Norges miljø- og biovitenskapelige universitet

Master's Thesis 2018 60 ECTS

Faculty of Chemistry, Biotechnology and Food Science
Morten Sørlie

Molecular Evolution of the Substrate Specificity of Bacterial Lytic Polysaccharide Monooxygenases (LPMOs)

Helene Bjølgerud
Master of Biotechnology
Faculty of Chemistry, Biotechnology and Food Science

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out in Morten Sørlie's (Bioorganic) research group and the Protein Engineering and Proteomics (PEP) group at the Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences (NMBU) in the period from January 2017 to May 2018. This thesis was part of a larger research project where the objective was to unravel the molecular basis of LPMO activity.

I would like to express my gratitude to my supervisor Professor Morten Sørlie for allowing me to work in his research group and a very special thanks to my co-supervisor Dr. Bastien Bissaro for his ideas, guidance and comments during all parts of this thesis. His inspiring thoughts, encouragement and support has been of great importance to me – despite his busy schedule, he always made time to check in with me concerning the progression of my project.

I would also like to thank Dr. Anne Grethe Hamre, Dr. Zarah Forsberg and Dr. Åsmund Kjendseth Røhr for your practical advises in the lab.

Thanks to my fellow students, especially Rim Al-Sadawi and Synne L. Rooth, for the good company at the lab and during the writing of this thesis, as well as being a constant source of motivation and kindness throughout this past year.

Last, but not least, thanks to my friends and family for their support all the way throughout my five-year long education at Ås.

Ås, May 2018

Helene Bjølgerud

SUMMARY

The transition to a more environment-friendly economy has spurred the research on how to efficiently convert recalcitrant polysaccharides into soluble sugars. One of the major breakthroughs in the field has been the discovery of enzymes capable of disrupting the crystalline structures of polysaccharides. These enzymes, often referred to as lytic polysaccharide monooxygenases (LPMOs), are able to oxidize recalcitrant polysaccharides such as chitin and cellulose and play thus an important role in biomass conversion. Today, LPMOs are classified in families AA9, 10, 11, 13, 14 and 15 in the Carbohydrate-Active Enzymes database. The AA10 family is composed of LPMOs that are mainly of bacterial origin and that present the particularity to be active either on chitin, or on cellulose, or on both. Understanding the evolutionary divergence in substrate specificity (EDSS) among the AA10s would allow for a better understanding of the molecular basis of LPMO catalysis, and also to improve our capacity to predict the enzymatic phenotype of an LPMO based on its sequence. To tackle this question, the natural diversity of AA10s sequences has been analyzed in order to pinpoint potential "hotspot" residues involved in the EDSS via a so-called "correlated mutation analysis" (CMA). In this work, a chitin-active AA10 from the gram-negative bacterium *Cellvibrio japonicus* (*Cj*AA10A) was selected as a starting model enzyme. The aim of this research is to evolve *Cj*AA10A from a chitin-active phenotype towards a cellulose-active one. Importantly, *Cj*AA10A is a multi-modular enzyme composed of a catalytic domain linked to several carbohydrate binding modules (CBMs) that are specific on chitin. In this context, a first phase of the project consisted in engineering a platform enzyme by exchanging the chitin-specific CBMs of *Cj*AA10A for a cellulose-specific CBM. The catalytic domain of the resulting platform hybrid enzyme (i.e. chitin-active catalytic domain and cellulose-binding CBM) was then subjected to site-directed mutagenesis by targeting the "hotspot" residues identified via the CMA evocated above. Cloning, expression, purification and characterization of different variants of *Cj*AA10A, with shuffled linkers and CBM domains, were successful. The binding and activity of the variants were analyzed by MALDI-TOF MS and HPLC in order to identify how the changes introduced were affecting the enzyme properties. In addition, the H_2O_2 production ability of each enzyme was quantified.

SAMMENDRAG

Overgangen til en mer bærekraftig bioøkonomi har satt fart på forskning tilknyttet effektiv konvertering av vanskelig nedbrytbare polysakkarider til løselige sukkerarter. Et stort gjennombrudd innen dette feltet har vært oppdagelsen av enzymer som kan bidra til depolymerisering av krystallinske polysakkaridstrukturer. Disse enzymene, ofte referert til som lytiske polysakkarid monooksygenaser (LPMOer), kløyver glykosidiske bånd i polysakkaridkjedene ved oksidasjon, og spiller dermed en viktig rolle i biomassekonvertering. LPMOer er i dag klassifisert i familiene AA9, 10, 11, 13, 14 og 15 i databasen for karbohydrat-aktive enzymer, CAZy. AA10 familien består i hovedsak av LPMOer av bakteriell opprinnelse som er substratspesifikke til enten kitin, cellulose eller begge. Forståelsen av den evolusjonære divergensen i substratspesifisitet (EDSS) blant AA10ene vil gi en bedre forståelse av det molekylære grunnlaget for LPMO katalyse, men også for å bedre vår evne til å forutsi den enzymatiske fenotypen av LPMOer basert på deres sekvens. Det naturlige mangfoldet av AA10-sekvenser har blitt analysert for å identifisere potensielle "hotspot"-residuer involvert i EDSS via en såkalt "korrelert mutasjonsanalyse" (CMA). I denne studien har en kitin-aktiv AA10 fra den gram-negative bakterien *Cellvibrio japonicus* (*Cj*AA10A) blitt valgt ut som et startmodell-enzym. Målet med denne studien var å utvikle *Cj*AA10A fra en kitin-aktiv fenotype mot en cellulose-aktiv fenotype. *Cj*AA10A er et multimodulært enzym bestående av et katalytisk domene koblet via en linker til flere karbohydratbindende moduler (CBMer) som er spesifikke til kitin. CBMer har vist seg å være avgjørende for LPMO-stabilitet under katalyse. I denne konteksten bestod den første fasen i prosjektet av å konstruere et plattformenzym ved å bytte den kitin-spesifikke CBMen til *Cj*AA10A med en cellulose-spesifikk CBM. Det katalytiske domenet til den resulterende plattformenzymet (dvs. kitin-aktivt katalytisk domene med cellulose-bindende CBM) ble så utsatt for seterettet mutagenese ved å velge "hotspot" residuene identifisert via den tidligere nevnte korrelert mutasjonsanalysen. Kloning, utrykking, rensing og karakterisering av forskjellige varianter av *Cj*AA10A med endrede linkere og CBM-domener var vellykkede. Binding og aktivitet hos varianter ble analysert ved MALDI-TOF MS og HPLC for å identifisere hvordan de innførte endringene påvirket enzymegenskapene. I tillegg ble H_2O_2-produksjonsevnen til hvert enzym kvantifisert.

TABLE OF CONTENTS

ABBREVIATIONS

AA	Auxiliary activity
Ala (A)	Alanine
Ba	*Bacillus amyloliquefaciens*
BG	β-glucosidases
bp	Base pair
CAZymes	Carbohydrate-active enzymes
CBH	Cellobiohydrolase
CBM	Carbohydrate-binding module
CBP	Chitin-binding protein
Cd	Catalytic domain
CE	Carbohydrate esterase
CHB	Chitobiase
Cj	*Cellvibrio japonicus*
Cu	Copper
C1	Carbon 1
C4	Carbon 4
Da	Dalton
D-Glc*p*	D-glucopyranose
dH_2O	Distilled water of Milli-Q® quality
DNA	Deoxyribonucleic acid
dsDNA	Double stranded deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
Ef	*Enterococcus faecalis*
EG	Endoglucanases
EtOH	Ethanol
ExPASy	Expert Protein Analysis System
GC-content	Guanine-cytosine content
GH	Glycoside Hydrolase
Glc	Glucose
GlcNAc	*N*-acetyl-D-glucosamine
Glu (E)	Glutamic acid/glutamate

Gly (G)	Glycine
GT	Glycosyl transferase
His (H)	Histidine
HPAEC-PAD	High Performance Anion Exchange Chromatography with pulse amperometric detection
ICS	Ion Chromatography System
IEC	Ion Exchange Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny Broth
LDS	Litium Docedyl Sulphate
LPMO	Lytic Polysaccharide Monooxygenase
MALDI	Matrix Assisted Laser Desorption/Ionization
MS	Mass Spectrometry
NMR	Nuclear magnetic resonance
PASC	Phosphoric acid swollen cellulose
PCR	Polymerase chain reaction
Phe (F)	Phenylalanine
PL	Polysaccharide lyase
PyMOL	Molecular visualization program of protein crystals
RNAse	Ribonuclease
Sc	*Streptomyces coelicolor*
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
Ser (S)	Serine
Sm	*Serratia marcescens*
SOC	Super Optimal Broth with Catabolite repression
ssDNA	Single stranded deoxyribonucleic acid
Ta	*Thermoascus aurantiacus*
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine
WT	Wild type

1. INTRODUCTION

Today, the high worldwide demand for energy, the uncertainty regarding petroleum sources and a general concern over global climate changes have led to a resurgence in development of alternative energies and processes that could replace fossil-based fuels and chemicals. In many countries, the initiation of extensive research and development programs in biofuels is a direct response. Biomass, whether as cellulosic material, sugar crops or starch crops, represents an abundant and renewable carbon resource for sustainable production that can provide liquid transportation fuels, organic fuels and chemicals that are now primarily made from petroleum (Faaij, 2006). In response, countries worldwide develop visions and scenarios regarding the amount of transportation fuels to be derived from biofuels within the next 20-30 years (Garland et al., 2012; Ryan et al., 2006).

Currently, production of biofuels, chemicals and biomaterials are mainly generated from first-generation feedstock (e.g. wheat, corn and sugarcane) (Hein & Leemans, 2012). However, a transition to using primarily second-generation feedstock (e.g. palm oil production) is of great interest due to environmental issues and the reliance of food sources. About 70% of the annual global production of plant biomass consists of plant cell-walls, and about three quarters of these being polysaccharides (Duchesne & Larson, 1989).

Plant carbohydrates ($C_m(H_2O)_n$), which are the conversion product of CO_2 and H_2O via photosynthesis, represent therefore a great opportunity to improve energy security. The two most abundant polysaccharides found in nature are (hemi)celluloses and chitin and have thus been subject to extensive research for decades. Because of their recalcitrant nature, depolymerization and thus, degradation, constitute both a biological and industrial challenge. The transition to a more environment-friendly economy has therefore spurred research on enzymes capable of efficiently degrading these recalcitrant polysaccharides (Himmel et al., 2007). On that note, one of the major breakthroughs in the field has been the discovery of enzymes able to disrupt the crystalline structure of polysaccharides via oxidation, namely the lytic polysaccharide monooxygenases (LPMOs), which are at the heart of the present project. In the next paragraphs, the nature of the main recalcitrant polysaccharides, namely chitin and cellulose, will be introduced. Then, microbial and enzymatic degradation of structural

polysaccharides will be discussed. The last paragraphs are dedicated to the discovery and increasing understanding of LPMOs.

1.1. CARBOHYDRATES

Alongside proteins, nucleic acids and lipids, carbohydrates (i.e. sugars) make up one of the four major classes of biomolecules. Glucose (Glc) was the first carbohydrate obtained in pure form with the chemical formula $C_6H_{12}O_6$. This led to the belief of glucose being a "hydrate of carbon", $C_6(H_2O)_6$, and the class of biomolecules was called carbohydrates.

Although carbohydrates are compounds of aldehydes or ketones with hydroxyl groups, the name remained. Monosaccharides are building blocks of di-, oligo- and polysaccharides. Carbohydrates play diverse roles in all forms of life with functions as nutrients for organisms, energy storage, cell to cell communication and as structural elements in cells walls (e.g. cellulose and chitin).

1.1.1. CELLULOSE

The main constituent of the plant cell wall is cellulose, a linear non-branched homopolymer made up of D-glucopyranose (D-Glc*p*) units linked by β-1,4-glycosidic bonds (**Figure 1.1**). Every second D-Glc*p* unit is rotated 180˚ along the sugar plane compared to the adjacent unit, allowing hydrogen bonding between parallel chains (Cocinero et al., 2009). The usual length of a cellulose chain is described by the degree of polymerization (DP; i.e. the number of monosaccharide units in the chain). Depending on the source of cellulose, the degree of polymerization is estimated to range from a few hundreds to several thousand glucose units (Hallac & Ragauskas, 2011). A cellulose microfibril consists of strong inter-chain interactions that generate a planar structure. This, and the huge potential for hydrogen bond formation, contributes to the insolubility of cellulose in water (Brown, 2004; Medronho et al., 2012). All the hydroxyl groups in elementary cellulose fibrils are positioned equatorially, while the hydrogens are placed axially. This arrangement gives the fibrils polar sides and a hydrophobic face both under and over the sheet (Beckham et al., 2011). This enables the cellulose sheets to pack together by inter-sheet hydrogen bonds and van der Waals interactions. A dense layer of water usually coats the

hydrophobic surface of the microfibril, which makes crystalline cellulose highly resistant to acidic hydrolysis (Himmel et al., 2007).

Seven polymorphs of cellulose (Iα, Iβ, II, III_I, III_{II}, IV_I and IV_{II}) can exist in nature (O'Sullivan, 1997). Cellulose Iα and Iβ are the two native crystalline forms that are the most abundant in nature (Brown, 2004). These two forms of cellulose differ in hydrogen bonding patterns, but have the same conformation with a parallel chain arrangement. The cell wall of higher plants, such as cotton, wood and ramie are rich in cellulose Iβ, whereas cellulose Iα can be found in green algae and *Valonia* (Payne et al., 2015). Cellulose II, naturally produced by some algae, is usually the most crystalline and thermodynamically stable form. The cellulose III and IV polymorphs are "artificial" polysaccharides resulting from chemical and/or physical pretreatment methods of cellulose, although certain treatments (Swatloski et al., 2002).

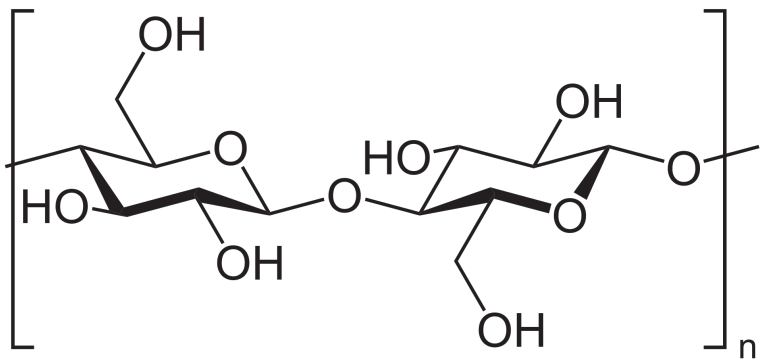


Figure 1.1. Chemical structure of cellulose. Cellulose is a linear polymer of D-glucopyranose (D-Glc*p*) units connected by β-1,4-glycosidic bonds, each rotated 180˚ with respect to its neighbor. Cellobiose is the repeated unit in cellulose, which is composed of two glucose saccharides joined together.

1.1.2. CHITIN

Chitin, the structural component in yeast and fungal cell walls, in the exoskeletons of arthropods and in insects, is often considered as the second most abundant polymer in nature, after cellulose (Tharanathan & Kittur, 2003). Chitin share similarities with cellulose since it consists of linear β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) units with each unit rotated 180˚ with respect to its neighbor (**Figure 1.2**). In nature, two crystalline polymorphs of chitin are found, α- and β-chitin, as well as a third less abundant allomorph, γ-chitin. The most abundant allomorph is α-chitin where the polysaccharide chains are arranged in an antiparallel fashion. This makes the α-chitin more dense and rigid, with thus lower solubility and swelling properties, compared to

β-chitin (Kumirska et al., 2010; Kurita et al., 2005), which consists of a more open structure due to a parallel chain arrangement (Gardner, K. H. & Blackwell, 1975; Minke & Blackwell, 1978). Due to strong intermolecular hydrogen bonding, chitin, alike cellulose, is insoluble in water (Khoushab & Yamabhai, 2010). In contrast to cellulose, chitin can be modified via by deacetylations (by chemical or enzymatic treatment), resulting in a mixed structure made of β-1,4-D-glucosamine (deacetylated) and GlcNAc (acetylated) units. The polysaccharide becomes soluble in water if the degree of deacetylation is higher than 50%, which is then referred to as chitosan (Younes & Rinaudo, 2015). Chitosan and chitooligosaccharides have a number of applications in agriculture such as cosmetics, wastewater treatment, and medicine (Aam et al., 2010) and are therefore valuable as biodegradable and biocompatible products.

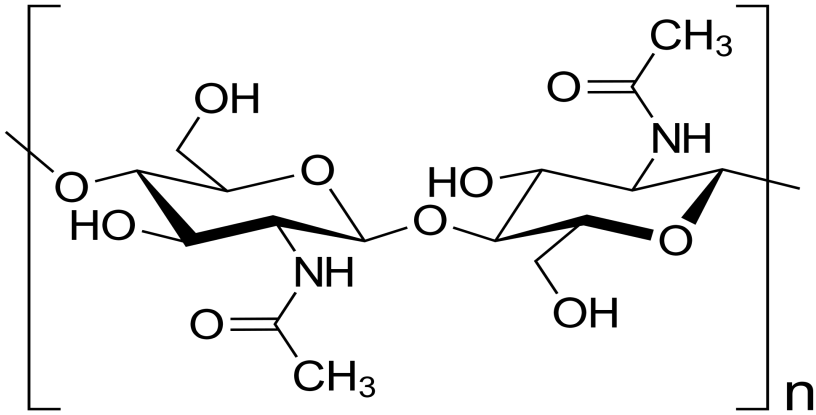


Figure 1.2. The chain structure of chitin. Chitin share similarities with cellulose since it consists of linear β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) units with each unit rotated 180˚ with respect to its neighbor.

