et al. 2012) measured Cellic-CTec2 conversion

on wheat straw and increasing yields were observed even after 140 hours (90% cellulose conversion). So it could be expected to observe yields over 90% in conversion of office paper with high Cellic CTec2 concentrations (25mg/g) over longer incubation durations.

Production of recombinant endoglucanases

Initial expression of *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A, in *Pichia pastoris* resulted in low yields. As explained in Section 3.4.3, yeast expression systems (such as *Pichia pastoris* with the control of GAP promotor) show high growth rates and a good production of enzymes on inexpensive media (Várnai, Tang et al. 2014).

High enzyme yields are normally obtained by utilizing shake flask (Várnai, Tang et al. 2014). In this study, a Harbinger system provided oxygen for the yeast. The oxygenation in the system was very efficient. This resulted in a rapid growth of cells, but, even though a lot of cells were observed, only small amounts of proteins were measured after purification with IEC. Longer incubation and added dextrose did not increase the yield. The results could indicate that the rapid growth of *P. pastoris* hampered the production of the desired endoglucanases. The yields were too small for the following experiments so a pre-produced batch containing *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A (further described in Section 4.3) was utilized because of time limitations.

*Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A required different degrees of purification. The broths with *Ma*Cel45A and *Ta*Cel5A contained an other protein at 70kDa, which was eluted with the endoglucanases after IEC. These proteins had similar affinity as *Ma*Cel45A and *Ta*Cel5A for the column material in IEC. The mass difference between the 70kDa protein and the endoglucanases would indicate an easy separation on SEC. Indeed, *Ma*Cel45A was completely separated from the 70kDa protein after SEC, but *Ta*Cel5A appeared not to be completely purified (figure 16 D and Figure D.4 in the Appendix). This was not expected as the proteins differ highly in molecular mass. The next run on SEC for the *Ta*Cel5A (Figure D5 in the Appendix) resulted in a good separation. The conditions were the same in the second run as for the first purification on SEC, so the unsuccessful purification of *Ta*Cel5A in the first run has to

be problems connected to the column. I could be that the column was not sufficiently equilibrated with equilibration buffer before enzyme loading.

After SEC separation of *Ma*Cel45A, a small contaminating protein with about the same mass as *Ma*Cel45A was detected (Figure 16B). This small band could be a result of unprocessed *Ma*Cel45A with a signal peptide attached, but it could also be a protein, difficult to separate from *Ma*Cel45A. Because of its low concentration this protein was considered negligible.

*Af*Cel12A was completely purified after IEC (Figure 16E) and no further purification was necessary.

Production of oligosaccharides

In the time course experiment with Avicel (Figure 16) showed that the yields of longer cellodextrines (DP 4, 5 and 6) reach their maximum levels after about 1hour for all three endoglucanases. After about one hour, the levels of these longer products remained about the same until the 6 hours sampling point. Similar phases of quasi-equilibrium were observed in other, similar reaction set-ups, e.g. for the reaction of *Af*Cel12A with CCB (Figure. 18; stable phase between 2 to 6 hours) After 24 hours incubation the reaction mixtures obtained with Avicel (Figure 16) did not contain detectable amounts of DP 5 or 6 in the solution. One could assume a state of quasi-equilibrium during the first hours of the reaction, meaning that the production and further hydrolysis of longer soluble cellodextrines happens at approximately the same rate. This would explain why we observed unchanged concentrations of longer cellodextrines (DP 4, 5 and 6) during the first hours. However, when endoglucanases struggle to find available cellulose on the insoluble material, the quasi-equilibrium ceases. No more soluble cellodextrines were released, but endoglucanases in the solution continue the hydrolysis of cellotetraoses, cellopentaoses and cellohexaoses. After 24 hours we see an increase in smaller molecules such as cellobiose and cellotriose as a result. It would have been interesting to include a 72-hour measurement in the time assay to analyze the endproducts – one would expect hydrolysis of cellotetraoses down to cellobiose for *Ma*cel45A in Avicel. Observations in the time course experiments indicate that endoglucanases are either unable to

hydrolyse triose further down to glucose and cellobiose or that they do so at a very slow rate. This could be confirmed with a late time point.

Even though *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A resulted in low yields with Avicel compared to hydrolyzing SOP, CCB and SNP, it had fast conversion rate and seemed to reach maximum yield (for the reaction setup) within the first 6 hours (small changes in yield from 6 to 24 hours). It would have been interesting to try and experiment with adding more enzymes after 10 hours, and see if the yields increase. If this would be the case, it would indicate that the enzyme in the initial loading was inactive. If no additional solubilized cellodextrines were observed, they might have reached maximum conversion or the enzymes might be hindered by inhibitors such as glucose and cellobiose which would inactivate or reduce the efficient of the enzymes(Zeng 2007). Another experiment would be to add more substrate and see if the yields increase. If the yields increased it would indicate that the endoglucanases hydrolysed all accessible cellulose in the original substrate loading (indicates maximum yields from endoglucanases when hydrolyzing Avicel). If no additional solubilized sugars were observed, it would indicate that the enzymes are inactive. Avicel has a high degree of crystalline regions, reducing the total yield of the hydrolysis. Treatments, such as steam explosion (explained in Section 1.5.2), could create more amorphous region, thus increasing the yield.