1.2. MICROBIAL DEGRADATION OF STRUCTURAL POLYSACCHARIDES

Because of the high degree of structural order in polysaccharides such as cellulose and chitin, they are resistant to microbial degradation. Their crystalline and insoluble structures (e.g. insect shell or plant cell wall) protect the organisms by providing mechanical strength and chemical resistance towards degradation. However, several microbes have developed efficient enzyme systems to allow the deconstruction of these crystalline polysaccharides. This provides a good starting point for conversion of polysaccharides into soluble sugars (i.e. nutrients for growth). Many bacteria, fungi

and protozoa have developed (ligno)cellulolytic arsenals, notably composed of enzyme systems such as carbohydrate-active enzymes (CAZymes) to degrade polysaccharides (Cragg et al., 2015). These microbes occupy a number of ecosystems, including lakes, rivers, soils, decaying vegetable matters and more (Ljungdahl & Eriksson, 1985). Aerobic cellulolytic and chitinolytic microorganisms, especially wood-decaying fungi, secrete vast arrays of free enzymes, and these enzymes act synergistically to degrade biomass (Resch et al., 2013). In the case of lignocellulose-degrading microorganisms, a wide variety of enzymes is usually recruited to target the different polymers found in plant cell wall: cellulose, hemicellulose and in some cases also lignin. Regarding chitinolytic systems, the machinery of *Serratia marcecens* is one of the most known systems for the conversion of insoluble polysaccharides. Four chitin-active enzymes are included in this machinery: ChiA, ChiB and ChiC including CBP21, a surface-active CBM33-type lytic polysaccharide monooxygenase that introduces chain breaks by oxidative cleavage (Fuchs et al., 1986; Horn et al., 2012; Vaaje-Kolstad et al., 2010). Interestingly, studies of chitin degradation show that the chitinolytic cocktail from the gram-negative soil bacterium *Cellvibrio japonicus* degrade β-chitin in a similar way as a *S. marcescens* cocktail (Tuveng et al., 2016). Inspired from microbial strategies, industrials have designed several enzyme cocktails to achieve relatively efficient biomass saccharification. Along with our increasing understanding of natural biomass conversion process, the efficiency of commercial enzymatic cocktails is constantly in progression (Johansen, 2016).

1.3. ENZYMATIC DEGRADATION OF STRUCTURAL POLYSACCHARIDES

1.3.1. CARBOHYDRATE-ACTIVE ENZYMES

The carbohydrate active enzymes (CAZymes) constitute a broad class of enzymes involved in the synthesis, modification and/or deconstruction of carbohydra-containing biomolecules. Carbohydrate compounds are among the most structurally diverse molecules in nature, this because of the variety of monosaccharides in combination with the variety of sugar linkages. Thus, facing such structural diversity, a large variety of enzymes acting on these substrates are needed in nature (Cantarel et al., 2009).

With the increasing number of CAZymes over the years, the CAZy database [www.cazy.org (Cantarel et al., 2009; Lombard et al., 2014)] has been built up. Because this classification system is primarily based on amino acid sequence similarities, the classification correlates better with mechanism and structural folding than with enzyme specificity. CAZymes include glycosyl transferases (GTs), polysaccharide lyases (PLs) and carbohydrate esterases (CEs). Notably promped by the discovery of LPMO activity (section 1.4.), the CAZy database has recently undergone a major expansion with the creation of a new class called auxiliary activities (AAs) gathering a collection of redox enzymes (Levasseur et al., 2013). This class, unlike the other enzyme classes, contains enzymes that act not exclusively on carbohydrates.

1.3.2. ENZYMATIC DEGRADATION OF CELLULOSE

For a complete enzymatic degradation of (pure) cellulose, three types of glycoside hydrolases (GHs) are involved. These three types of enzymes are *exo*-β-1,4-glucanase (also termed cellobiohydrolase, CBH), *endo*-β-1,4-glucanases (endoglucanases (EGs)), and β-glucosidases (BGs) (Merino & Cherry, 2007).

Hydrolysis of internal glycosidic bonds by EG in the cellulose chain results in the generation of new reducing and non-reducing chain ends. These can be accessed by CBHs, which in turn hydrolyse in a processive (bound to a carbohydrate binding module) or non-processive manner, the cellulose chain into cellobiose units (Henrissat et al., 1985; Taylor et al., 2018).

The cellobiose released is the substrate for the BG, enzymes that mitigates CBH product inhibition arising from cellobiose by hydrolyzing cellobiose into monomeric product inhibition arising from cellobiose by hydrolyzing cellobiose into monomeric glucose. The synergy between these enzymes can be described in two different types: the synergism between EGs and CBHs, termed the endo-exo-synergism and the synergism between different CBHs referred to as exo-exo-synergism (Igarashi et al., 2011). By working in synergy, these enzymes enhance the activity of each other, and contribute to an efficient degradation of lignocellulosic biomass.

These three main catalytic activities in an enzyme cocktail are not enough to achieve the maximum theoretical yield of glucose that could be obtained from lignocellulosic biomass (Harris et al., 2010). Both CBHs and EGs struggle to perform hydrolysis on crystalline cellulose, even after pretreatment. Depolymerization of crystalline

cellulose is much more arduous compared to solubilized chains because of the rigid hydrogen-bonding network that contributes to recalcitrance in crystalline cellulose. Thus, for CBHs to carry out several turnovers before dissociating from the substrate, strong binding to the chain end is required. However, the strong binding reduces the rate of depolymerization. The latter features can be referred to as the "cost and benefits of processivity" (Horn et al., 2016), so the hydrolytic catalysis of EGs will therefore require a transition of conformation to the glucose ring to a more unfavorable conformation (Rye & Withers, 2000). Neither EGs nor CBHs are specialized in degrading highly crystalline regions of cellulose since they can hardly bring into their active site the cellulose chain that is embedded in a crystalline lattice. Therefore, enzyme cocktails composed of only GHs are therefore too inefficient and expensive to be profitable in a biorefinery context (Hemsworth et al., 2016). This is where the discovery of AAs comes in handy.

1.3.3. ENZYMATIC DEGRADATION OF CHITIN

Similarly to the cellulolytic system, the degradation of chitin occurs as chitinases, belonging to GH families 18, 19, 23 and 48, randomly attacks a point along the chains (endochitinase) or by attacking either the reducing or the non-reducing end of the chitin chain (exochitinase) (Horn et al., 2006). The resulting soluble chitin oligomers, mainly N,N'-diacetylchitobiose ($[GlcNAc]_2$), are subsequently cleaved into monomeric GlcNAc units by D–N-acetylhexosaminidases (also called chitobiase) that are found in GH families 3, 20, 84 and 116 (Beier & Bertilsson, 2013). Alternatively, monomeric glucosamine can be produced as the result of deacetylation of chitin to chitosan and subsequent hydrolysis by chitosanases and glucosaminidases (Zhao et al., 2010). Hydrolysis of the insoluble chitin is therefore the result of the synergistic action of various chitinases (Suzuki et al., 2002; Vaaje-Kolstad et al., 2013).

The recently discovered LPMOs have an activity that was first demonstrated on chitin and have been shown to contribute to the depolymerization of the rigid and crystalline substrate by catalyzing oxidative cleavage of insoluble chitin chains (Vaaje-Kolstad et al., 2010).

1.4. LYTIC POLYSACCHARIDE MONOOXYGENASES (LPMOs)

1.4.1. DISCOVERY AND EXPANSION

Up until the early years of the 21st century, the traditional view of enzymatic degradation of cellulose and chitin involved the action of mainly hydrolytic enzymes (e.g. cellulases and chitinases). However, it was proposed that biological solubilization (i.e. degradation) of cellulose would require at least two steps as early as in 1950. The first step would be performed by an activity that disrupts the cellulose structure, before hydrolysis of the β-(1,4)-glucan chain (Reese et al., 1950). We know today that this first step involves enzymes that are collectively referred to as LPMOs. In 2010 and 2011, it was discovered that proteins/enzymes belonging to family 33 of carbohydrate binding modules (CBMs) and family 61 of GHs are oxidative enzymes capable of cleaving chitin and cellulose chains (Forsberg et al., 2011; Phillips et al., 2011; Quinlan, R. J. et al., 2011; Vaaje-Kolstad et al., 2010). The enzymes of CBM33 and GH61 were reclassified as auxiliary activities (AAs) in the families 9 and 10 respectively. Since the establishment of these founding LPMO families, four new LPMO families were reported, in 2014 as AA11 (Hemsworth et al., 2014) and AA13 (Leggio et al., 2015; Vu et al., 2014a) and in 2018 as AA14 (Couturier et al., 2018) and AA15 (Sabbadin et al., 2018). Families AA9, AA11, AA13 and AA14 are strictly of fungal origin, family AA10 comprises enzymes from bacteria, viruses and some eukaryotic organisms and AA15 have been so far mainly found in insects. Family AA9s have been shown to target cellulose and some hemicelluloses (Agger et al., 2014; Phillips et al., 2011; Quinlan, R. J. et al., 2011), oligosaccharides and xylan (Frommhagen et al., 2015; Isaksen et al., 2014), whereas AA11s and AA13s have been demonstrated to cleave chitin and starch, respectively (Hemsworth et al., 2014; Leggio et al., 2015; Vu et al., 2014a). AA10s, mainly found in bacterial genomes (with a few viral members), have been shown to act on cellulose, chitin or both (Forsberg et al., 2011; Forsberg et al., 2014b; Vaaje-Kolstad et al., 2010). Fungal AA14s target xylan structures only in plant cell wall context (Couturier et al., 2018) while the insect AA15s have been shown to oxidize either cellulose or chitin (Sabbadin et al., 2018). The known structures of the different LPMOs show that they all share a high degree of structural similarity (Vaaje-Kolstad et al., 2017). Today, AAs are known as enzymes that carry out oxidative cleavage of glycosidic bonds in

crystalline substrates such as cellulose and chitin (Hemsworth et al., 2013; Horn et al., 2012).

1.4.2. PHYLOGENETIC RELATIONS

As of April 2018, the number of characterized LPMOs was of 26 in AA9s; 21 in AA10s; 1 in AA11s; 3 in AA13s; 2 in AA14s and 2 in AA15s. The LPMOs that are encoded in the genomes of biomass-degrading organisms show great variation. For instance, in the saprophytic fungus *Cheatomium globosum*, as many as 40 AA9 genes have been reported (Busk & Lange, 2015). Although a few bacterial genomes have been shown to harbor up to seven LPMO genes, bacterial genomes usually contain only one or two AA10 genes. As detailed above, LPMOs vary greatly in substrate preferences. The sequence identity between members of different LPMOs families is usually lower than 20% and very little is known about the evolution of LPMO domains between enzymes of bacterial and fungal origins. It is plausible that LPMOs of bacterial origin, mainly found in the AA10 family, are direct descendants of a putative LPMO common ancestor. As an illustration of the latter point, AA10 is the only family containing LPMOs active on chitin, cellulose or both whereas other LPMO families usually contain more "specialized" enzymes. A phylogenetic study conducted by Book et al. (2014) on AA10 suggested that these LPMOs are under selection to change their function, which may optimize cellulolytic activity. The structural similarity but absence of significant sequence similarity between AA9 and AA10 families suggests that these enzyme families share an ancient ancestral protein. Therefore, a basis for identifying and classifying additional cellulolytic or chitinolyitc LPMOs is provided.

Although the majority of family AA10s exists as single domain enzymes, several enzymes also contain one or more additional carbohydrate-binding modules (CBMs, discussed further in section 1.4.7.) that may reflect substrate preferences (Cuskin et al., 2012; Hervé et al., 2010). While most of chitin-active AA10s are not attached to a chitin-binding domain, many cellulose-active AA10s seem to have acquired a cellulose-binding domain, a fact that may reflect a compensatory mechanism required during the course of the evolution of the catalytic domain from chitin to cellulose-specificity (**Figure 1.3**). The cellulose-active AA9s have proved to release C1-oxidized or C4-oxidized products, or a mixture of both, with different product profiles. In 2011, a phylogenetic sub-classification was proposed according to the

oxidative regioselectivity. These enzymes were divided into three main groups: type 1 (C1-oxidizers), type 2 (C4-oxidizers) and type 3 (C1/C4-oxidizers). A subset of type 3 LPMOs was also added (LPMO3* subfamily), with enzymes that appear to have lost C4 activity and only carries out C1 oxidation (Phillips et al., 2011).

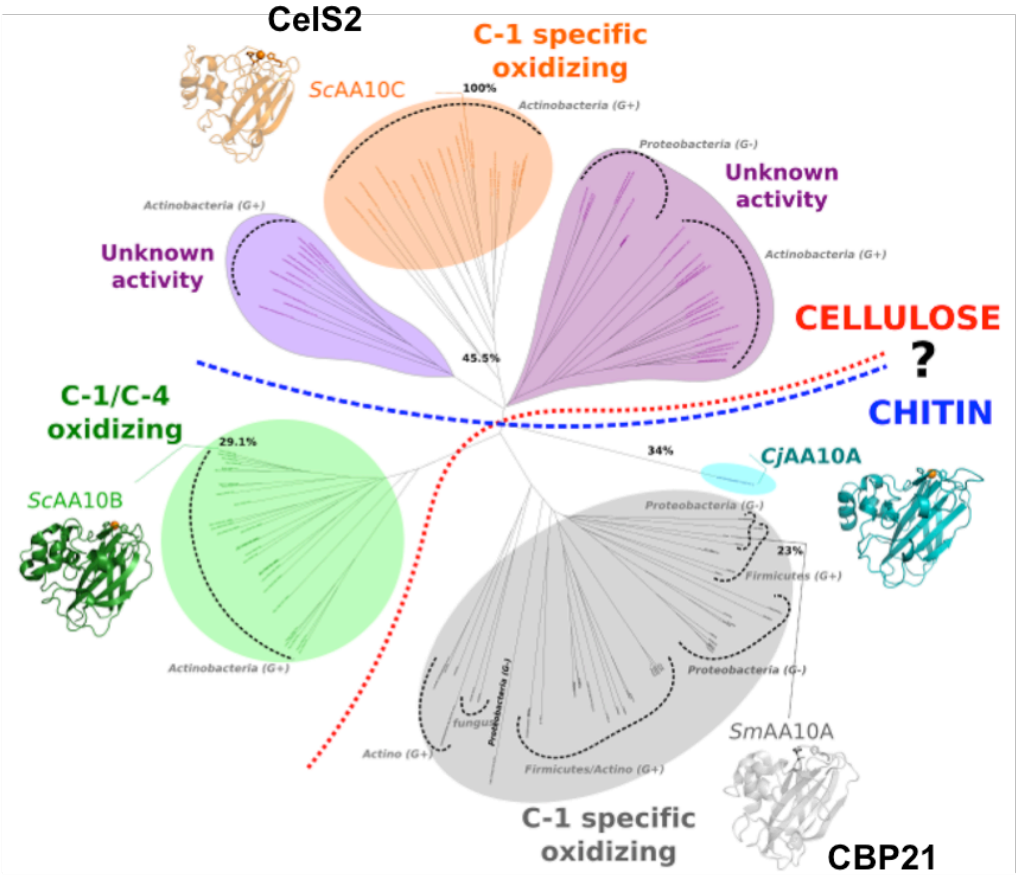


Figure 1.3. Phylogenetic Tree of LPMOs from Family 10 (bacterial origin). Three main groups can be observed: the C1 oxidizing chitin-active (grey circle) and the cellulose-active oxidizing at C1 (orange circle) or C1/C4 (green circle).

1.4.3. THREE-DEMENTIONAL STRUCTURES OF LPMOs

In general, LPMOs share a similar overall fold involving a compact β-sandwich core consisting of two β-sheets with a varying number of β-strands (usually 8-10). The strands are connected by a series of short loops located between the β-strands, with a varying number of α-helix insertions. Longer loops often contain short helices that build the flat surface for binding to carbohydrates. While the β-sandwich is highly conserved, most of the structural diversity in LPMOs is found in a loop referred to as loop 2 (L2). This loop differs in size and conformation, and may play an important role in binding and substrate specificity, as it constitutes large parts of the substrate-binding surface (Book et al., 2014; Forsberg et al., 2014a; Forsberg et al., 2016) (**Figure 1.4**). Other differences occur outside the L2 loop, which may explain functional variation (Vu et al., 2014a). The active site that catalyzes oxidative cleavage of carbohydrates is located in the middle of the flat binding-surface.

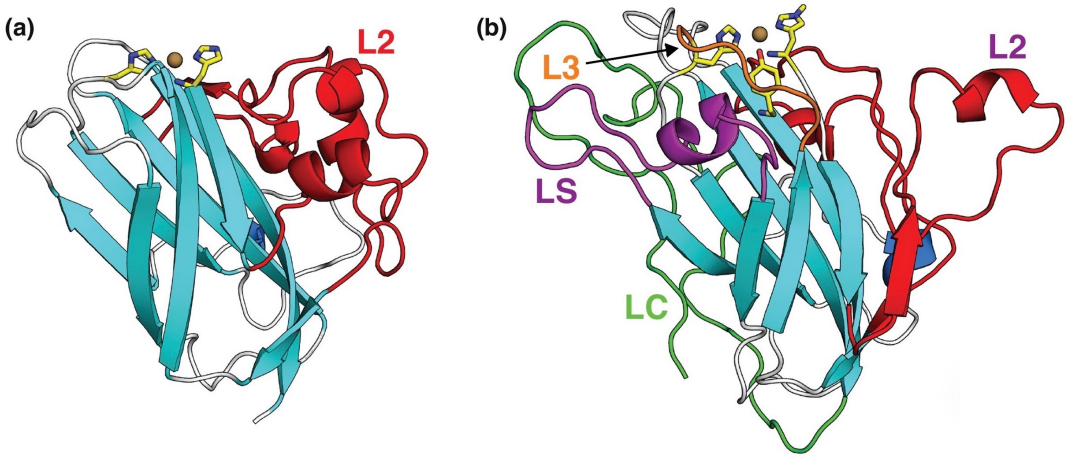


Figure 1.4. Structural diversity of LPMOs. Panels (a) and (b) show the typical fold of an LPMO10 illustrated by the structure of CBP21 from *Serratia marcescens* (PDB id: 2BEM) and an LPMO9 illustrated by *Nc*AA9M from *Neurospora crassa* (PDB id: 4EIS), respectively. Loops important for forming the substrate-binding surface (L2, LS and LC) are indicated. Figure taken from (Vaaje-Kolstad et al., 2017).

1.4.4. STRUCTURAL DESCRIPTION OF THE COPPER ACTIVE SITE

The active site of LPMOs consists of a metal binding site, a type 2 copper site essential for the LPMO activity (Vaaje-Kolstad et al., 2012; Westereng et al., 2011). The copper ion is tightly bound by three nitrogen ligands contributed by two histidine residues in a motif referred to as the histidine brace (Quinlan, R. J. et al., 2011; Aachmann et al., 2012). A square plane geometry with a 30˚ tilt relative to the binding site plane is formed by the three N ligands together with a co-substrate (Gudmundsson et al., 2014). The T-shaped histidine brace that coordinates the copper ion is the key structural element found in all LPMO active sites and is essential for catalysis. The axial positions differ between the LPMO families. Cu(II) is, when inactive, coordinated by tyrosine in the axial position, a conserved residue for AA9s, AA11s and AA13s. In the AA10A family, a hydrophobic residue occupies these positions, usually a phenylalanine (e.g. Phe187 in the AA10 from *Serratia marscecens* (*Sm*AA10A, also known as CBP21). Also, a conserved alanine (Ala112 in CBP21) is found on the solvent access side (Hemsworth et al., 2013; Vaaje-Kolstad et al., 2017). The AA11 structure shows features from both AA9s and AA10s.

The copper can only be coordinated by two to four ligands when reduced from Cu(II) to Cu(I) with a delivery of an electron donor (E^0_{red} = +0.15 V) (Beeson et al., 2015).

1.4.5. REACTION MECHANISM

Since the discovery of LPMOs, several plausible scenarios for the reaction mechanism have been suggested (Beeson et al., 2015; Walton & Davies, 2016). The first establishment of the oxidative properties of LPMOs (Vaaje-Kolstad et al., 2010), was made through the detection of unidentified modified chito-oligosaccharides in reactions carried out with the AA10 from *Serratia marscecens* (CBP21) in presence of O_2 and a reductant. By experiments performed in $^{18}O_2$-saturated conditions the "mysterious" products were identified as chito-oligosaccharides with an oxidized sugar at the reducing end. By monitoring the product profile with MALDI-TOF MS, the oxidized products showed a mass increase of two atomic mass units (amu) compared to those that did not contain isotope-labeled molecular oxygen. By performing this isotope labeling, it was demonstrated that the two oxygen atoms introduced at the chain ends of oxidized products comes from water and molecular oxygen (O_2) (**Figure 1.5**).

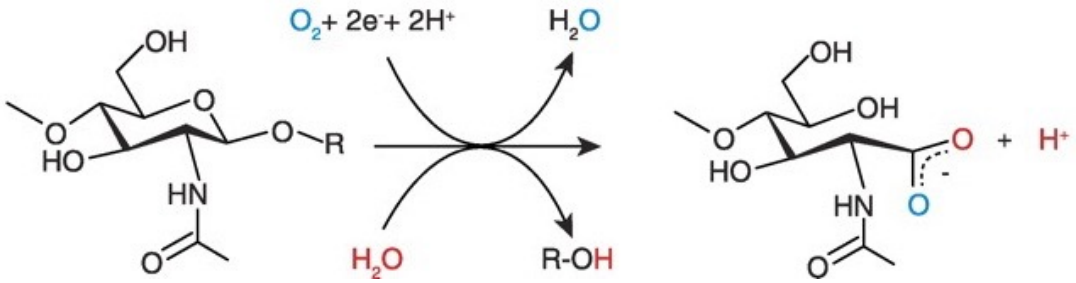


Figure 1.5. Scheme for the enzymatic reaction catalyzed by *Sm*AA10A (CBP21). The two oxygens in the final oxidized product come from molecular oxygen (blue) and from water (red). Figure taken from Vaaje-Kolstad et al. (2010).

The requirement of molecular oxygen for catalysis was confirmed when removal of dissolved molecular oxygen in the reaction solution inhibited CBP21 activity (Vaaje-Kolstad et al., 2010). The activity has been shown to be dependent on the presence of a divalent metal ion at the active site. It was also shown that the activity was greatly enhanced by an external electron donor [i.e. ascorbic acid (Vaaje-Kolstad et al., 2010)]. The exact electron delivery mechanisms to the active site are still unclear. The O_2 and reductant-depencies of reaction, along with the oxygenated nature of the products, led scholars to widely recognize these enzymes as monooxygenases, hence later called LPMOs (Horn et al., 2012), or more controversially PMOs (Beeson et al., 2012). The classical monooxygenase reaction can be written R-H + $2e^-$ + $2H^+$ + O_2 → R-OH + H_2O. In addition to O_2, the LPMO requires therefore the delivery of two electrons and protons during each catalytic cycle. When binding to a substrate, the LPMO active site, located at the interface between the polysaccharide and enzyme surfaces (**Figure 1.6**), will most likely be inaccessible for external electron donors. Thus, it was proposed that long-distance electron delivery would take place (Hemsworth et al., 2013).

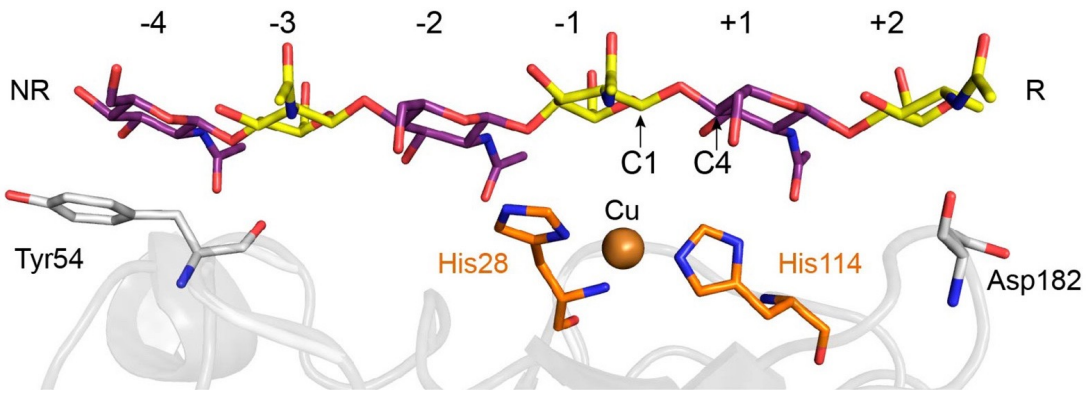


Figure 1.6. Model of *Sm*AA10A (performed on β-chitin) interacting with $(NAG)_6$. The histidine brace composed of His28 and His114 and bound to Cu are colored in orange. The side chains of Tyr54 and Asp182, in subsites −4 and +2, respectively, are also shown (gray sticks). C1 of the NAG unit located in subsite −1 is the carbon subject to hydroxylation during catalysis by *Sm*AA10A. Figure taken from Bissaro et al. (2018).

In 2011, it was demonstrated that the metal occupying the active site is a copper atom (Quinlan, J. R. et al., 2011) and it was suggested that an electron donor reduces the active site Cu(II) to Cu(I) which then binds molecular oxygen (Phillips et al., 2011). From this point, several mechanistic scenarios have been proposed (Beeson et al., 2015; Kim et al., 2014; Walton & Davies, 2016) but none of them has been experimentally fully demonstrated. A Cu(II)-superoxide enzyme species has been spectroscopically characterized (Kjaergaard et al., 2014), but this only in absence of polysaccharide substrate. Anyhow, the common outcome of all putative pathways is the hydroxylation of the glucan chain at the C1 or C4 carbons involved in the glycosidic bond. Such hydroxylation was proposed to destabilize the glycosidic bond and induce its cleavage by an elimination reaction (Beeson et al., 2012). The latter cleavage leads to the production of an oxidized end (e.g. a lactone when oxidation occurring at C1 position) and a non-modified end (**Figure 1.7**). The resulting mono-oxygenated product undergoes a spontaneous hydrolysis yielding an aldonic acid or a gem-diol for C1- and C4-oxidizers, respectively. Such reaction is irreversible. In absence of carbohydrate substrate, it has been shown that LPMO could indeed activate O_2 leading *in fine* to the production of H_2O_2 ($O_2 + 2e^- + 2H^+ \rightarrow H_2O_2$) (Kittl et al., 2012) through a mechanism that remains to be clarified (Span et al., 2017).

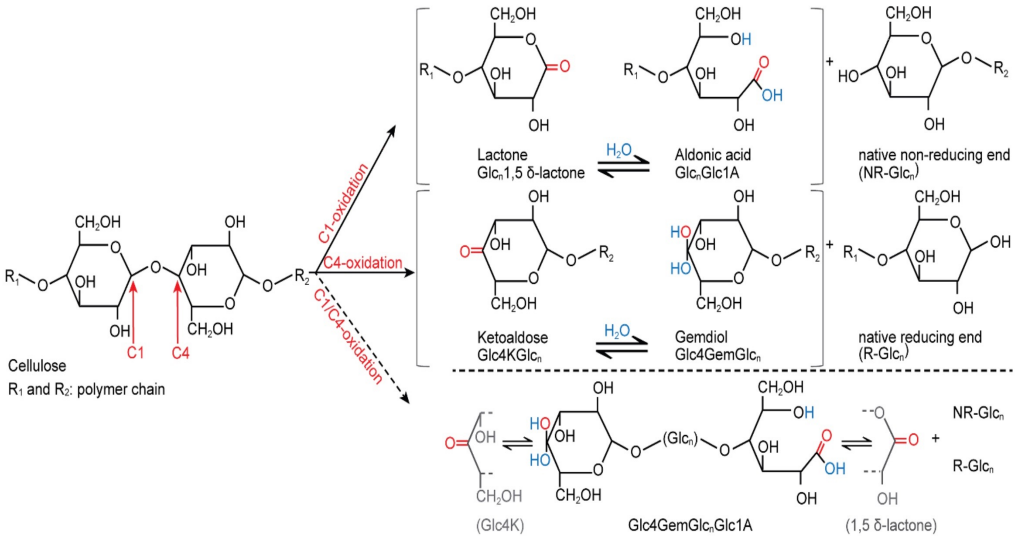


Figure 1.7. Oxidative regioselectivity in cellulose-active LPMOs. C1-oxidation (upper scheme) results in formation of lactones, which are hydrated to aldonic acids, and generates native non-reducing ends. C4- oxidation (middle scheme) leads to the formation of ketoaldoses and the corresponding hydrated gemdiols and generates native reducing ends. LPMOs with mixed C1/C4-oxidation can, in addition to the abovementioned compounds, also produce oligosaccharides that are oxidized at both ends (i.e. C1/C4 double-oxidized products). Figure taken from Forsberg et al. (2018).

1.4.6. HYDROGEN PEROXIDE AS A CO-SUBSTRATE

As exposed hereinbefore, the general belief of catalysis by LPMOs includes the requirement of molecular oxygen and a reductant that would deliver two electrons per catalytic cycle. In light of recent studies with the use of enzyme assays, mass spectrometry and experiments with labeled oxygen atoms, it has been shown that H_2O_2, rather than O_2, is the preferred co-substrate of LPMOs (Bissaro et al., 2017). Stable reaction kinetics can be achieved by controlling the H_2O_2 supply and it also allows the LPMOs to work in seemingly anaerobic conditions (i.e. in absence of O_2, but presence of H_2O_2). As a result, the reductant is only required in priming amounts to bring the LPMO in its active state from Cu(II) to Cu(I). Once activated, the LPMO-Cu(I) can carry out several catalytic cycles provided that H_2O_2 is available because H_2O_2 brings together the oxygen, electrons and protons equivalents ($H_2O_2 = O_2 + 2e^- + 2H^+$) necessary for a complete catalytic cycle. A significant increase in initial LPMO rates with 26-fold more oxidized products released was shown by a reaction with an AA10 in the presence of a reductant and low concentrations of exogenous H_2O_2. Prior to the discovery on H_2O_2 (Bissaro et al., 2017), it has been reported a

dramatic increase (up to 100-fold) for the reaction with an AA9 (fungal LPMO) in the presence of chlorophyllin, light and ascorbic acid (Cannella et al., 2016), and this increase in activity was proposed by Bissaro et al. (2017) to be due to an efficient generation of H_2O_2 by the photocatalytic system. Isotopic labeling was performed to verify the role of H_2O_2 as a co-substrate, with $H_2{}^{18}O_2$ showing that the oxygen introduced into the polysaccharide chain comes from H_2O_2 rather than O_2. Furthermore, reactions with lower concentrations of $H_2{}^{18}O_2$ showed that even in the presence of a 10-fold surplus of $^{16}O_2$, the oxidized products carried ^{18}O from $H_2{}^{18}O_2$ and not from $^{16}O_2$. Moreover, a competition experiment with peroxidase and LPMO showed a complete inhibition of LPMO activity by the peroxidase, despite the presence of O_2 and a reductant. To conclude, the experiments showed that H_2O_2 is the catalytically relevant co-substrate for LPMOs. These findings might explain why hitherto published catalytic rates for LPMOs are usually low and similar and independent of the LPMO or the substrate used (Agger et al., 2014; Frandsen et al., 2016; Vaaje-Kolstad et al., 2010). By one of several possible pathways, after reaction of the LPMO-Cu(I) with H_2O_2 it was proposed that a Cu(II)-oxyl intermideiate is formed, responsible of the hydrogen atom abstraction on the polysaccharide (Bertini et al., 2018; Bissaro et al., 2017; Kim et al., 2014; Wang et al., 2018). The subsequent steps lead to polysaccharide hydroxylation and bond cleavage as described above. Of note, in reaction settings containing H_2O_2, a quite low amount of ascorbic acid was also added, suggesting that the LPMO remains in the reduced state after the reaction. Also, levels of oxidized products are much higher than the total amount of ascorbic acid added, thus agreeing with the proposed mechanism in which a reduced LPMO can catalyze several reactions provided that H_2O_2 is supplied.

It was also observed that the LPMO reactions stopped very rapidly at a higher concentration of H_2O_2 (Bissaro et al., 2017). The inactivation was due to oxidative damage, especially in the active site, observed using proteomics technologies. Additionally, the inactivation was partly avoided by the presence of substrate. In light of the findings, that the modifications occurred very locally, it was concluded that oxidative damage is not caused by a random generation of hydroxyl radical derived from H_2O_2 in solution, but by *in situ* enzyme-generated hydroxyl radicals. The recruitment of two electrons and protons in automatically delivered with H_2O_2 as the co-substrate ($O_2 + 2e^- + H_2^+ = H_2O_2$).

Despite the light shed on the reaction mechanism of LPMOs by this research, it is still unclear how the active site coordinates the reactive oxygen species to perform regioselective oxidation.

1.4.7. CARBOHYDRATE BINDING MODULES

Carbohydrate binding modules (CBMs) are non-catalytic proteins often appended to catalytic domains allowing the later to get into close proximity with the substrate to be modified (Bolam et al., 1998; Boraston et al., 2004). It appears that these binding domains simply assist the enzyme catalytic domain in adsorption onto the substrate, as they have minimal effect on the catalytic mechanism (Várnai et al., 2013). More efficient degradation of the polysaccharide is a result of an increase of enzyme concentration on the substrate surface (Bolam et al., 1998). The CMBs are classified according to sequence similarities in the CAZy database, currently into 83 families.