SOP contains high amounts of cellulose without the protective cover of lignin, which were removed under the production of paper. As a result, hydrolysis of SOP was expected to result in high yields compared to CCB and SNP, in the time course experiment. Observations from Figures 18, 19 and 20 did not support these expectations, because similar yields were observed on all three substrates when hydrolyzed with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A. There was no pH control (use of citric acid to control the pH) in these tests so the yield in SOP might be have been higher if optimal pH was set and controlled in the hydrolysis with *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A. 20mM Na-citrate pH5 were used in the reaction, but the alkaline properties of calcium carbonate could elevate the pH.

SNP resulted in the highest yield in the experiment when hydrolyzed by *Ta*Cel5A. This could potentially mean that hydrolysis with endoglucanases was not hindered by lignin. In the enzyme hydrolysis with Cellic CTec2 the biomasses were

less accessible for processive cellulases, because lignin creates a protective barrier around the cellulose. Endoglucanases might manage to cut at available spots around the lignin, not covering cellulose, resulting in higher yields in SNP as it might have more accessible cellulose microfibrils (less crystalline regions). SOP contains long cellulose chains as described in section 1.6; in these cellulose chains there might be a considerable amount of crystalline regions, reducing yields and increasing the need for cellobiohydrolases to obtain high yields. However, lignin and hemicellulose could have an impact on the effectiveness of the enzymes. In the biomasses SOP, CCB and SNP (Section 4.4.1), drastic increases in the yields were observed between 6 and 24 hours, meaning that there was a lot of available carbohydrates available after 6 hours. In Avicel (no lignin and hemicellulose) however, it was concluded that the reaction was close to maximum yield after 6 hours.

As described earlier, a quasi-equilibrium could develop when endoglucanases hydrolyse insoluble cellulose and release cellodextrines (DP2-6). However, because *Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A hydrolyze the substrates (SOP, CCB and SNP) so inefficiently, this quasi-equilibrium were not achieved, because the hydrolysis of longer soluble cellodextrines happens at a faster rate than the endoglucanases release new cellodextrines from the insoluble material. As a result of this, low amounts of longer cellodextrines (4, 5 and 6) were observed. One might speculate that use of CBM-containing endoglucanases could increase the endoglucanases concentration on the insoluble cellulose (explained in Section 1.6.4), and less endoglucanases would be necessary to efficiently hydrolyze SOP, CCB and SNP. This could potentially result in additional yields of longer cellodextrines (DP 4, 5, and 6).
Figures 17 and 18 show that *Ma*Cel45A releases more cellotetraose from Avicel and SOP, than *Ta*Cel5A and *Af*Cel12A. There are several possible explanations for these results. It could be that *Ma*Cel45A effectively concentrates it self around the insoluble material resulting in increased yields of longer cellodextrines, or that it has a low activity in the solution and difficulties hydrolyzing the soluble oligomers.

Results from the time course experiment show that *Ma*Cel45A performed best when hydrolyzing SOP and CCB; *Ta*Cel5A performed best on SNP. There could be several answers to this, but one can speculate that *Ma*Cel45A has the best conversion on SOP because it tolerates calcium carbonate better than the *Ta*Cel5A

and *Af*Cel12A. Furthermore, CCB contains the most hemicellulose of the three substrates and some endoglucanases have the ability to hydrolyse hemicellulose (Hasper, Dekkers et al. 2002). It might be that *Ma*Cel45A was capable of hydrolyzing the hemicellulose in CCB thereby creating better access to the cellulose and increasing the yield. *Ma*Cel45A had very low yields on SNP, and this could be a result of the lignin fractions inhibiting the enzyme from hydrolyzing the substrate. Jenni Liisa Rahikainen and colleagues (Rahikainen, Evans et al. 2013) discovered that *Ma*Cel45A could interact with negatively charged groups on lignin which contribute to non-productive enzyme adsorption. The lignin residues might not affect *Ta*Cel5A in the same way, and this could explain the high yields. The time course experiment was performed on SOP, CCB and SNP pretreated at 210°C for 14min. If Avicel were pretreated at the same conditions, increased yields would be expected because steam explosion defibrillates cellulose as a result of the intense conditions in the procedure. The pretreatment would generate additional amorphous regions, which endoglucanases could hydrolyze.

Effect of MaCel45A, Tacel5A and AfCel12A on insoluble residues

The low DP in newspaper and CCB were a result of the treatment under production (described in Section 1.6). They are mechanically handled and rough treatments lowers the DP (also observed in Figure 21 with the T-18 basic homogenizer). Chemical pulping of SOP preserves the long fibers in softwoods required for high quality paper. The difference between the DP in SOP, CCB and SNP are visualized in Section 4.5.2 (Figure 21).

From the time course experiments it was decided to use MaCel45A when studying the effect of endoglucanases on insoluble fractions of cellulose. *Ma*Cel45A was chosen because it had good yields overall with each substrate. The three EGs vary significantly depending on which substrate they hydrolyse.

In Section 4.5.3, *Ma*Cel45A hydrolyze Avicel at different enzyme concentrations. It was initially predicted that lower enzyme loading would create less soluble cellodextrines, the hypothesis was that low concentrations of endoglucanases introduce less cuts leaving more insoluble chains still attached to the amorphous regions compared to high enzyme concentrations which would be more effective

and solubilize more cellodextrines. As a result it could be measure additional reducing ends (at the low enzyme concentration) resulting in a reduced DP.

The results (Figure 23) show that reduced amounts of *Ma*Cel45A (0.05mg/g DM) were less capable of solubilizing cellulose in amorphous regions (Figure 24, solubilized 0.78% of DM) compared to high enzyme loadings (0.5mg/g DM), which solubilized 1.1% of DM. These results were put in context with the results in Figure 23, where we observe a low DP (180) at the reduced enzyme concentration (0.05mg/g DM). Whereas we observed a high DP (200) for the experiment with high enzyme concentrations (0.5 mg/g DM.

In theory, we should observe the same effect in the 24-hour time course conversion of wastepaper in Section 4.5.2, with the 30 minutes sample (Figure 22). After 30 minutes we would expect that the endoglucanases would create cuts in the cellulose chain, without solubilizing many cellodextrines, thereby lowering the DP of the insoluble material compared to the 24-hour sample. It was expected that endoglucanases would solubilize all available cellulose in amorphous regions within 24 hours, leaving only crystalline areas, which resulted in an increased DP of the insoluble material. The results, however, showed that the DP was higher after 30 minutes compared to 24 hours. One can conclude that the 30 minutes experiment was too short for the endoglucanases to create enough insertion to lower the DP. In the time course experiments of SOP, CCB and SNP (Figure 22), we see a general low activity (release of soluble cellodextrines) during the first 30 minutes. In post perspective, a 2 hours time point would be interesting and we might observe a decrease in DP compared to 24 hours.

Figure 23 illustrates the soluble products. Even at a low MaCel45A concentration (0.05mg/g DM) it was observed relative high values of soluble Cellodextrines. If these cellodextrines were not solubilized, but rather remained insoluble (DP>6), we would observe a decreased DP. If the enzyme concentration was reduced even more, we might observe less soluble products and a lower DP.

5.1 Concluding remarks and future work

This thesis focused on the conversion of pretreated waste biomass to monosugars with a pre-produced enzymatic cocktail (Cellic CTec2) and produceing cellooligosaccharides with three recombinantly expressed endoglucanases (*Ma*Cel45A , *Ta*Cel5A and *Af*Cel12A).

In this thesis the compositional analysis of untreated and pretreated SOP, CCB and SNP was performed. The results from this assay were used to accurately determine the total yields in the subsequent enzyme hydrolysis experiments.

The first goal in this thesis was to optimize the yields when converting SOP, CCB and SNP to sugar monomers with Cellic CTec2. SOP was a difficult substrate because of its alkaline properties. Experiments with varied acid loading were therefore performed for all substrates, in order to set the pH of the reaction to according to the preference of the enzymes. To maximize the yields, increased enzyme loadings and incubation times were performed. SOP was solubilized up to 89.6% of the theoretical maximum sugars, CCB was solubilized up to 55.2% and SNP was solubilized up to 56.3%. Based on these data, SOP was the most accessible substrate of the three to Cellic CTec2. By comparing the total yields to the compositional data of the substrates, it was clear that lignin and hemicellulose content affected the total yield.

Furthermore three endoglucanases (*Ma*Cel45A, *Ta*Cel5A and *Af*Cel12A) were successfully purified from premade broths. The endoglucanases were produced in *P. pastoris* and the purifications were performed in two steps with IEC followed by SEC.

In experiments with endoglucanases I wanted to observe two things. First, the release of soluble cellodextrines when hydrolyzing Avicel, SOP, CCB and SNP, secondly the effect of endoglucanases on insoluble cellulose chains to better understand how the endoglucanases work. Time course experiments were performed in order to observe released soluble cellodextrines at different time points, and the results showed a release of all cellodextrines ranging from DP 2 to 6 during the course of the experiment. Longer cellodextrines (DP 4-6) were observed in the initial phase of the hydrolysis (30 min to 6 hours).

Cellopentaoses and cellohexaoses were further degraded to shorter chains after 24 hours.

When analyzing the insoluble material, results demonstrate how the endoglucanases reduce the degree of polymerization by introducing cuts on the insoluble cellulose chains. It thus seem that the endoglucanases hydrolyze the accessible regions of cellulose chain but not to an extent that leads to a high degree of conversion of soluble material.

For future work, it would be interesting to study the distribution of length of the cellulose chains after endoglucanases hydrolyze the insoluble cellulose. It could be presume that the endoglucanases only work on the surface of the cellulose, so we would observe short chains on the cellulose surface and longer chains at the core of cellulose. If this is correct, we could fractionate the shorter insoluble chains (maybe ranging from DP 7 to 30) with organic solvents from the insoluble residues for appliances in pharmaceutical and food industries. The longer chains, which remain after fractionating, could be re-hydrolyzed with endoglucanases now that new accessible cellulose is exposed. Furthermore, It would also be interesting to determine why the endoglucanases hydrolyze differently on the individual substrates (SOP, CCB and SNP). This research can give information about how the endoglucanases are affected by additives in paper and improve our understanding of endoglucanases.

For the conversion of waste paper to monosaccharides it would be beneficial to study which of the two materials, SOP or CCB, would be more economical on a large scale. SOP requires high amounts of acid, but the hydrolysis with Cellic CTec2 solubilizes monosugars more efficiently. CCB, on the other hand, requires less acid but Cellic CTec2 solubilizes less monosugars.

6 References

Abbasi, T., et al. (2011). "The return to renewables: Will it help in global warming control?" Renewable and Sustainable Energy Reviews **15**(1): 891-894.

Agulia, J. L. D. (2007). Comparison of Catalytic Activities of Wild-type and Mutant Endoglucanases from Pyrococcus Horikoshii.

Albersheim, P., et al. (2010). Plant Cell Walls.

Alvira, P., et al. (2010). "Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review." Bioresource Technology **101**(13): 4851-4861.

Bastawde, K. B. (1992). "Xylan structure, microbial xylanases, and their mode of action." World Journal of Microbiology and Biotechnology **8**(4): 353-368.