CBM binding to cellulose are classified into family number 1, 2, 3, 4, 6, 9, 10, 16, 17, 30, 37, 44, 46, 49, 59, 63, 64 and 72, while those binding to chitin are found in families 5, 12, 14, 18, 19, 37, 50, 54, 55 and 73. There is a large variation in binding specificity of the CBMs to the extent that cellulose-binding CBMs seem optimized to bind different types and faces of cellulose (Blake et al., 2006). Furthermore, several studies indicate discrepancies between the binding preferences of CBMs and the substrate specificity of the catalytic domains they are appended too (Cuskin et al., 2012; Hervé et al., 2010). By binding to one polysaccharide type in the plant cell wall (which is of multi-polymeric nature), the CBM brings the catalytic domain in proximity of its substrate, which can be another type of polysaccharide (Hervé et al., 2010). Subsequently, it has been shown that the beneficial effect of CBMs on enzyme efficiency diminishes at high substrate concentration, which is important for practical applications for the biorefinery (Várnai et al., 2013). Therefore, in industrial settings that often employ high solid loadings, the role of CBMs is likely to be less important.

The ability of LPMOs to bind efficiently to various substrates is often promoted by the presence of one or more CBMs. For instance, by using a “module walking” approach, i.e. looking for LPMO-like sequences appended to starch-binding CBMs, researchers have been able to discover a new family of starch-active LPMOs (AA13 family) (Vu et al., 2014b). Although the majority of AA10s exist as single domain enzymes, several enzymes also contain one or more CBMs attached to the catalytic domain of LPMOs via a flexible linker (section 1.4.8.). This may reflect substrate

preferences (Horn et al., 2012), as the cellulose-binding CBM2 in the bi-modular *Streptomyces coelicolor* AA10, *Sc*AA10C (also known as CelS2), contributes to activity (Forsberg et al., 2011; Forsberg et al., 2014a). The loss of a CBM module from two AA10s has been shown to cause a reduction in enzyme activity towards Avicel and PASC [phosphoric acid swollen cellulose; (Arfi et al., 2014; Crouch et al., 2016; Forsberg et al., 2014a)].

The CBMs have been classified as three different types, based on their structure and ability to influence the function of associated catalytic domains (Boraston et al., 2004). CBMs type A interacts with the planar surface of crystalline polysaccharides (e.g. cellulose) through interactions between aromatic amino acid side chains of Trp, Tyr and Phe (Morag et al., 1995; Tormo et al., 1996) and the polysaccharide. Type B CBMs can bind polysaccharides found in amorphous regions of cellulose and hemicellulose, due to their open cleft. Type C CBMs are suggested to bind short soluble oligosaccharides (Boraston et al., 2004). Therefore, CBMs of different types can target an attached catalytic domain to a particular substrate (Cuskin et al., 2012; Gao et al., 2013; Hervé et al., 2010; Montanier et al., 2009).

1.4.8. THE LINKER

As explained above, many lignocellulose-degrading enzymes contain non-catalytic CBMs to facilitate a high concentration of enzymes at polymer surfaces (Boraston et al., 2004; Hervé et al., 2010; Lehtiö et al., 2003). These non-catalytic CBMs are connected to the catalytic domain by flexible linkers. It has long been thought that these linkers simply serve as a tether between structures domains or to act in an inchworm-like fashion during catalytic action (Payne et al., 2013). However, it is now generally believed that linker regions maximize substrate accessibility when the enzyme in bound to the plant cell wall via CBMs and thus display a great deal of structural flexibility (Pell et al., 2004).

Linkers have been reported to vary between 6 and 67 amino acids in length (Hansson et al., 2017) and are typically rich in proline and hydroxyl amino acids (e.g. serine and threonine). The proline and hydroxyl amino acid content differs among the linkers and they rarely share any apparent sequence homology.

Linkers can exert diverse functions and adopt various structures to fulfill the application of fusion proteins. Flexible linkers vary in length adjusted to different purposes and are generally rich in small or polar amino acids such as Gly and Ser.

This provides good flexibility and solubility, and although they do not have rigid structures, they can serve as passive linkers to keep a distance between functional domains. In contrast, rigid linkers exhibit relatively stiff structures by containing multiple Pro residues or by adopting α-helical structures. The length can easily be adjusted to achieve an optimal distance between domains and can, under many circumstances, separate the functional domains more efficiently than the flexible linkers. In contrast to the flexible and rigid linkers that do not allow the separation of joined proteins, cleavable linkers permit the release of free functional domain *in vivo* via reduction or proteolytic cleavage (Chen et al., 2013). Based on similarity to proline-rich linkers from other proteins, it has been suggested that cellulose linkers found in cellolose-active enzymes represent extended, flexible hinges between the two domains facilitating their independent function. The importance of the linker on the enzyme function is emphasized by their widespread occurrence (Srisodsuk et al., 1993). Interestingly, the linker has also been shown to increase cellulose affinity by absorbing to cellulose (Payne et al., 2013). While there is evidence of conserved function in linkers, biophysical analyses and sequence analyses have demonstrated that these regions are highly divergent and do not exhibit considerable structural elements (Abuja et al., 1988; Abuja et al., 1989; Crasto & Feng, 2000; Schmuck et al., 1986; Uversky & Dunker, 2012). Therefore, the challenge is to identify important features in linker regions and to understand the role these play in protein function (Sammond et al., 2012). In general, linkers have so far attracted very little attention. A very recent study has investigated the distinct roles of cellulases glycosylation on the catalytic domains relative to glycans found decorating intrinsically disordered linkers (e.g. for binding, activity). Overall, the need for deeper understanding of multidomain architectures in plant cell-wall-degrading enzymes is highlighted by these results (Amore et al., 2017). Of note, there is no report in the public literature on the effect of the composition or length of linkers on the activity of LPMOs.

1.5. THE ENZYMATIC MACHINERY OF *Cellvibrio japonicus*

The AA10 from *Cellvibrio japonicus* (*Cj*AA10A) has been the main model enzyme of the present master thesis project. Therefore, the next paragraphs will introduce the source microorganism itself and the different properties of *Cj*AA10A.

Cellvibrio japonicus, a Gram-negative soil bacterium, is primarily known for its ability to degrade plant cell wall polysaccharides. Even though the plant cell wall polysaccharide-degrading machinery of *C. japonicus* has been studied for decades, not much is known about the bacterium's ability to degrade chitin or about the enzymes involved. Notably, a type II secretion system (T2SS) secrete the majority of the biomass degrading enzymes in *C. japonicus*, and the bacterium shows limited growth on cellulose and chitin in the absence of this secretory system (Forsberg et al., 2016; Gardner, J. G. & Keating, 2010).

Characterization of the chitin-active LPMO *Cj*AA10A, a tri-modular enzyme containing a catalytic family AA10 LPMO module (called *Cj*AA10A^{cd}), a family 5 chitin-binding module and a C-terminal family 73 chitin-binding module, revealed that the active site of the enzyme combines structural features hitherto only observed in either cellulose or chitin active LPMO10s. The full-length LPMO has been reported to show no activity towards cellulose, but is able to bind and cleave both α- and β-chitin. In 2016 it was shown that this enzyme shares features with both chitin and cellulose-oxidizing AA10s (**Figure 1.8**) (Forsberg et al., 2016).

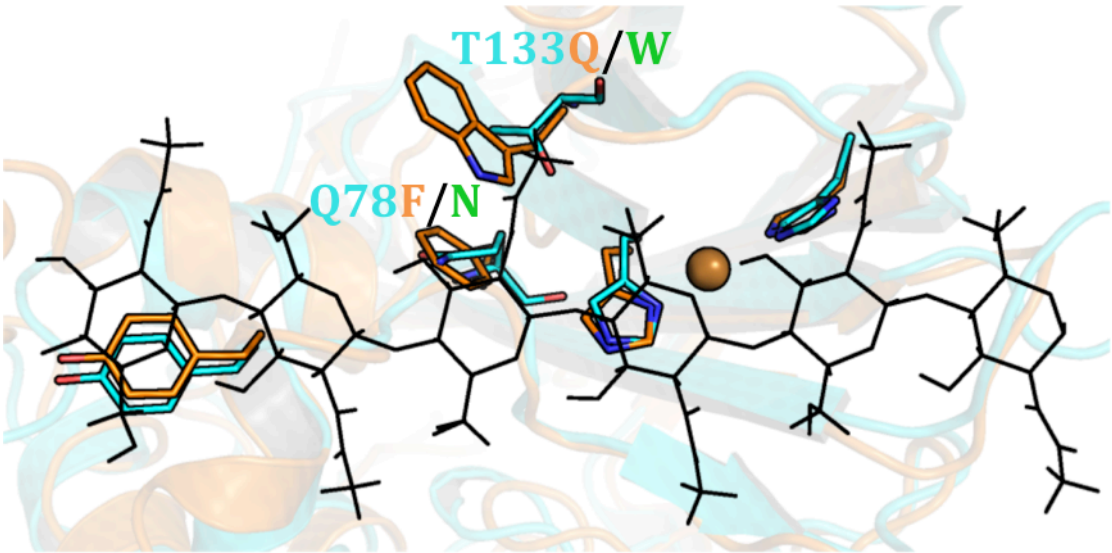


Figure 1.8. Superimposition of catalytic domains of *Cj*AA10A (blue) and *Sc*AA10C (orange). The figure shows how *Cj*AA10A shares features with cellulose-oxidizing AA10s (here shown with *Sc*AA10C). The putative positioning of chitin is modeled in black lines (as described in section 1.1.2.).

In the *Cj*AA10A^{cd} the planar face of the substrate binding surface protrusion is extended by one amino acid compared to the chitin-active LPMOs *Sm*AA10A (CBP21) and *Ef*AA10A. *Cj*AA10A also shows similarity to the cellulose-active *Sc*AA10C (CelS2), *Sc*AA10B and *Ba*AA10A. The active site of *Cj*AA10A show similarity to other AA10s, containing a copper ion coordinated by two histidines (His37 and His136) in a T-shaped histidine brace (Forsberg et al., 2016). *Cj*AA10A is an interesting LPMO, as it is a chitin-active enzyme with many features of cellulose-active LPMOs and constitutes an evolutionary intermediate.

1.6. AIM OF THIS STUDY

The prelude to this thesis was the characterization of the chitin-active *Cj*AA10A. In the family 10 of AAs, some LPMOs can be found to be active on chitin and others active on cellulose, or both. However, their structure is very similar and it has so far been impossible to attribute clearly the molecular determinant driving this substrate preference, with no clear-cut frontier (Forsberg et al., 2014b; Forsberg et al., 2016).

The evolution of proteins often occurs in a framework of functional constraints, leading to the co-evolution of several residues to allow a divergence of specificity (e.g. cellulose vs. chitin) while keeping the main biological function (e.g. oxidation). In this theoretical context, we aimed at identifying co-evolved residues by performing a correlated mutation analysis based on multiple sequence alignment gathering cellulose as well as chitin-active AA10 (in-house unpublished results). It appears that the chitin-active *Cj*AA10A constitute an evolutionary intermediate between the chitin-active *Sm*AA10A and the cellulose-active *Sc*AA10C, with many features of cellulose-active AA10s already in place. Therefore, *Cj*AA10A was selected as a starting point in the present project to get insights into the evolutionary divergence in substrate specificity (EDSS) amongst the AA10s to allow a better understanding of the molecular basis of LPMO catalysis, but also to improve the predictive capacity to link an LPMO sequence to its phenotype.

The first part of this study focused on the cloning, expression, purification and characterization of different versions of *Cj*AA10A, with shuffled linkers and CBM domains, in order to generate a platform enzyme suitable for molecular evolution of the catalytic domain (section 4.2.1.).

The second part of the study focused on generating and characterizing mutants of the aforementioned platform enzyme (section 4.2.3.). The binding and activity of the variants were analyzed by MALDI-TOF MS and HPLC in order to identify how the swapping of linker, CBMs and introduction of mutations affecting the enzyme properties (section 4.4. and 4.6.).

In addition, the variants were performed in a reaction assay without the substrate to quantify their H_2O_2 production ability (section 4.5.), as LPMOs are widely known to produce H_2O_2 when the substrate is not present (Kittl et al., 2012).

2. MATERIALS

2.1. LABORATORY EQUIPMENT

Table 2.1. Laboratory equipment.

EQUIPMENT	SPECIFICATION	SUPPLIER
Autoclave tape	12 mm 18 mm	VWR
Automatic pipettes	Finnpipette F2 Pipettes, single channel pipettes	Thermo Scientific
Balances	Sartorius basic	Sartorius
Biosafety cabinets	Av-100	TelStar
Blue-cap bottles	1000 mL, 500 mL, 250 mL, 100 mL, 50 mL	VWR
Centrifuge	AvantiTM J-25 Setrifuge – 5430R	Beckman Eppendorf
Centrifuge rotors	JA10	Beckman
Centrifuge tubes	500 mL 25 mL	Nalgene
Centrifuge filters for protein concentration	Amicon Ultra-15 10K	Merck Millipore
LP Chromatographic system	BioLogic LP system BioFrac Fraction Collector	Bio-Rad
Cryotubes	2 mL	Sarstedt
Culture flasks	2000 mL	Nalgene
Electrophoresis equipment	Vertical electrophoresis tank Power supply	VWR
Eppendorf tubes	1.5 mL – Axygen®	VWR
Falcon tubes	15 mL; 50 mL	
Filters	Steritop^{TM} 0.22 µm	Millipore
Filter plate, 96 well	0.45 µm hydrophilic, low protein binding	Merck-Millipore
Freezer (-18˚C)		Bosch Whirlpool

Freezer (-80˚C)	Ultra-Low	SANYO
Fume hood	Mc6	Waldner
Gel imager	Gel Doc EZ Imager	Bio-Rad
Glass equipment		Schot-Duran/VWR
Gold electrode, PAD detection	Disposable Electrode Dionex Carbohydrate Certified	Thermo Scientific
HiTrap DEAE FF HiLoad™ 16/600	5 mL	GE Healthcare
HPLC-columns	Dionex™ CarboPac™ PA1 Analytical and Guard column	Thermo Scientific
HPLC system	ICS 3000	Thermo Scientific
Ice maker	KF 145	PORKKA
Incubator		Termaks
Inoculation loops	1 µL	Sarstedt
Magnetic stirrer		VWR
Microbalance	Sartorius CP-2P	VWR
Multiskan™ FC Microplate Photometer		Thermo Scientific
Parafilm	5 cm	VWR
Pasteurpipettes	Plastic, 5 mL; 10 mL	VWR
PCR cooler		Eppendorf
PCR system	PCR Doppio SimpliAmp	VWR Thermo Scientific
PCR tubes	0.2 mL	VWR
Petri dish	9 cm	Heger
pH benchtop meter	FiveEasy Plus	Mettler Toledo Sentron
Pipette tips	Next Generation Tip Refill	VWR
Refrigerator (4˚C)		Whirlpool
Shaking incubator	Multitron HT	Infors
Size Exclusion Chromatography	HiLoad™ 16/60 Superdex™ 75 prep grade	GE Healthcare
Sonicator bath	Transonic 460/H	Elma
Spectrophotometer	AG Biophotometer	Eppendorf

	WPA CO8000 Cell Density Meter	Biochrom
Syringe filters	0,22 μm; 0,45 μm	Sarsted
Syringes	50 mL, 30 mL, 20 mL, 10 mL, 2mL, 1 mL	BD-plastipak
Table top centrifuge	Sigma 1-14	LABEX
Thermomixer	Comfort C	Eppendorf
Ultrasound bath	Transsonic 460/H	Elma
Uvettes	1 x 1 cm	Eppendorf
Vacuum manifold		Millipore
Vortex	Vortex-2 Genie MS 3 basic	Scientific Industries IKA
Water bath	Eco Temp TW12	Julabo
Weighing boats		VWR
Whirlmixer	Vortex-Genie 2 MS 3 Basic	Scientific Industries IKA
Qubit fluorometer		Life Technologies

2.2. SOFTWARE FOR ANALYSIS

Table 2.2. Software for analysis.

SOFTWARE	APPLICATION	SUPPLIER
ExPASy ProtParam tool	Calculations of pI, MW and ε	ExPASy
ImageLab	Gel visualization	BioRad
LP Data View	LP-chromatographic system	BioRad
Primer design tool	Primer design	Agilent Technologies
SkanIt	Multiskan™ FC Microplate Photometer	Thermo Scientific

2.3. CHEMICALS

Table 2.3. Chemicals.