Beckham, G. T., et al. (2014). "Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases." Current Opinion in Biotechnology **27**(0): 96-106.

Bilek, M. M. M., et al. (2011). "Free radical functionalization of surfaces to prevent adverse responses to biomedical devices." Proceedings of the National Academy of Sciences **108**(35): 14405-14410.

Boraston, A. B. (2004). "Carbohydrate-bining modules: fine-tuning polysaccharide recognition."

Brander, J. and I. Thorn (1997). "Surface Application of Paper chemicals." (First edition): 146.

Cannella, D., et al. (2012). "Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content." Biotechnology for Biofuels **5**(1): 26.

Chami Khazraji, A. and S. Robert (2013). "Self-Assembly and Intermolecular Forces When Cellulose and Water Interact Using Molecular Modeling." Journal of Nanomaterials **2013**: 12.

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

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

Decker, S., et al. (2009). "High-Throughput Screening Techniques for Biomass Conversion." BioEnergy Research **2**(4): 179-192.

Di Risio, S., et al. (2011). "Large-scale, high-solids enzymatic hydrolysis of steam-exploded poplar." Biofuels, Bioproducts and Biorefining **5**(6): 609-620.

Dodd, D. and I. K. O. Cann (2009). "Enzymatic deconstruction of xylan for biofuel production." Global change biology. Bioenergy **1**(1): 2-17.

DuBois, M., et al. (1956). "Colorimetric Method for Determination of Sugars and Related Substances." Analytical Chemistry **28**(3): 350-356.

Dunlop, A. P. (1948). "Furfural Formation and Behavior." Industrial & Engineering Chemistry **40**(2): 204-209.

Eijsink, V. G. H., et al. (2008). "Towards new enzymes for biofuels: lessons from chitinase research." Trends in Biotechnology **26**(5): 228-235.

Estella, R. (2013). "Hydrolysis of Biomass Mediated by Cellulases for the Production of Sugars."

Fengel, D. and G. Wegener (1983). Wood: Chemistry, Ultrastructure, reactions.

Fox, J. M., et al. (2012). "Initial- and Processive-Cut Products Reveal Cellobiohydrolase Rate Limitations and the Role of Companion Enzymes." Biochemistry **51**(1): 442-452.

Ganter, J. L. M. S., et al. (2001). "Real-time monitoring of enzymatic hydrolysis of galactomannans." Biopolymers **59**(4): 226-242.

Gao, D., et al. (2013). "Increased enzyme binding to substrate is not necessary for more efficient cellulose hydrolysis." Proceedings of the National Academy of Sciences of the United States of America **110**(27): 10922-10927.

Hasper, A. A., et al. (2002). "EglC, a new endoglucanase from Aspergillus niger with major activity towards xyloglucan." Applied and environmental microbiology **68**(4): 1556-1560.

Henrissat, B. (1991). "A classification of glycosyl hydrolases based on amino acid sequence similarities." Biochemical Journal **280**(Pt 2): 309-316.

Henrissat, B. and G. Davies (1997). "Structural and sequence-based classification of glycoside hydrolases." Curr Opin Struct Biol **7**(5): 637-644.

Hubbe, M. A. and L. A. Lucia (2007). The "love-hate relationship present in lignocellulosic material.

Hughes, E. (2000). "Biomass cofiring: economics, policy and opportunities." Biomass and Bioenergy **19**(6): 457-465.

Igarashi, K., et al. (2009). "High Speed Atomic Force Microscopy Visualizes Processive Movement of Trichoderma reesei Cellobiohydrolase I on Crystalline Cellulose." The Journal of Biological Chemistry **284**(52): 36186-36190.

Jacobs, A., et al. (2002). "Characterization of water-soluble hemicelluloses from spruce and aspen employing SEC/MALDI mass spectroscopy." Carbohydrate Research **337**(8): 711-717.

Jørgensen, H., et al. (2007). "Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities." Biofuels, Bioproducts and Biorefining **1**(2): 119-134.

Kang, Q., et al. (2014). "Bioethanol from Lignocellulosic Biomass: Current Findings Determine Research Priorities." The Scientific World Journal **2014**: 13.

Ketudat Cairns, J. and A. Esen (2010). "β-Glucosidases." Cellular and Molecular Life Sciences **67**(20): 3389-3405.

Kim, J.-H., et al. (2010). "Simultaneous consumption of pentose and hexose sugars: an optimal microbial phenotype for efficient fermentation of lignocellulosic biomass." Applied Microbiology and Biotechnology **88**(5): 1077-1085.

Kipper, K., et al. (2005). "Processive action of cellobiohydrolase Cel7A from Trichoderma reesei is revealed as 'burst' kinetics on fluorescent polymeric model substrates." Biochemical Journal **385**(Pt 2): 527-535.

Kolenová, K., et al. (2006). "Mode of action of endo-β-1,4-xylanases of families 10 and 11 on acidic xylooligosaccharides." Journal of Biotechnology **121**(3): 338-345.

Kostylev, M. and D. Wilson (2012). "Synergistic interactions in cellulose hydrolysis." Biofuels **3**(1): 61-70.

Lan, T. Q., et al. (2013). "Enzymatic Saccharification of Lignocelluloses Should be Conducted at Elevated pH 5.2–6.2." BioEnergy Research **6**(2): 476-485.

Li, J., et al. (2007). "Lignin depolymerization/repolymerization and its critical role for delignification of aspen wood by steam explosion." Bioresource Technology **98**(16): 3061-3068.

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

Manning, et al. (2000). A reader in preservation and conservation. page 65.

Martin, N., et al. (2000). Opportunities to improve energy efficiency and reduce greenhouse gas emissions in the U.S. pulp and paper industry.

McDonald, K., et al. (2008). "In-Gel Protein Quantification Using the Criterion Stain Free Gel imaging system."

Medie, F. M., et al. (2012). "Genome analyses highlight the different biological roles of cellulases." Nat Rev Micro **10**(3): 227-234.

Mehdi Dashtban, h. S., wensheng Qin (2009). "Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives Int J Biol Sci 2009; 5(6):578-595.".

Moreira, L. (2008). "An overview of mannan structure and mannan-degrading enzyme systems." Applied Microbiology and Biotechnology 79(2): 165-178.

Moreira, L. R. S. and E. X. F. Filho (2008). "An overview of mannan structure and mannan-degrading enzyme systems." Applied Microbiology and Biotechnology 79(2): 165-178.

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

Motta, F. L., et al. (2013). A Review of Xylanase Production by the Fermentation of Xylan: Classification, Characterization and Applications.

Novozymes (2015). "Application sheet Cellic CTec 2."

Oner, S. Y. a. B. (2014). "Mannan as a Promising Bioactive Material for Drug Nanocarrier Systems, Application of Nanotechnology in Drug Delivery, PhD. Ali Demir Sezer (Ed.), ISBN: 978-953-51-1628-8, InTech, DOI: 10.5772/58413. Available from: http://www.intechopen.com/books/application-of-nanotechnology-in-drug-delivery/mannan-as-a-promising-bioactive-material-for-drug-nanocarrier-systems."

Pandey, A., et al. (2015). Recent Advances in Thermochemical Conversion of biomasses.

Patel, S. and A. Goyal (2010). "Functional oligosaccharides: production, properties and applications."

Patel, S. and A. Goyal (2011). "Functional oligosaccharides: production, properties and applications." World Journal of Microbiology and Biotechnology **27**(5): 1119-1128.

Pérez, J., et al. (2002). "Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview." International Microbiology **5**(2): 53-63.

Pierce, b. (2002). "Extinction Coefficients: A guide to understanding extinction coefficients, with emphasis on spectrophotometric determination of protein concentration.". from https://classes.soe.ucsc.edu/bme220l/Spring11/Reading/Extinction-coefficients.pdf.

Qin, X., et al. (2011). "GAP Promoter Library for Fine-Tuning of Gene Expression in Pichia pastoris." Applied and Environmental Microbiology **77**(11): 3600-3608.

Rahikainen, J. L., et al. (2013). "Cellulase–lignin interactions—The role of carbohydrate-binding module and pH in non-productive binding." Enzyme and Microbial Technology **53**(5): 315-321.

Rath, A., et al. (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins." Proceedings of the National Academy of Sciences **106**(6): 1760-1765.

Redinbaugh, M. G. and R. B. Turley (1986). "Adaptation of the bicinchoninic acid protein assay for use with microtiter plates and sucrose gradient fractions." Analytical Biochemistry **153**(2): 267-271.

Reusch, W. (2013). "Carbohydrates ".

Richard, T. (1996). "The effect of lignin on biodegradability." from http://compost.css.cornell.edu/calc/lignin.html.

Ritchie, C. (2012). "Protein purification." from http://www.labome.com/method/Protein-Purification.html.

Sartorius (2000). "Ultrafiltration ".

Shackford, L. D. (2003). A comparison of pulping and bleaching of kraft softwood and eucalyptus pulps. Proceedings of the 36th International Pulp and Paper Congress.

Shallom, D. and Y. Shoham (2003). "Microbial hemicellulases." Current Opinion in Microbiology **6**(3): 219-228.

Shen, J., et al. (2010). "The prediction of elemental composition of biomass based on proximate analysis." Energy Conversion and Management **51**(5): 983-987.

Sluiter, A., et al. (2008). "Determination of structural carbohydrates and lignin in biomass." Laboratory analytical procedure.

Sun, J. X., et al. (2004). "Fractional extraction and structural characterization of sugarcane bagasse hemicelluloses." Carbohydrate Polymers **56**(2): 195-204.

Sun, Y. and J. Cheng (2002). "Hydrolysis of lignocellulosic materials for ethanol production: a review." Bioresource Technology **83**(1): 1-11.

Sørlie, M., et al. (2012). "Processivity and substrate-binding in family 18 chitinases." Biocatalysis and Biotransformation **30**(3): 353-365.

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

Tenkanen, M., et al. (1996). "Effect of side groups on the action of β-xylosidase from Trichoderma reesei against substituted xylo-oligosaccharides." FEBS Letters **399**(3): 303-306.

Timell, T. E. (1967). "Recent progress in the chemistry of wood hemicelluloses." Wood Science and Technology **1**(1): 45-70.

Ulbricht, R. J., et al. (1984). "A review of 5-hydroxymethylfurfural (HMF) in parenteral solutions." Fundamental and Applied Toxicology **4**(5): 843-853.

Vaaje-Kolstad, G., et al. (2010). "An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides." Science **330**(6001): 219-222.

Varnai, A. (2012). Improving Enzymatic conversion of lignocellulose to platform sugars. University of Helsinki.

Várnai, A., et al. (2014). "Expression of endoglucanases in Pichia pastoris under control of the GAP promoter." Microbial Cell Factories **13**: 57-57.

Vlasenko, E., et al. (2010). "Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases." Bioresource Technology **101**(7): 2405-2411.