CHEMICAL		SUPPLIER
Acetic acid	CH_3COOH	Sigma-Aldrich
Agar agar		Merck Millipore
Ampicillin sodium salt	$C_{16}H_{18}N_3NaO_4S$	Sigma-Aldrich
Amplex® Red reagent	$C_{14}H_{11}NO_4$	Thermo Scientific
Avicel® PH-101, cellulose microcrystalline		Sigma-Aldrich
Bacto™ Yeast Extract, granulated		Merck
Bacto™ Tryptone (peptone from casein)		Merck
Bis-Tris methane	$C_8H_{19}NO_5$	VWR
Copper (II) sulfate	$CuSO_4$	VWR
Dimethylsulfoxide (DMSO)	$(CH_3)SO$	
Distilled water, dH2O (Milli-Q quality)		Merck Millipore
Ion-free water, FLUKA TraceSELECT®		Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	$C_{10}H_{16}N_2O_8$	Sigma-Aldrich
Ethanol 96 %	C_2H_5OH	VWR
Glycerol 85 %	$C_3H_5O_3$	VWR
Hydrogen chloride	HCl	Sigma-Aldrich
Isopropanol	C_3H_8O	VWR
L-Ascorbic acid	$C_6H_8O_6$	Sigma-Aldrich
Magnesium chloride	$MgCl_2$	VWR
Potassium chloride	KCl	Merck
Potassium phosphate dibasic	K_2HPO_4	Sigma-Aldrich
Potassium phosphate monobasic	KH_2PO_4	Sigma-Aldrich
Protease inhibitor tablets		Sigma-Aldrich
SOC medium	2 % Tryptone, 0.5 % Yeast extract, 10 mM NaCl,	Life Technologies

	2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose	
Sodium acetate	$NaCH_3COOH$	Sigma-Aldrich
Sodium chloride	NaCl	Sigma-Aldrich
Sodium hydroxide	NaOH	VWR
Sodium phosphate dibasic	$HNa_2O_4P * H_2O$	Sigma-Aldrich
Sodium phosphate monobasic	$H_2NaO_4P * 2H_2O$	Sigma-Aldrich
Trizma® base	$C_4H_{11}NO_3$	Sigma-Aldrich

2.4. SELF-MADE MEDIA AND BUFFERS

Table 2.4. Self-made media and buffers.

MEDIA/BUFFER	CONTENT
LB media	10 g Tryptone 5 g Yeast 10 g NaCl Add dH_2O to 1L volume. Autoclave.
LB media with agar	10 g Tryptone 5 g Yeast 10 g NaCl 15 g Agar Add dH_2O to 1 L volume. Autoclave and cool down to ~50 ˚C, add 1 mL 100 mg/mL ampicillin in a biosafety cabinet. Poor the solution into petri dishes and let solidify. Store at 4 ˚C.
Sodium phosphate buffer 50 mM, pH 7.0	68.9 g NaH_2PO_4 88.9 g Na_2HPO_4 Add dH_2O to 200 mL, adjust pH, filtrate with 0.2 μm filter.
Spheroplast buffer	10 mL 1 M Tris-HCl, pH 8.0 17.1 g Sucrose 100 μL EDTA, pH 8.0

	Add dH_2O to 100 mL, filtrate with 0.2 μm filter.
Tris-HCl 1 M, pH 8.0	121.1 g Tris-Base Add dH_2O to 1000 mL, adjust pH, autoclave.
Tris-HCl 50 mM $(NH_4)_2SO_4$ 1 M, pH 8.0	50 mL 1 M Tris-HCl pH 7.0 132.1 g $(NH_4)_2SO_4$ Add dH_2O to 200 mL, adjust pH, filtrate with 0.2 μm filter.

2.5. KITS

Table 2.5. Kits.

KITS	CONTENT	SUPPLIER
Amplex® Red Hydrogen Peroxide/ Peroxidase Assay Kit	5 x reaction buffer Molecular Probes (0.25 M sodium phosphate, pH 7.4) (Thermo Scientific) Amplex® Red (154 μg) Dimethylsulfoxide (DMSO)	10 U Horseradish peroxidase (HRP)* Hydrogen Peroxide (H_2O_2)
E.Z.N.A.® Plasmid Mini Kit I, V (capped) spin protocol	**Solution I** (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 μg/mL RNaseA) **Solution II** (200 mM NaOH, 1% SDS) **Solution III** (4.2 M Guanidine-HCl, 0.9 M potassium acetate pH 4.8) **Equilibrium Buffer** (3 M NaOH) **HBC Buffer** (5 M Guanidine-HCl, 30% isopropanol) **DNA Wash Buffer** (10 mM Tris-HCl pH 7.5, 80% ethanol) **Elution Buffer** (10 mM Tris-HCl, pH 8.5) **HiBind™ DNA mini columns** **Collection tubes** (2 mL)	Omega BIO-TEK

In-fusion® HD Cloning Kit	5x In-Fusion HD Enzyme Premix	Clontech
iProof™ HF Master Mix Kit	iProof™ High-Fidelity Master Mix	Bio-Rad
Ni-NTA Purification System	**5X Native Purification Buffer** (250 mM NaH_2PO_4, pH 8.0, 2.5 M NaCl. Mix well and adjust pH to 8.0 with NaOH or HCl) **1X Native Purification Buffer** (80 mL of sterile distilled water, 20 mL of 5X Native Purification Buffer. Mix well and adjust pH to 8.0 with NaOH or HCl) **Native Binding Buffer** (30 mL of 1X Native Purification Buffer, 100 µL of 3 M Imidazole, pH 6.0. Mix well and adjust pH to 8.0 with NaOH or HCl) **Native Wash Buffer** (50 mL of 1X Native Purification Buffer, 335 µL of 3 M Imidazole, pH 6.0 Mix well and adjust pH to 8.0 with NaOH or HCl) **Native Elution Buffer** (13.75 mL of 1X Native Purification Buffer, 1.25 mL of 3 M Imidazole, pH 6.0. Mix well and adjust pH to 8.0 with NaOH or HCl)	Thermo Scientific
NucleoSpin® Gel and PCR Clean-up Kit	Kit Buffers NT1 and NT3	Thermo Scientific
QuikChange II Site-Directed Mutagenesis	PfuUltra High-Fidelity DNA polymerase (2.5 U/µl) [a]	Agilent Technologies

Kit	10x reaction buffer *DpnI* restriction enzyme (10 U/µl) [a] dNTP mix	

[a] "1 unit (U) is defined as the amount of enzyme that will form 1.0 mg purpurogallin from pyrogallol in 20 seconds at pH 6.0 and 20 ˚C" (Thermo Fisher Scientific, 2017).

2.6. LADDERS AND MIXTURES

Table 2.6. DNA/protein ladders and dNTP mixtures.

LADDER/MIXTURE	SUPPLIER
BenchMark™ Protein Ladder	Thermo Fisher
dNTP mix (Phusion® High-Fidelity PCR Kit)	NEB
dNTP mix (QuikChange II XL Site-Directed Mutagenesis Kit)	Agilent Technologies
GeneRuler 1 kb DNA Ladder	Thermo Fisher
GeneRuler 100 bp DNA Ladder	Thermo Fisher
NEBuffer 2 (10x)	NEB
PfuUltra HF DNA Polymerase	Agilent Technologies
Phusion® HF DNA Polymerase	NEB
Red Taq 2x DNA Polymerase Master Mix	VWR

2.7. CARBOHYDRATE SUBSTRATES

Table 2.7. Polysaccharide substrates used in binding or activity assays.

SUBSTRATE	SOURCE	SPECIFICATIONS	SUPPLIER
α-chitin	Shrimp shell	Dried and milled (~400 µm particle size, ash 1.7 %, 4.7 % moisture)	Sea garden
β-chitin	Squid pen	Dried and milled (~400 µm particle size)	France chitin, Marseille, France
Avicel® PH-101	Cellulose	~50 µm particle size	Sigma-Aldrich
PASC	Avicel® PH-101	Phosphoric swollen acid	Made in-house

2.8. BACTERIAL STRAINS

Table 2.8. Bacteria Strain.

BACTERIA STRAIN	SPECIFICATION	SUPPLIER
One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*	Chemical competent cells for protein expression	Life Technologies
One Shot® TOP10 Chemically Competent *E. coli*	Chemical competent cells for plasmid production	Life Technologies
XL1-Blue® Supercompetent Cells	Chemical competent cells for protein expression	Agilent Technologies

2.9. PROTEINS AND ENZYMES

Table 2.9. Produced LPMOs; wild type and variants.

ENZYME	ORIGIN OF CONSTRUCT	PRODUCTION
CBP21 (*Sm*AA10A-WT)	(Vaaje-Kolstad et al., 2005)	This study
CelS2 (*Sc*AA10C-WT = *Sc*AA10C^{cd}-*Sc*Linker-*Sc*CBM2)	(Forsberg et al., 2011)	This study
*Cj*AA10A-WT	(Forsberg et al., 2016)	This study
*Cj*AA10A^{cd} (construct 1)	(Forsberg et al., 2016)	This study
*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2)		This study
*Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3)		This study
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4)		This study
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5)		This study
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag - Q78F		This study
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag - Q78N		This study
CHB from *Serratia marcescens*	(Loose et al., 2014)	This study

2.10. PRIMERS FOR CLONING OF *Cj*AA10A VARIANTS

Table 2.10. Table of primers used for homologous recombination. Primers with overhangs containing the desired encoding sequences were designed. The overhang is found in the reverse primer dedicated to amplify the insert and in the forward primer dedicated to amplify the plasmid part. The overhang sequence is underlined. On each plasmid a forward and a reverse sequencing has been ordered with the primers used for the amplification PCR of the insert fragment (therefore the reverse primer is the same for both constructs).

PRIMERS	SEQUENCE (5'-3')
F-ins-1_*Sc*CBM2	AGTGTGGCATCCTCATCGTGTATGGCCGTCTATAGC
R-ins_*Sc*CBM2	AGCCGGATCAAGCTTTTACGGGGCAACACAACCGA
F-ins-2_*Sc*Linker	GACGTTGATTTCGGCGGCGGTAATGGTGAAGTTACG
F-plasmid	AAGCTTGATCCGGCTGCTAA
R-plasmid-1	TGAGGATGCCACACTAGAGG
R-plasmid-2	GCCGAAATCAACGTCAATGC

Table 2.11. Expected PCR amplification products.

Forward / Revers	F-ins-1_ScCBM2	F-ins-2_Sclinker	F-plasmid
R-ins_*Sc*CBM2	**I1** = [a]Rec1-*Sc*CBM2-Rec2 (336 bp)	**I2** = [c]Rec1'-*Sc*Linker-*Sc*CBM2-Rec2' (441 bp)	
R-plasmid-1			**P1** = [b]Rec2-Cj[cd]-*Cj*Linker-Rec1 (3500 bp)
R-plasmid-2			**P2** = [d]Rec2'-Cj[cd]-Rec1' (3400 bp)

[a]Rec1 in I1 contains DNA bases encoding for *Cj*Linker and is complementary to Rec1 in P1.
[b]Rec2 in I1 contains DNA bases encoding for pRSET-B and is complementary to Rec2 in P1.
[c]Rec1' in I2 contains DNA bases encoding for *Cj*[cd] and is complementary to Rec1' in P2.
[d]Rec2' in I2 contains DNA bases encoding for pRSET-B and is complementary to Rec2' in P2.

The couples of PCR products (I1-P1) and (I2-P2) were then used in separate homologous recombination reactions to yield constructs 3 and 4, see **Figure 4.3**.

2.11. HIS-TAG PRIMERS

Table 2.12. Primers used for His-Tag insertion. Primers with overhangs containing the His-Tag encoding sequence were designed. The overhang is found in the reverse primer dedicated to amplify the insert and in the forward primer dedicated to amplify the plasmid part. The overhang sequence is underlined and introduced base pairs are shown as boldface letters.

PRIMERS	SEQUENCE (5'-3')
F_plasmid_HisTag	**GCTCATCATCATCATCATCAC**TAAAAGCTTGATCCG GCTGCTAA
R-plasmid-2	GCCGAAATCAACGTCAATGC
F-ins-2_*Sc*Linker	GACGTTGATTTCGGCGGCGGTAATGGTGAAGTTACG
R_Ins_HisTag	TTA**GTGATGATGATGATGATGAGC**CGGGGCAACACA ACC

2.12. SEQUENCING PRIMERS

Table 2.13. Primers used for sequencing.

PRIMERS	SEQUENCE (5'-3')
Fo primer pRSET-B	GGTTGTGTTGCCCCGGCTCATCATCATCATCACTA AAAGCTTGATCCG
F-ins-2_Sclinker	GACGTTGATTTCGGCGGCGGTAATGGTGAAGTTACG
Rv primer pRSET-B	CGGATCAAGCTTTTAGTGATGATGATGATGATGAGCC GGGGCAACACAACC

2.13. MUTATIONS ON *Cj*AA10A

Table 2.14. Primers used for QuikChange II Site-Directed Mutagenesis to incorporate the desired mutation on the gene encoding *Cj*AA10A. The targeted codons that express the desired amino acids are underlined and the mutated base pair(s) is shown as boldface letters.

PRIMERS	SEQUENCE (5'-3')
*Cj*_Q78F_Fo	CGGTCTGTATACCCCG**TTC**GAAGTTGCAGTCGGCG
*Cj*_Q78F_Rv	CGCCGACTGCAACTTC**GAA**CGGGGTATACAGACCG
*Cj*_Q78N_Fo	CGGTCTGTATACCCCG**AAT**GAAGTTGCAGTCGGCG
*Cj*_Q78N_Rv	CGCCGACTGCAACTTC**ATT**CGGGGTATACAGACCG

*Cj*_T133Q_Fo	TTGGTTTTATGGGCCGCC**CAG**TTGGTCCAGACAAATTC
*Cj*_T133Q_Rv	GAATTTGTCTGGACCAA**CTG**GGCGGCCCATAAAACCAA
*Cj*_T133W_Fo	TTTGGTTTTATGGGCCGCC**TGG**TTGGTCCAGACAAATTCG
*Cj*_T133W_Rv	CGAATTTGTCTGGACCAA**CCA**GGCGGCCCATAAAACCAAA

2.14. BIOINFORMATICS

Table 2.15. Overview of physiochemical properties and domain structures.

PROTEIN	**Number of aa**	**Mw (Da)**	**pI**	**ε ($M^{-1}cm^{-1}$)**
CHB from *Serratia marcescens*	885	98548.25	6.30	
*Cj*AA10A-WT	361	38734.31	6.13	99810
*Cj*AA10A^{cd}	180	20344.53	6.23	49390
*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5	271	29794.59	6.30	72225
*Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2	312	33391.58	5.49	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2	316	34043.31	5.25	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag	323	34937.23	5.71	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F	323	34956.28	5.71	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N	323	34923.20	5.71	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-T133Q	323		5.71	72880
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-T133W	323		5.71	72880

3. METHODS

3.1. GENERAL METHODS

3.1.1. PRIMER PREPARATION

Materials:

- Primers (**Table 2.10**, **Table 2.12**, **Table 2.13** and **Table 2.14**)

Method:

The primers came as powder and were dissolved in dH_2O to a concentration of 100 µM, vortexed and kept on ice. A dilution was made to reach 10 µM where 5 µL from each sample and 45 µL dH_2O was mixed in new Eppendorf tubes. All primers solutions were stored at -20 ˚C after use.

3.1.2. PREPARATION OF GLYCEROL STOCKS

Glycerol stocks were made of the bacterial cells for saving at – 80 ˚C.

Materials:

- Cell culture (section 3.4.3.)
- Glycerol (85 %)

Method:

1 mL cell culture and 400 µL 85 % glycerol were mixed well in a CryoTube. The glycerol stocks were kept at – 80 ˚C.

3.1.3. QUICK START™ BRADFORD PROTEIN ASSAY

The protein concentration was determined according to the Quick Start™ Bradford Protein Assay (Bradford, 1976) with a spectrophotometer and a cuvette. This method can be used to determine the protein concentrations in the range 1.25 – 10 ng/µL.

Materials:

- Eppendorf BioPhotometer
- Protein Assay Dye Reagent Concentrate (Bio-Rad)
- Polystyrene cuvettes, 1.5 ml (Brand)
- Storage buffer (**Table 2.4**)
- Protein of interest

Method:

A dilution of each sample was made by transferring 5 µL of isolated plasmid to 795 µL storage buffer. A blank sample (i.e. 800 µL of storage buffer) was prepared for each set of experiment. Each of the (diluted) samples (800 µL) were treated with 200 µL Bradford solution, vortexed and incubated at room temperature for 5 minutes. Then, the absorbance was read at 595 nm (A_{595}), which is the characteristic wavelength for protein-bound dye molecules. The measured absorbance was converted into concentration using a standard curve made with BSA, or with a reference solution of the protein under study, and prepared in the same conditions (buffer, pH) as unknown samples.

3.1.4. DIRECT PHOTOMETRIC MEASUREMENT OF PROTEIN AND DNA

Alternatively to the Bradford Protein Assay method (see above), the protein concentration was also determined by reading the absorbance at 280 nm (A_{280}). Beer-Lamberts law relates the absorption of monochromatic light to concentration, the length of the light pathway through the solution and the extinction coefficient, and can be used to determine the concentration of purified protein.

Equation: $A_\lambda = \varepsilon_\lambda lc$

A_λ is absorbance at wavelength λ, ε_λ is the extinction coefficient (**Table 2.15**) at wavelength λ, l is the light path through the solution and c is the concentration. DNA concentrations were also determined by direct photometric measurement using the absorbance at 260 nm (A_{260}).

Materials:

- Eppendorf BioPhotometer
- Eppendorf® UVette®
- Storage buffer (**Table 2.4**)
- Isolated DNA (section 3.1.8.)

Method:

The absorbance at 280 nm was measured by making a 10-fold dilution of the samples to a total volume of 70 µL. The samples were transferred to an UVette and measured in a spectrophotometer.

3.1.5. AGAROSE GEL ELECTROPHORESIS

Materials:

- Mini-Sub® Cell GT cell
- Agarose gel
  - SeaKem® LE Agarose
  - 1 x TAE buffer
  - PeqGREEN DNA/RNA dye
- GeneRuler™ 1 kb DNA ladder
- GeneRuler™ 100 bp DNA ladder
- DNA samples
  - UView™ 6x Loading Dye (Bio-Rad)

Method:

The gel is prepared with an agarose concentration appropriate for the size of DNA fragments to be separated. Here, 2% and 0.8 % agarose gels were prepared to separate DNA fragments in the range 3000-4000 bp or 300-500 bp, respectively. Agarose was weight out to desired concentration and dissolve in 50 mL 1x TAE buffer by heating the mix in a microwave oven. The solution was cooled down to approximately 60 ˚C before a volume of 2.5 µL PeqGREEN solution was added. This solution facilitates visualization of DNA fragments during the run and for UV detection. The agarose was poured in a gel chamber with a comb and left to solidify for 30 minutes. After drying, the comb was removed and the gel was placed in a Mini-Sub Cell GT cell and submerged in 1x TAE buffer. The samples were mixed with 6x Loading Dye (1x final) and applied on the gel, alongside a ladder with known DNA fragment sizes to estimate the sample DNA size. Finally, the gel was run at 90 V for 40 minutes using a power supply and a Benchtop UV Transilluminator revealed the DNA bands. Gel imaging was performed using Gel Doc™ EZ Imager (Bio-Rad) with a UV Sample Tray.

3.1.6. DNA ISOLATION BY GEL EXTRACTION

By running a (complex) DNA sample on gel electrophoresis one can isolate and purify DNA fragments based on their size (often after a PCR for cloning purposes).

Materials:

- Agarose gel with DNA bands
- UV transilluminator
- NucleoSpin® Gel and PCR Clean-up Kit (**Table 2.5**)

Method:

After running a standard agarose gel electrophoresis, as described in section 3.1.5., the position of the desired DNA bands was identified by short exposure to UV light and the bands were then cut out of the agarose gel and placed in a pre-weighted falcon tube. The weight of the bands cut from the gel determines the volume of buffer to add. For bands extracted from < 2% agarose gels 200 µL of NT1 buffer is added/100 mg of gel. For bands extracted from ≥ 2% agarose gels 400 µL buffer/100 mg of gel is added. The tubes with the gel and the added buffer were incubated for 5-10 minutes in a pre-heated water bath at 50 ˚C until the gel was completely dissolved. The tubes were regularly vortexed to make sure all the gel was dissolved. The samples were transferred to collection tubes with DNA-binding columns and centrifuged for 30 sec at 11 000 x g. The flow-through was discarded. Then 700 µL buffer NT3 was added to the columns and a new centrifuge step followed, 11 000 x g for 30 sec. The flow-through was discarded and the step was repeated. To remove all NT3 buffer, the tubes were centrifuged once more for 1 min at 11 000 x g and then incubated for 1 min at room temperature. Finally, the silica membranes were transferred to new, clean 1.5 mL Eppendorf tubes and 30 µL dH_2O was added straight to the filter. After 1 min incubation at room temperature, the DNA was eluted from the column by centrifugation for 1 min at 11 000 x g. The samples were stored at – 20 ˚C before further use.

3.1.7. BASIC TRANSFORMATION PROCEDURE

Materials:

- One Shot® TOP10 Chemically Competent *E. coli*
- XL1-Blue® Supercompetent Cells
- One Shot® BL21 Star™ (DE3) Chemically Competent Cells
- SOC-media
- LB agar plates with ampicillin (100 mg/mL)
- Protein-encoding plasmids

Method:

To generate plasmid DNA or to express proteins, the protein-coding plasmids were transformed into cells suited for each purpose, respectively One Shot® TOP10 Chemically Competent *E.coli*-cells or XL1-Blue® super competent cells and One Shot® BL21 Star™ (DE3) Chemically Competent cells. After thawing the competent cells (50 µL) on ice, 2 µL of plasmid (ca. 100 ng) was added and further incubated on ice for 30 min. The cells were then heat-shocked in a preheated water bath at 42 ˚C for 30 seconds before being set on ice for 10 additional minutes. SOC media (5 equivalent volumes of the cells; i.e. 250 µL) was added to each sample before incubation at 37 ˚C at 200 rpm for one hour. Finally, the transformation mixture (250 µL or appropriate dilution) was plated on LB media-agar plates containing ampicillin (100 µg/mL final). The plates were incubated at 37 ˚C over night.

3.1.8. PLASMID ISOLATION

Plasmid incorporated by transformants issued from transformation was extracted by inoculating with a single colony 5 mL LB-ampicillin liquid medium into a 15 mL Falcon tube. This mini-culture was incubated at 37 ˚C during ca. 18 h before plasmid DNA extraction and sequencing (section 3.1.9.). In the case of transformants issued from site-directed mutagenesis experiments, at least 3 colonies were selected to inoculate separate mini-cultures. To isolate plasmid DNA from cell mini-cultures, the E.Z.N.A.® Plasmid DNA Mini Kit I was used. This kit allows extraction of plasmid DNA from bacterial cells by alkaline-SDS lysis. By using HiBind® DNA Mini Columns, the purification of plasmid is facilitated in three steps, binding, washing and elution. This enables multiple samples to be processed simultaneously (OMEGA BIO-TEK, 2018).

Materials:

- E.Z.N.A.® Plasmid DNA Mini Kit I Spin Protocol
- Cell culture

Method:

5 mL cell culture was centrifuged in 15 mL Falcon tubes at 5000 rpm for 10 minute at 4 ˚C. The supernatant was removed and the cell pellet was resuspended in 250 µL solution I/RNase A. The samples were transferred to a 2 mL Eppendorf tube, 250 µL solution II was added and the samples were inverted several times to obtain a clear

lysate. After 2-3 minutes 350 µL solution III was added. The samples were immediately inverted several times until a flocculent white precipitate was formed. The solutions were centrifuged for 10 minutes at 13 000 x g. At the same time HiBind® DNA Mini Columns were equilibrated by centrifuging 100 µL 3M NaOH for 1 minute at 13 000 x g. The cleared supernatants were transferred to the equilibrated columns and centrifuged at 13 000 x g for 1 minute. In the next step, 500 µL HBC buffer was added and the columns were centrifuged again at 13 000 x g for 1 minute. For the washing step 700 µL DNA wash buffer was added and the columns were centrifuged once again at 13 000 x g for 1 minute. The empty columns were centrifuged at 13 000 x g for 2 minutes to dry the column matrix and remove residual ethanol. The HiBind® columns were transferred to a clean 2 mL Eppendorf tube before 50 µL sterile deionized water was added directly to the center of the column membrane to elute the plasmid. After 1 minute the columns were centrifuged at 13 000 x g for 1 minute. The concentration of isolated plasmids was then determined as described in section 3.1.4. The isolated plasmid was kept at – 20 ˚C.

3.1.9. SEQUENCING

To verify that the desired mutations have been incorporated in the transformants, the isolated plasmids were sent for sequencing at GATC Biotech (Konstanz, Germany). The transformants were cultured and isolation of the mutated plasmids was performed with the E.Z.N.A.® Plasmid DNA Mini Kit I Spin Protocol (section 3.1.8).

Materials:

- DNA
- Sequencing primers (**Table 2.13**)
- dH_2O

Method:

In a 1.5 mL Eppendorf tube, DNA (400 ng), primer (forward or reverse, 10 µM) and dH_2O were mixed to a total volume of 11 µL. The tubes were labeled and sent for sequencing.

3.2. GENE CLONING

Polymerase chain reaction (PCR) is a technique used for amplifying DNA in vitro. With the help of a DNA polymerase, new strands of DNA complimentary to the template strand can be synthesized. The PCR follows three steps: denaturation, hybridization and extension. In the first step, denaturation, the sample is warmed up to 98 ˚C. At this temperature, the template is denatured to single-stranded DNA. In the second step, the temperature is lowered to an optimal individual annealing temperature to let the primers be able to associate with the vector. Lastly, in the extension step, the temperature arises to a temperature that is optimal for the polymerase in use. This three-step cycle is repeated several times over to synthesize large amounts of DNA. The result of the PCR is the synthesis of new DNA either bearing the desired mutation (for SDM) or representing a portion of the template DNA (for DNA fragment amplification and recombination purposes).

3.2.1. SITE-DIRECTED MUTAGENESIS

To create site-specific mutations such as insertions, deletions and point mutations in the double-stranded vector, site-directed mutagenesis (SDM) is utilized.

One forward and one reverse primer, each containing the desired mutations, are required for the mutagenesis. Each of the oligonucleotide primers, complementary to the opposite strands of the vectors, are extended during temperature cycling by DNA Polymerase called Phusion® (Hill & Stewart, 1992). The designed primers are 25-45 base pairs long, with the point mutation in the center and the ends need to be rich with GC-nucleotides to stabilize the binding with the vector. The melting temperature can be calculated with a formula given by the instruction manual to Agilent Technologies. The mutagenic primers (**Table 2.10**) were designed by Bastien Bissaro and synthesized by Invitrogen (Thermo Fisher Scientific).

Materials:

- QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) (**Table 2.5**)
- DNA template
- Primers (**Table 2.10**)

Method:

The sample reactions were made by solutions provided by the QuikChange II XL Site-Directed Mutagenesis Kit. The samples were prepared in PCR tubes on a cooler and *PfuUltra* DNA polymerase was added right before running the PCR program, shown in **Table 3.2**. For the PCR cycle (**Table 3.1**) two separate PCR reactions were prepared, one with forward primer and one with reverse primer. After completed PCR reaction, 1 µL *DnpI* was added and the samples were incubated at 37 ˚C for one hour. *DpnI* is an endonuclease that will recognize hemi-methylated and methylated DNA and degrade this to remove non-mutated DNA (Agilent Technologies, 2018).

Table 3.1. Content of sample reaction, first cycle.

REAGENT	QUANTITY
10x reaction buffer	5 µL
dsDNA template	X µL (50 ng)
Forward primer	1.5 µL (10 µM) [a]
Reverse primer	1.5 µL (10 µM) [a]
dNTP mix	1 µL
dH_2O	till total volume of 49 µL
PfuUltra DNA Polymerase	1 µL

[a] Replace with water where primer is not used.

Table 3.2. Cycling Parameters, first cycle.

STEP	REPETITIONS	TEMPERATURE	LENGTH
Heating	x 1	110 ˚C	
Initialization	x 1	95 ˚C	30 sec
Denaturing	x 5	95 ˚C	30 sec
Annealing		55/60/65 ˚C	1 min
Extension		68 ˚C	X min [a]
Final hold	x 1	10 ˚C	

[a] The time length for extension is 1 minute per kilobase of plasmid length.

3.2.2. DOMAIN SHUFFLING

In order to transfer the linker and CBM2 domain from *Sc*AA10C to *Cj*AA10A a strategy was developed consisting in amplifying plasmid and insert using overhang primers (see section 3.2.3.). The resulting amplified DNA fragments (**Table 2.11**) were purified by gel extraction (section 3.1.6.) and used in homologous recombination reaction (section 3.2.4.).

3.2.3. "OVERHANG" PCR

Materials:

- Q5 High-Fidelity 2x Master Mix
- Template DNA
- Primers (**Table 2.10**)

Method:

The sample reactions were prepared in PCR tubes on a cooler with Q5 High-Fidelity 2x Master Mix, template DNA and both forward and reverse primers. Water was added to adjust the volume to a total of 50 µL. See **Table 3.3** for content of sample reactions.

For the annealing step, NEB calculator (http://tmcalculator.neb.com/#!/main) was used to calculate the temperature (T_m). The expected length of the plasmids is used to calculate the time of the extension step (**Table 3.5**). Each kilobase (kb) of plasmid length requires 1 minute at 68 ˚C per cycle. The cycling parameters are listed in **Table 3.4**. The samples were kept on -18 ˚C after completed PCR.

Table 3.3. Reaction setup for plasmid and insert cloning by PCR.

REAGENT	QUANTITY
Q5 High-Fidelity 2x Master Mix	25 µL
Forward Primer (10 µM) [a]	2.5 µL
Reverse Primer (10 µM) [b]	2.5 µL
Template DNA [c]	1 µL
Nuclease-Free Water	19 µL

[a, b] See **Table 2.11** for the combination of primers.

[c] *Cj*AA10A full-length for P1 and P2, *Sc*AA10C full-length for I1 and I2.

Table 3.4. Cycling Parameters.

STEP	REPETITIONS	TEMPERATURE	LENGTH
Initial denaturation	x 1	98 ˚C	30 sec
Denaturation	x 35	98 ˚C	10 sec
Annealing		X ˚C [b]	30 sec
Extension		72 ˚C	X min [a]
Final extension	x 1	72 ˚C	2 min
Final hold		4 ˚C	∞

[a] The time length for extension is 1 minute per kilobase of plasmid length.
[b] The temperature (T_m) for the annealing is calculated with regard to the expected bp length.

Table 3.5. Time length for extension and temperature for annealing.

PLASMID/INSERT	T_m ANNEALING STEP	LENGTH EXTENSION STEP
P1	63 ˚C	1 min 30 sec
P2	66 ˚C	1 min 30 sec
I1	68 ˚C	20 sec
I2	66 ˚C	20 sec

3.2.4. HOMOLOGOUS RECOMBINATION

Materials:

- Plasmid (P1 or P2) and insert (I1 or I2) (section 3.2.3.)
- In-fusion® HD Cloning kit (**Table 2.5**)
  - Enzyme Premix (5x)
- dH_2O

Method:

Plasmid and insert (**Table 2.11**) were combined in a 2:1 molar ratio by mixing them in PCR-tubes with 2 µL Enzyme Premix (5x) and dH_2O to a total volume of 10 µL (**Table 3.6**). The mix was incubated at 50 ˚C for 15 minutes. The mixture was diluted in 90 µL Tris-EDTA pH 8.0 and incubated on ice for 15 minutes. Then, 5 µL of the mixture was transformed into One Shot® TOP10 Chemically Competent *E. coli* as described in section 3.1.3.

Table 3.6. Reaction setup for homologues recombination of plasmid[a] and insert[b].

REAGENT	QUANTITY
5x In-Fusion HD Enzyme Premix	2 µL (1x)
Purified PCR fragments, plasmid	1.9 µL (100 ng) [a] 1.2 µL (100 ng) [b]
Purified PCR fragments, insert	0.7 µL (50 ng) [a] 0.8 µL (50 ng) [b]
dH_2O	5.4 µL [a] 6 µL [b]

3.3. HIS-TAG INSERTION ON PLATFORM ENZYME *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2

The engineered platform enzyme consists of the *Cj*AA10A^{cd} and swapped Linker and CBM with *Sc*AA10C yielding a final construct: *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2. This enzyme is called a platform enzyme because it will be used for site-directed mutagenesis. For a more time saving purification of these mutants, a 6x poly-histidine tag was added (see below).

A poly-histidine tag (6 x His-Tag) was fused to the C-terminus of the platform enzyme (i.e. structurally located on top of the CBM domain) to facilitate purification of this enzyme. The "overhang" PCR strategy (section 3.3.1) led to the generation of "Insert 3" made of *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2 (construct 4) with forward and reverse insert primers (**Table 2.12**) and "Plasmid 3" made of pRSET-B with His-Tag overhang and forward and reverse plasmid primers (**Table 2.12**). Both fragments were then assembled by homologous recombination (section 3.3.2.) to yield to the final gene product encoding for "*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2-HisTag" (construct 5).

3.3.1. "OVERHANG" PCR

Materials:

- iProof^{TM} High-Fidelity Master Mix (**Table 2.5**)
- DNA template: *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2 (construct 4)
- Primers (**Table 2.12**)

Method:

For the His-Tag insertion on the platform enzyme, PCR mixtures were prepared with iProof^{TM} HF Mix Kit, a DNA template and both forward and reverse primers. Water was added to adjust the volume to a total of 50 µL, see **Table 3.7** for content of sample reactions.

The expected length of the plasmids is used to calculate the time of the extension step. Each kilobase (kb) of plasmid length requires 1 minute at 68 ˚C per cycle. The plasmid and insert are 3500 bp and 447 bp long, respectively, and the extension step was therefore set to 2 min and 20 sec, respectively. The cycling parameters are listed in **Table 3.8**.

Table 3.7. Content of sample reaction.

REAGENT	QUANTITY
iProof™ HF Master Mix	25 µL
DNA template	X µL (50 ng)
Forward primer	2.5 µL (10 µM)
Reverse primer	2.5 µL (10 µM)
dH_2O	Till total volume of 50 µL

Table 3.8. Cycling Parameters.

STEP	REPETITIONS	TEMPERATURE	LENGTH
Initial denaturation	x 1	98 ˚C	1 min
Denaturation	x 35	98 ˚C	10 sec
Annealing		X ˚C [b]	30 sec
Extension		72 ˚C	X min [a]
Final extension	x 1	72 ˚C	5 min
Final hold		4 ˚C	∞

[a] The time length for extension is 1 minute per kilobase of plasmid length.
[b] The temperature (T_m) for the annealing is calculated with regard to the expected bp length.