Vocadlo, D. J., et al. (2001). "Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate." Nature **412**(6849): 835-838.

Willför, S., et al. (2008). "Spruce-derived mannans – A potential raw material for hydrocolloids and novel advanced natural materials." Carbohydrate Polymers **72**(2): 197-210.

Williams, S. (2013). "Retaining and inverting mechanism." from http://www.cazypedia.org/index.php/File:General_acid_base.png - filehistory.

Wyman, C. E. (1999). "BIOMASS ETHANOL: Technical Progress, Opportunities, and Commercial Challenges." Annual Review of Energy and the Environment **24**(1): 189-226.

Zeng, M. (2007). Characterization of cell wall deconstruction induced by aqueous pretreatment and enzyme hydrolysis.

Zhang, Y. H. P. and L. R. Lynd (2005). "Determination of the Number-Average Degree of Polymerization of Cellodextrins and Cellulose with Application to Enzymatic Hydrolysis." Biomacromolecules **6**(3): 1510-1515.

7 Appendix A

Chromatograms from protein purification; using Ionexchange and size exclution chromatography.

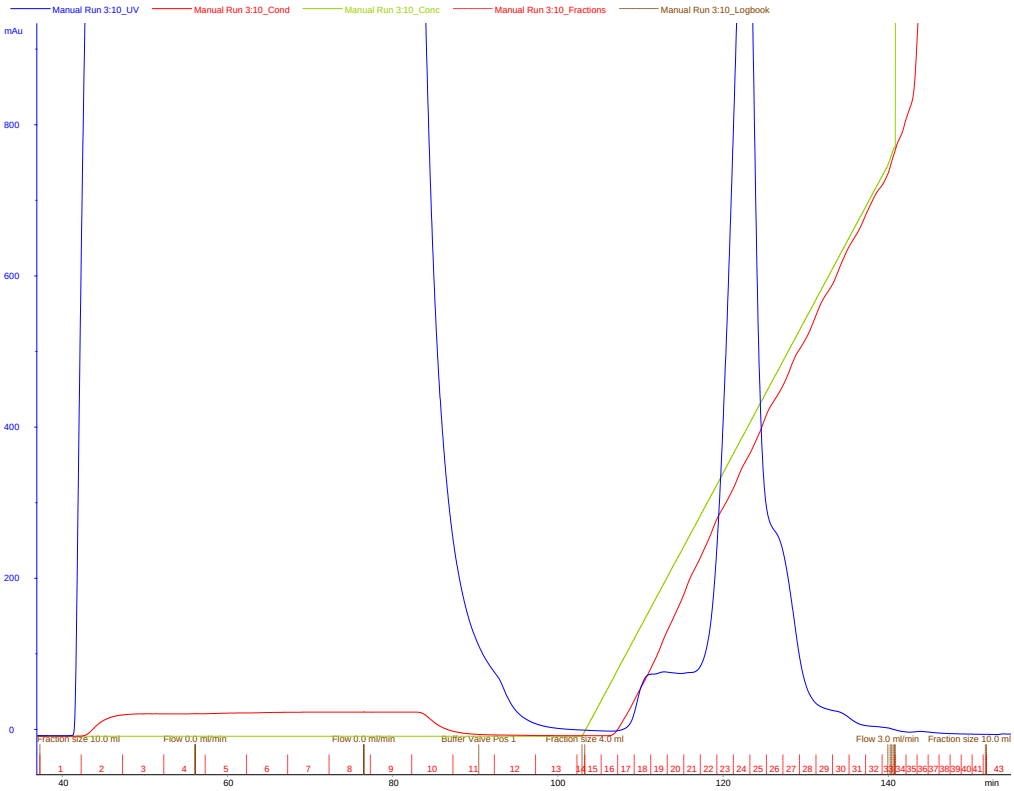


Figure D.1 *Ion exchange chromatography for MaCel45A. Fraction numbers are shown in red. Using a 5 ml-SP-FF- column in 20mM Na-Citrate buffer (pH4.0). Proteins were eluted with a linear gradient of 20 Column volumes ending 0.5M NaCl in 20mM Na-Citrate buffer 1M NaCl (pH4.0). Proteins were detected at 280nm (blue line) and collected in 4ml fractions. MaCel45A is found in fractions 22-31. The green line represents the NaCl gradient used for elution*