3.3.2. HOMOLOGUES RECOMBINATION

After completed PCR reaction, the samples were run in an agarose gel electrophoresis, as described in section 3.1.5., then extracted and purified as described in section 3.1.6. The concentration of the extracted DNA was measured (section 3.1.4.) before the plasmid and the insert was combined by homologous recombination. After, 1 µL *DnpI* was added and the samples were incubated at 37 ˚C for one hour. Finally, the products were transformed in One Shot® TOP10 Chemically Competent *E. coli* (section 3.1.7.) and plated on agar plates with ampicillin.

After incubation at 37 ˚C overnight, three clones from each plate were picked up and transferred to 15 mL Eppendorf tubes with 5 mL LB and 5 µL 100 mg/mL ampicillin. The tubes were incubated at 37 ˚C, 160 rpm overnight. The plasmid and insert were then isolated (section 3.1.8.) and the concentration was determined (section 3.1.4.), before the samples were sent for sequencing (see section 3.1.9.).

3.4. EXPRESSION AND PURIFICATION OF ENZYMES

3.4.1. CULTIVATION OF CHITOBIASE

The GH20 β-N-acetylhexosaminidase, also called chitobiase (CHB), from *Serratia marscecens* (*Sm*GH20A, Genebank ID: L43594) was produced and purified according to previously described protocol (Loose et al., 2014).

Materials:

- LB medium (**Table 2.4**)
- Kanamycin (50 mg/mL)
- Glycerol stock of BL21 DE3 transformed with pET-30Xa/LIC-*chb*
- IPTG (0.1 M)

Method:

In a 2 L Erlenmeyer flask, 1 L LB medium and 1 mL kanamycin (50 mg/mL) were mixed and inoculated with a few µL from the glycerol stock. The flask was incubated at 37 ˚C on vigorous shaking (200 rpm), and OD was measured every 2nd hour until it reached 0.4 – 0.6. When the desired OD was reached, 0.1 mM IPTG was added to induce the expression and the culture was further incubated at 30 ˚C for 3 hours before harvesting the cells (see section 3.4.2.).

3.4.2. CYTOPLASMIC EXTRACTION BY SONICATION

Materials:

- 200 mM Tris-HCl pH 8.0
- 20 mM Tris-HCl pH 8.0, 20 mM Imidazole
- Chitobiase cultures (section 3.4.1.)

Method:

After 3 hours of induction with IPTG, the cultures were harvested by centrifugation at 5500 rpm for 12 min. The supernatant was discarded and the pellet was resuspended in 10 mL 200 mM Tris-HCl pH 8.0. The cultures were centrifuged again at 5500 rpm for 12 min and the supernatant was discarded. The pellet was resuspended in lysis/binding buffer, 20 mL 20 mM Tris-HCl pH 8.0 with 20 mM Imidazole followed by disruption by sonication. The *Vibra cell Ultrasonic Processor (Sonics)* was used for the sonication with a cycle of 1 sec sonication and 3 sec pause for 8 min using 30 % amplitude. The samples were kept on ice throughout the sonication. The samples

were then centrifuged at 10 304 g for 30 min at 4 ˚C to remove cell debris, and the supernatant was filtered using a 0.2 µm syringe filter and stored at 4 ˚C before His-Tag purification (see section 3.4.9.).

3.4.3. CULTIVATION OF LPMOs

Materials:

- LB medium (**Table 2.4**)
- 100 mg/mL ampicillin
- Protein to cultivate

Method:

Inside biosafety cabinets, 500 mL LB-media (containing 100 µg/mL ampicillin) in 2 L Erlenmeyer flask was inoculated with a single colony picked up from a plate of fresh BL21 (DE3) *E. coli* transformants bearing the plasmid encoding for the LPMO of interest. The flasks were incubated overnight at 37 ˚C with vigorous shaking (160 rpm) before recovering the expressed protein by periplasmic extraction (see section 3.4.4.).

3.4.4. PERIPLASMIC EXTRACTION

E. coli is a gram-negative bacterium and has an area between the outer cell membrane and the peptidoglycan layer called periplasm. The sequence representing the catalytic domain of LPMO is always preceded by a signal peptide that ensures addressing of the protein to the periplasm of the cell. During translocation of the protein into the periplasmic space the signal peptide is cleaved off leading to a mature protein. The N-terminal residue of LPMO is always a histidine and defines in part the catalytic center (copper coordination). Therefore, correct processing of the signal peptide is crucial to obtain a functional LPMO. The proteins are released from the periplasm by the cold osmotic shock method (Manoil & Beckwith, 1986). For this, a cold spheroplast buffer (which contains a high concentration of sucrose) is added to the harvested cells. This removes the outer layer; the cells take up the sucrose and are then treated with cold dH_2O afterwards forcing the cells to experience osmotic shock and lyse.

Materials:

- Cell culture (section 3.4.3.)

- Spheroplast buffer (**Table 2.4**)
- $MgCl_2$ (20 mM)
- Cold dH_2O

Method:

After incubation of cell culture overnight (section 3.4.3.), the culture were distributed in centrifuge bottles and centrifuged at 5500 g, 4 ˚C for 12 min. After centrifugation, the supernatant was discarded and the bottles with the cell pellets were kept on ice. While kept on ice, the cell pellets were resuspended in 30 mL spheroplast buffer and the content was gathered in two centrifuge bottles. After incubation on ice for a few minutes, the bottles were centrifuged at 5500 g, 4 ˚C for 12 min. The supernatant was discarded and the pellet was placed in a heating cabinet, 30 ˚C, to reach room temperature. After 10 min, the pellets were resuspended in 35 mL cold dH_2O and 350 µL 20 mM $MgCl_2$. The samples were kept on ice for 10-15 min before new centrifugation, 11325 g for 20 min. The supernatants were filtered using 0.2 µm syringe filters and kept at 4 ˚C until purification (section 3.4.5., 3.4.8. and 3.4.9.).

3.4.5. ANION EXCHANGE CHROMATOGRAPHY

Materials:

- Protein solution (section 3.4.4.)
- HiTrap™ DEAE FF, 5 mL bed volume
- Buffer A: 50 mM Tris-HCl pH 8.5
- Buffer B: 50 mM Tris-HCl pH 8.5, 0.5 M NaCl
- 20% ethanol

Method:

The first step to separate and purify LPMOs was done by anion exchange chromatography (AEC) coupled to a UV detector. The column used was a HiTrap™ DEAE FF column and is a weak anion exchanger with positively charged tertiary ammonia groups (diethylaminoethyl, DEAE) (GE Healthcare, 2018). The column was operating with a flow rate of 3 mL/min in 50 mM Tris-HCl pH 8.5 (Buffer A). Negatively charged proteins (including the LPMO) bind to the columns positively charged ammonia groups. When increasing the ion concentration with NaCl (Buffer B) the proteins are forced to elute.

Firstly, the column (stored in 20 % ethanol) was washed with dH_2O (2 column

volumes (CV); 10 mL) and then equilibrated with 100 % buffer A with 3 column volumes (15 mL). The eluted proteins will be shown as peaks in the chromatogram as they contain aromatic amino acids that are detected by an UV detector. When the UV-baseline and conductivity had been stable for more than one column volume (at least 5 mL), the sample, filtrated and equilibrated with at the same concentration as buffer A, was loaded (1.5 mL/min flow rate). Proteins that do not bind to the column will elute directly and be shown as peaks in the chromatogram as they contain aromatic amino acids that are detected by an UV detector. Bound proteins are eluted by applying a gradient of buffer B from 0 to 50 % over 100 min at 3 mL/min flow rate (60 CV). Proteins with weak negatively charged surface will elute first, and the proteins with an increase in negatively charged surfaces will elute with an increase of the gradient. The eluted proteins were collected in fractions of 3 mL. At the end of the purification, the column was washed by two cycles of (2 CV of 100 % buffer B; 2 CV of 100 % buffer A) and then dH_2O (2 CVs) and 20 % ethanol (2 CVs). The column was stored at 4 ˚C and the collected fractions were analyzed on SDS-PAGE gel electrophoresis (section 3.4.6.) to pool the fractions containing the desired protein. Before the next purification step, size exclusion chromatography (section 3.4.8.), the samples were up-concentrated as described in section 3.4.7.

3.4.6. SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS

SDS-PAGE is a method used to separate proteins in a solution after size. The size and the purity of the solution can be determined.

Materials:

- Cell pellet and fractions from purification (section 3.4.4. and 3.4.5.)
- SDS-PAGE working solution
  - 5 µL LDS sample buffer (4x)
  - 2 µL sample reducing agent (10x)
  - 3 µL dH_2O
- Any kD™ Mini-PROTEAN® TGX Stain-Free™ Protein Gel, 15 wells
- BenchMark™ Protein Ladder
- 1 x TAE (Tris-Acetate-EDTA) Running Buffer

Method:

10 µL of the purified fractions were mixed with 10 µL of the SDS-PAGE working solution. The samples were then heated up to 98 ˚C for 10 min in a water bath to denature the proteins. This treatment unfolds the proteins completely, and thus the proteins have the same charge and shape. The negatively charged proteins will migrate towards the positively pol in the electric field. The gel was placed in a gel electrophoresis chamber and the chamber was filled with 1 x TAE Running Buffer. To identify the bands containing the desired proteins, 7 µL of the BenchMark™ Protein Ladder was added. The gel was run on 280 V for 17 min. After, the gel was transferred to a Stain-Free Sample Tray (Bio-Rad) and gel imaging was performed using Gel Doc™ EZ Imager (Bio-Rad). This method uses the intrinsic fluorescence of the aromatic amino acids in proteins to detect the protein bands. The fractions containing the desired proteins were further processed in a second purification step (section 3.4.8.).

3.4.7. CONCENTRATING PROTEIN SOLUTION WITH CENTRIFUGAL FILTERS

Materials:

- Protein solution (section 3.4.5.)
- Amicon Ultra-15 Centrifugal Filter Unit, 10K

Method:

To further up-concentrate samples, centrifugal filters (10K) which retains proteins larger than 10 kDa is used. Particles smaller than 10 kDa will go through the filter, making it suitable for up concentration of proteins larger than this. The protein solutions were applied to the filters and centrifuged at 4300 g, 4 ˚C until 1.0-1.5 mL sample was left in the filter. Afterwards, the samples were kept at 4 ˚C.

3.4.8. SIZE EXCLUSION CHROMATOGRAPHY

Materials:

- HiLoad™ 16/600 Superdex™ 75 prep grade
- Buffer A: degassed 50 mM Tris-HCl pH 8.5, 150 mM NaCl
- Degassed dH_2O
- Degassed 20 % ethanol
- Up-concentrated protein solution (1.0 mL) (section 3.4.7.)

Method:

Size exclusion chromatography (SEC) was the second purification step to separate and purify proteins. The column used was a HiLoad^{TM} 16/600 Superdex^{TM} 75 prep grade (120 mL) coupled to a UV detector. The column was operated at a flow of 1 mL/min. Firstly; the column was washed for 30 min with degassed dH_2O (0.5 CV). Then, the column was equilibrated with buffer A for at least 1 CV. Up-concentrated protein (section 3.4.7.) was filtrated, applied on the SEC column and eluted with buffer A. Fractions of 1 mL were collected. The smaller proteins diffuses easily into the gel beads pores and have thus a longer retention time than larger proteins that do not "visit" as many pores and elute therefore first (GE Healthcare, 2018). After eluting the proteins, buffer A ran on the column for two hours. After this, the column was washed with degassed dH_2O for 30 min and 20% ethanol for two hours. The column was stored at room temperature. The protein fractions were kept at 4 ˚C until SDS-PAGE analysis (section 3.4.6.).

3.4.9. PURIFICATION OF HIS-TAGGED PROTEINS

Imobilized metal affinity chromatography (IMAC) was used to purify the His-Tagged proteins. IMAC consist in loading the solution containing the protein of interest onto a prepacked column with Ni-NTA resin containing a metal ion (Ni^{2+}) for which the target proteins present an affinity (via the His-Tag). Proteins that lack the His-Tag (i.e. all unwanted proteins), and thus lack affinity for the metal ion are washed out, while the protein of interest binds to the column resin until elution. Elution and recovery of captured His-Tagged protein is accomplished by using a high concentration of imidazole.

Materials:

- Ni-NTA Agarose resin
- Purification Column
- Buffers (**Table 2.4**)
  - Native binding buffer
  - Native wash buffer
  - Native elution buffer
- Degassed dH_2O
- Degassed 20 % ethanol

- Protein solution (section 3.4.4.)

Method:

Disposable columns were packed with Ni-NTA resin by applying gravity-assisted flow. The resin was washed with dH_2O (2 CV; 2 x 14 mL) and equilibrated with native binding buffer (2 CV) before the protein solution was loaded. After loading the protein solution, native wash buffer (2 CV) was loaded on the column to wash out the proteins lacking the affinity to the resin. By using a gradient of imidazole concentration (50 mM, 100 mM and 250 mM), elution of the His-Tagged protein was accomplished.

3.5. ENZYMATIC ASSAYS

3.5.1. COPPER SATURATION

Materials:

- PD MiniTrap^{TM} G-25 column
- $CuSO_4$ (50mM)
- Storage buffer: 20 mM TrisHCl pH 8.0
- Purified *Cj*AA10A variants (section 3.4.8. and section 3.4.9.)

Method:

LPMOs are copper-dependent enzymes and it is therefore crucial to ensure that each active site is loaded with a copper atom. Isolated variants of the *Cj*AA10A were copper-loaded by incubating them at room temperature for 15 min with 3 times the molar concentration of $CuSO_4$ compared to the enzyme concentration. To equilibrate the column 3 x 5 mL 20 mM TrisHCl pH 8 was added, before applying the samples (between 200 and 400 µL). By applying the enzymes on the PD MiniTrap G-25 column the excess of copper is removed. The total sample volume applied should be 1 mL, thus 20 mM TrisHCl pH 7.0 was added additionally to reach the desired volume. To elute the copper-saturated *Cj*AA10A variants 20 mM TrisHCl pH 8.0 was applied and collected in three fractions: two times 1 mL and one 500 µL. The protein concentration in each fraction were determined as explained in section 3.1.4., and the copper-loaded LPMO solution kept at 4 ˚C until further experiments.

3.5.2. BINDING EXPERIMENTS

Materials:

- Copper-saturated *Cj*AA10A variants (section 3.5.1.)
- 0.5 M sodium phosphate buffer pH 7.0
- β-chitin (20 mg/mL)
- Avicel (50 mg/mL)
- Eppendorf Comfort Thermomixer
- 96-well filter plate
- Vacuum manifold

Method:

A binding test was carried out to determine how well the *Cj*AA10A variants bind to different substrates. A mix of buffer (50 mM final), substrate (10 mg/mL final) and water to a total volume of 600 µL was incubated in a Thermomixer at 40 ˚C, 1000 rpm for 10 min. β-chitin and Avicel was used as substrates for the assay and the copper-saturated enzymes were added (1 µM final) to start the reactions. The reactions were sampled after 10, 30, 60, 120 and 240 min and immediately filtrated using a 96-well filter plate operated by a vacuum manifold to stop the reactions. The concentrations were measured by the Bradford method, described in section 3.1.3.

3.5.3. HYDROGEN PEROXIDE ASSAY

When not bound to the substrate, LPMOs are known to produce hydrogen peroxide (H_2O_2) (Kittl et al., 2012). The H_2O_2 produced by the unbound LPMOs serves as co-substrate in a coupled reaction where the horseradish peroxidase (HRP) converts the Amplex® Red reagent (**Table 2.5**) into resorufin, a chromogenic and fluorescent product (Kittl et al., 2012). With the use of Amplex® Red reagent that reacts with H_2O_2, the H_2O_2 produced by the unbound AAs can be detected. The production of H_2O_2 was quantified by measuring the absorbance at 540 nm using a Multiskan FC spectrophotometer. The stoichiometry of the HRP-catalyzed reaction is (1:1).

Materials:

- Copper-saturated *Cj*AA10A variants (section 3.5.1.)
- Multiskan^{TM} FC Microplate Photometer
- Amplex® Red, Hydrogen Peroxide/Horseradish peroxidase
- 96-well microplate

- L-Ascorbic acid
- dH_2O

Method:

A 180 µL mixture consisting of a Amplex® Red reagent (100 µM final), HRP (5 U/mL final) and of the *Cj*AA10A variants (1 µM) in 50 mM sodium phosphate buffer, pH 7.0 was distributed in a 96-well microplate. By adding 20 µL L-Ascorbic acid (final concentration of 50 µM) the reaction started (200 µL final volume). Three parallels of the sample reactions were incubated for 30 min at room temperature and H_2O_2 production was quantified. The H_2O_2 standard stock was diluted to make a standard curve, ranging from 10 to 40 µM final concentration.