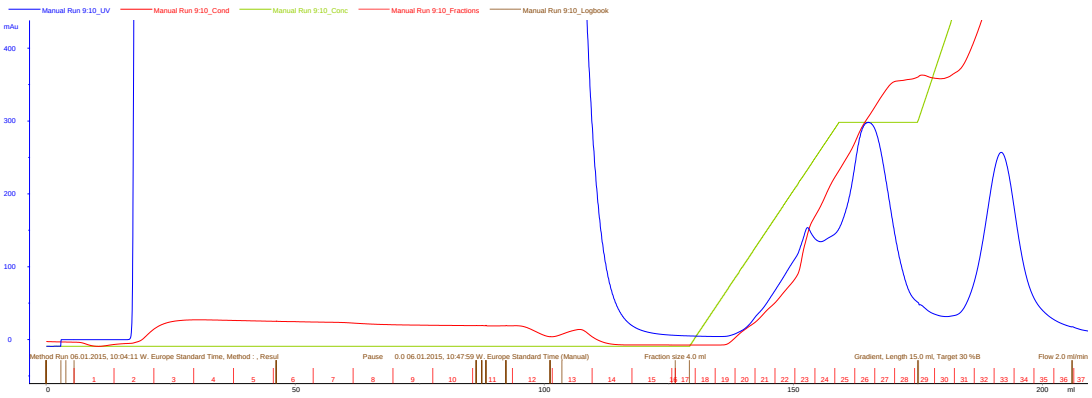


Figure D.2 *Ion exchange chromatography for TaCel45A. Fraction numbers are shown in red. Purified with a 5 ml-DEAE column in 20mM Na-Acetate buffer (pH 5). Proteins were eluted with a stepwise elutions of 20 Column volumes ending at 0.15M NaCl in 20mM Na-Acetate buffer 1M NaCl (pH5). Proteins were eluted between 0.12-0.15M. TaCel5A are found in fractions 20-29*
Eluted proteins was detected at 280nm (blue line) and collected in 4ml fractions. Green line represents the NaCl gradient used for elution

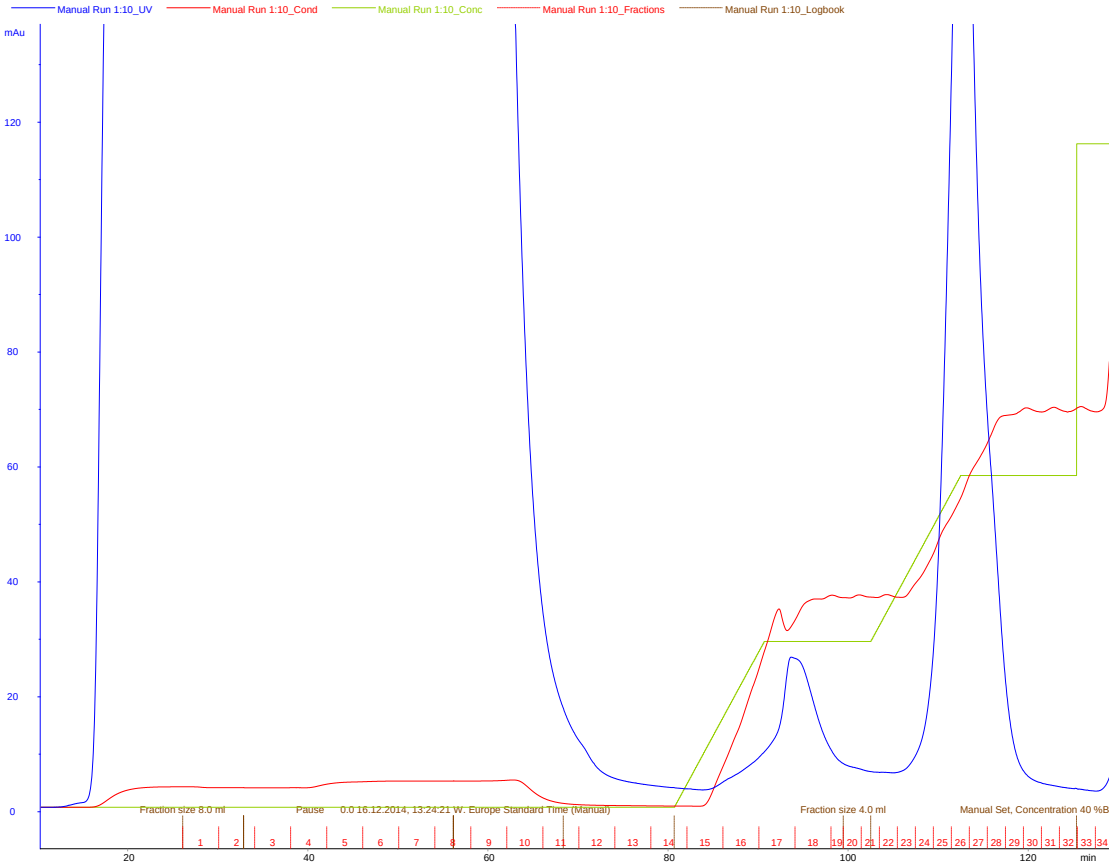


Figure D.3 *Ion exchange chromatography for AfCel12A. Fraction numbers are shown in red. Purified with a 5 ml-SP-FF- column in 20mM Na-Citrate buffer (pH3). Proteins were eluted with a Stepwise gradient of 20 column volumes, NaCl concentrations increasing from 0-0.1M and the further till all proteins are eluted with 20mM Na-Citrate buffer 1M NaCl (pH3). Eluted proteins was detected at 280nm and collected in 4ml fractions. AfCel5A are found in fractions 24-33. Green line represents the NaCl gradient used for elution*

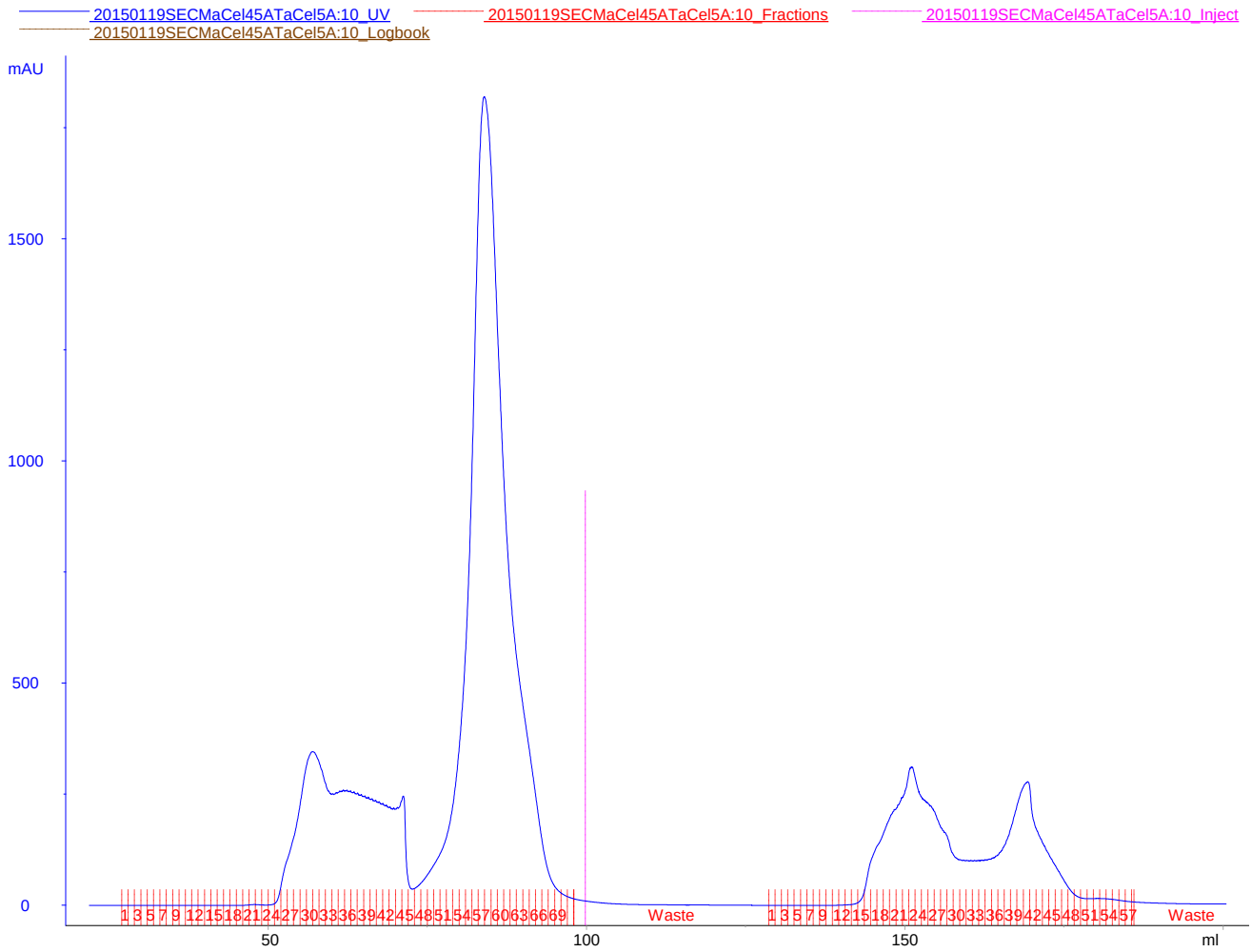


Figure D4. *Size exclution chromatography used for further purification of MaCel45A and, after ion exchange chromatography. Fractionation numbers are shown in red. Fractions containing the protein of interest from ion exchange are pooled and concentrated down to approximately 2 ml. The sample was applied through a 5ml loading loop to a high load 16/60 superdex column. The fractions are checked with SDS-PAGE. The chromatogram to the left is MaCel45A, The protein of interest is found in fractions 48-69. TaCel5A is to the right. Protein of interest is found in fraction 36 to 48. It was not a good separation of TaCel5A, so it needed to be separated one more time.*

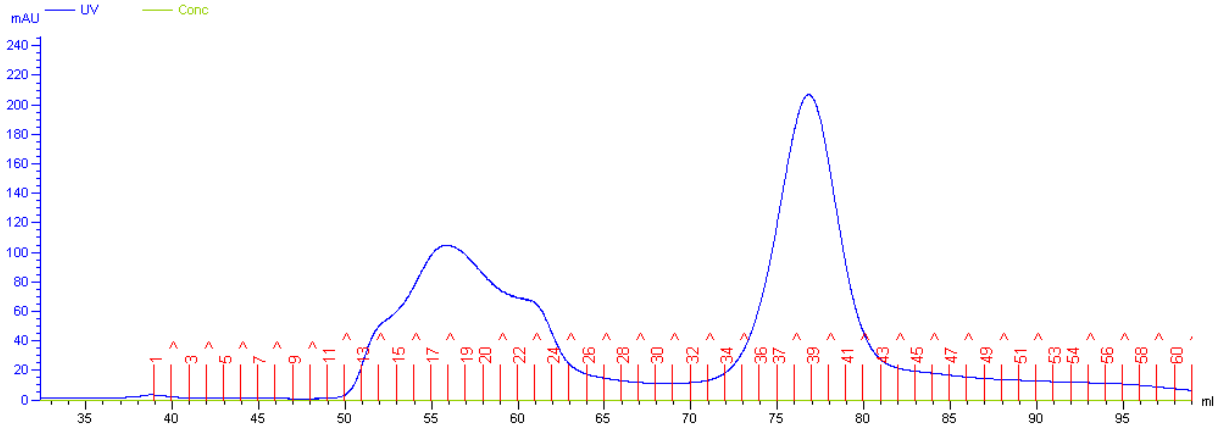
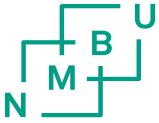


Figure D5*. Size exclusion chromatography for TaCel5A. Rerun after first round of SEC. The proteins were not separated well the first round. This chromatogram, however, show a good peak separation, confirmed by SDS.*



Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no