3.5.4. ACTIVITY ASSAY

Materials:

- Copper-saturated *Cj*AA10A variants (section 3.5.1.)
- 0.5 M sodium phosphate buffer pH 7.0
- β-chitin (20 mg/mL)
- α-chitin (20 mg/mL)
- Avicel (50 mg/mL)
- PASC (12 mg/mL)
- L-Ascorbic acid (100 mM)
- Eppendorf Comfort Thermomixer
- 96-well filter plate
- Vacuum manifold

Method:

An activity assay was carried out to determine the product profile of *Cj*AA10A variants on various substrates. A mix of buffer (50 mM final), substrate (10 mg/mL final), copper-saturated enzymes (0.5 µM final) and water to a total volume of 500 µL was incubated in a Thermomixer at 40 ˚C, 1000 rpm for 10 min. α-chitin, β-chitin, Avicel and PASC was used as substrates for the assay and 5 µL L-Ascorbic acid was added to start the reactions. The reactions were sampled after 30, 60, 120, 240 min and 24 hours and immediately filtrated using a 96-well filter plate operated by a vacuum manifold to stop the reactions. The filtered samples were stored at – 20 ˚C before product analysis by high performance anion-exchange chromatography

(section 3.5.6.). The remaining of the samples not filtered was stored at – 20 ˚C before MALDI-Tof MS analysis (section 3.5.5.).

3.5.5. MALDI-TOF MS ANALYSIS

Materials:

- Ultraflex MALDI-TOF/TOF Mass Spectrometer (Bruker Daltonics)
- MALDI TOF/MS 348 target plate ground steel TF (Bruker Daltonics)
- 2,5-dihydroxybenzoic acid (9 gL^{-1}) dissolved in H_2O:acetonitrile (1:3 v/v)
- LPMO reaction samples (section 3.5.4.)

Method:

Matrix Assisted Laser Desorption/Ionization-Time of Flight (MALDI-ToF) was used to analyze the samples. The samples were prepared by mixing 1 µL sample and 2 µL 2,5-dihydroxybenzoic acid (9 gL^{-1}, dissolved in a mixture of H_2O:acetonitrile, 1:3 v/v). The mix was applied to a MTP 384 target plate followed by drying under a steam of air. The samples were analyzed with an Ultraflex MALDI-TOF/TOF instrument (Bruker Daltonics GmbH, Bremen, Germany) equipped with a Nitrogen 337 nm laser beam, using Bruker FlexAnalysis software.

3.5.6. PRODUCT ANALYSIS BY HIGH PERFORMANCE ANION-EXCHANGE CHROMATOGRAPHY

Soluble oxidized products released from activity assay were analyzed using high-performance anion exchange chromatography (HPAEC) coupled with pulsed amperometric detection (PAD). The weakly acidic nature of carbohydrates at high pH is used to force the carbohydrates to be electrocatalytically oxidized at the gold electrode by applying a positive potential. As the carbohydrates are oxidized when touching the working electrode surface, the PAD detector measures the potential variation resulted by oxidizing and reducing conditions on the electrode surface. Many potentially interfering species cannot be oxidized or reduced, and thus not detected, making the PAD a sensitive and highly selective detector. The PAD only detects electroactive species, the current generated is proportional to the carbohydrate concentration, and therefore carbohydrates can be detected and quantified.

Materials:

- LPMO reaction samples (section 3.5.4.)
- Dionex Bio-LC (ICS 3000) equipped with a CarboPac^{TM} PA1 column
- Pulsed amperometric detector (PAD)
- Gold electrode
- Eluent A: Degassed 0.1 M NaOH
- Eluent B: Degassed 0.1 M NaOH, 1 M $NaCH_3COOH$
- Eluent C: Degassed dH_2O

Method:

Prior to the injection of the samples into the column, the CarboPac^{TM} PA1 column was washed with eluent C and then equilibrated using the eluent A. The samples were applied to the column by an automatic sampler. The carbohydrate products in each sample were separated using a linear gradient in several steps, with a flow rate of 0.25 mL/minute, a column temperature of 30 ˚C and a gradually increasing concentration of eluent B (0.1 M NaOH, 1 M NaOAc), as follows: 0-10 % B over 10 min, 10-18 % B over 10 min, 18-30 % B over 1 min, 30-100 % B over 1 min, 100-0 % B over 0.1 min and 0 % B over 13.9 min. Oxidized dimers were quantified using GlcGlc1A standards that were prepared in-house (Bissaro et al., 2017).

Eluted cello-oligosaccharides were detected by PAD and chromatograms recorded using Chromeleon 7.0 (Dionex) software.

4. RESULTS

4.1. ANALYSIS OF CBM DISTRIBUTION THROUGH THE SEQUENCE DIVERSITY OF AA10As

As of April 22nd, 21 members of the AA10 family have been characterized (biochemically and/or structurally) while more than 3000 sequences are predicted to encode for AA10s. As explained in section 1.4., the AA10 family that is mainly of bacterial origin is the only one composed of enzymes targeting the two most abundant biopolymers cellulose and/or chitin, with C1 and/or C4 oxidative regioselectivity. This is likely to reflect a more ancestral character of AA10s, before acquisition by other microorganisms (e.g. fungi or insects) and substrate specialization occurred. In the course of this project we were interested in understanding the determinants of molecular evolution between chitin and cellulose-active AA10s. Beyond the importance of substrate recognition by the catalytic domain Nature has evolved CBMs (see section 1.4.7.) to ensure that the catalytic domains better recognize and anchor to their substrate. The few chitin-active AA10s hitherto characterized are usually devoid of such CBM (Vaaje-Kolstad et al., 2005) while several of the cellulose-active AA10s did acquire cellulose-binding domains (Forsberg et al., 2014a). However, beyond the few examples available there is no clear global view about the extent of CBM absence or acquisition throughout the AA10 family. Here, the aim was to map the presence or absence of CBMs (and their predicted substrate affinity) on a global phylogenetic tree representative of the AA10 diversity (**Figure 4.1**). B. Bissaro previously performed a phylogenetic analysis based on MSA of catalytic domains. Here, for each sequence, using the sequence reference number a search in the Pfam database (http://pfam.xfam.org/) allowed us to trace back the predicted modular structure of each enzymes which allowed us to annotate the tree.

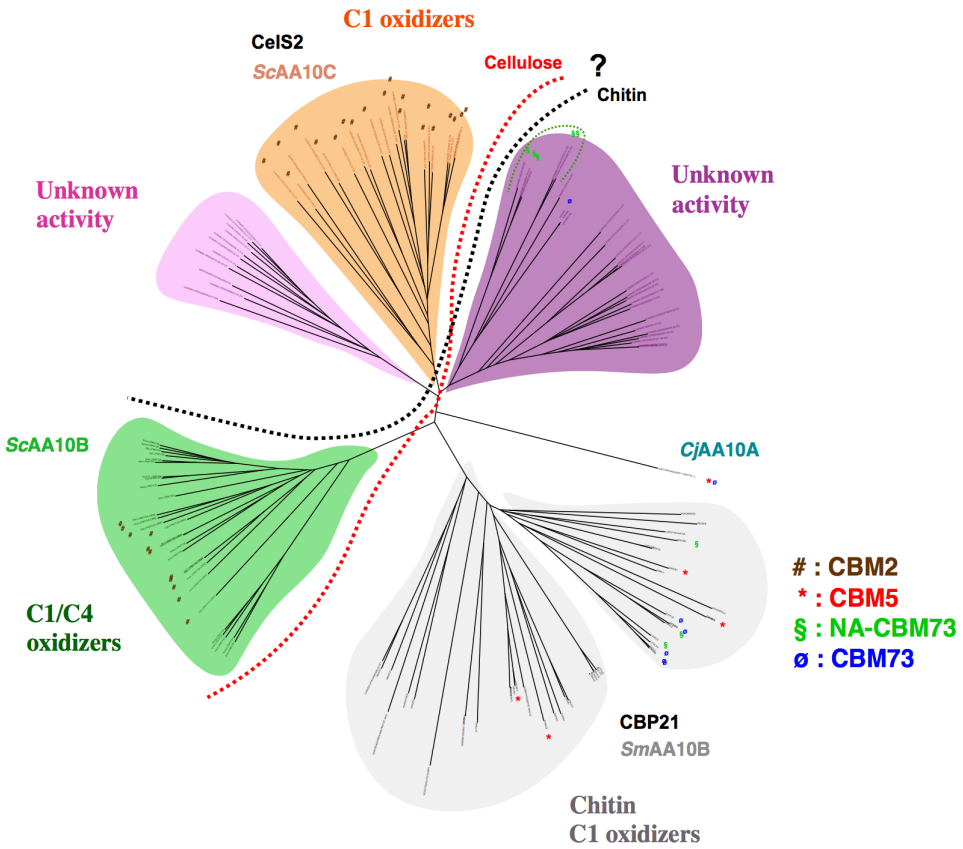


Figure 4.1. Global view on the extent of CBMs in AA10As represented in a phylogenetic tree. The presence or absence of CBMs is mapped on a global phylogenetic tree representative of the AA10 diversity. Three main groups can be observed: the C1 oxidizing chitin-active (grey circle) and the cellulose-active oxidizing at C1 (orange circle) or C1/C4 (green circle).

One can observe that the C1 oxidizing chitin-active AA10s (grey circle in **Figure 4.1**, e.g. CBP21) mainly do not harbor any CBMs, in some cases a CBM5, a CBM73 or a combination of an unknown sequence and a CBM73 is observed. The cellulose-active C1 oxidizing AA10s (orange circle in **Figure 4.1**, e.g. CelS2) either have a cellulose-binding CBM2 or no CBM at all. The magenta circle on the left in **Figure 4.1** was believed to contain cellulose-active enzymes, but after conducting this CBM analysis, it is more likely that the enzymes here are chitin-active, due to the fact that they harbor no CBM2 (cellulose-binding CBM) and in a few cases they even display an unknown sequence and a CBM73 (chitin-binding CBM). Regarding the mix of cellulose and chitin-active AA10s (green circle in **Figure 4.1**, e.g. *Sc*AA10B), either no CBMs or CBM2 were observed in this pool of sequences. Overall, LPMOs containing one or several CBMs represent 36 % of CBP21-like enzymes, 94 % of CelS2-like enzymes and 39 % of *Sc*AA10B-like enzymes, respectively.

4.2. CLONING, MUTAGENESIS AND TRANSFORMATION

In light of the fact that CBMs contribute to a better recognition and binding to the substrate, the CBM and linker of *Cj*AA10A (chitin-specific) was switched with the CBM and linker of *Sc*AA10C (cellulose-specific) to examine the effect this would have on the substrate specificity. To yield these constructs, one with the original *Cj*Linker and a new *Sc*CBM2 and one with both new *Sc*Linker and *Sc*CBM2, homologues recombination of designed plasmids and inserts was carried out.

4.2.1. CLONING OF *Cj*AA10A AND VARIANTS

Amplification of plasmid 1, insert 1, plasmid 2 and insert 2 was successful. The resulting gel electrophoresis images from the PCR products run on a 0.8 % and 2 % agarose gel (section 3.1.5.) are displayed in **Figure 4.2**. (Plasmid 1-Insert 1) and (Plasmid 2-Insert 2) pairs were used in homologues recombination cloning to yield *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4) constructs, respectfully.

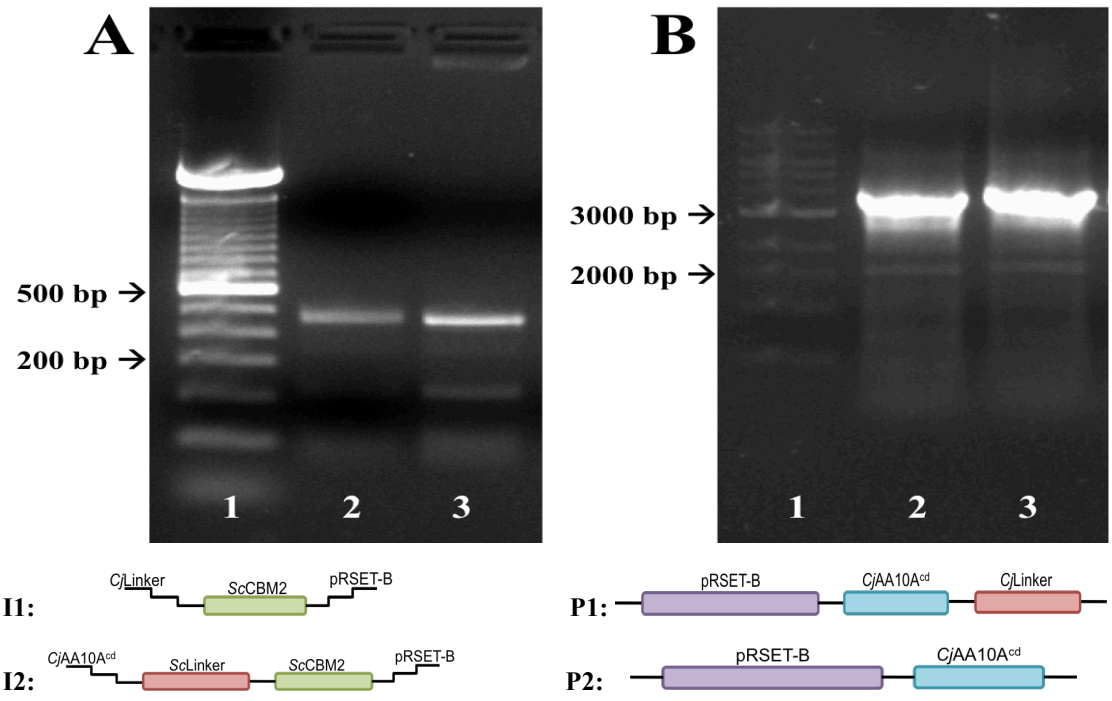


Figure 4.2. Agarose (2% (A) and 0.8% (B)) gel electrophoresis of insert (A) and plasmid (B) generated during PCR. In **panel A**, line 1 shows the GeneRuler 100 bp ladder, followed by the PCR products generated for the insert 1 (336 bp, lane 2) and insert 2 (441 bp, lane 3). **Panel B** contains the GeneRuler 1 kb ladder in line 1 followed by the PCR products generated for the plasmid 1 (3500 bp, lane 2) and plasmid 2 (3400 bp, lane 3). The drawings below each gel indicate the expected corresponding protein architecture.

Sequencing of the isolated plasmids (section 3.1.9.) verified that the desired genes had been successfully generated and that no changes, other than the desired swapping of CBM and linker, had occurred.

The modular architecture of *Cj*AA10A, *Sc*AA10C and variants thereof are shown in **Figure 4.3**.

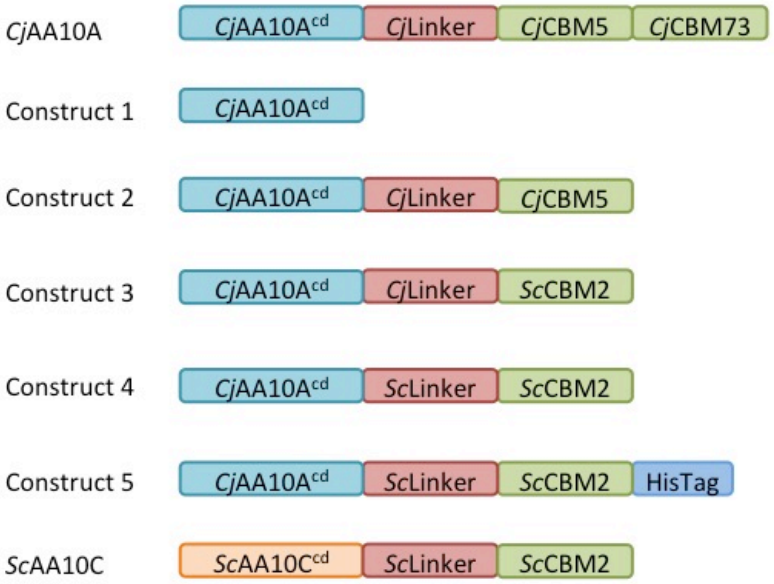


Figure 4.3. Modular architecture of *Cj*AA10A, *Sc*AA10C and variants thereof. *Cj*AA10A and *Sc*AA10C have C-terminal carbohydrate-binding modules (CBMs) from families 5/73 and 2 respectively, connected to the catalytic domain (cd) via polypeptide linkers. Construct 1 shows *Cj*AA10A^{cd}, construct 2 shows *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5, construct 3 shows *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2, construct 4 shows *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 and construct 5 shows *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag. *Cj*AA10A and *Sc*AA10C show the full length "natural" version of *Cj*AA10A and CelS2.

4.2.2. HIS-TAG ADDITION ON *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2

In order to accelerate the purification procedure for any future mutants generated from the platform enzyme, *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4), a 6 x His-Tag was added. Several approaches have been tried to introduce the His-Tag encoding sequence in construct 4, but only the successful one is presented here.

Following the cloning of *Cj*AA10A variants, *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 was used as template in a PCR procedure employing primers with overhang containing a His-Tag-encoding sequence (see **Table 2.11**). **Figure 4.4** shows agarose gel electrophoresis of amplified insert and vector generated during PCR. The plasmid and insert amplicons were then used in a homologous recombination cloning reaction to yield the construct 5; *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag. This cloning procedure adds a poly-histidine tag (6 × His-Tag) to the C-terminus of the protein.

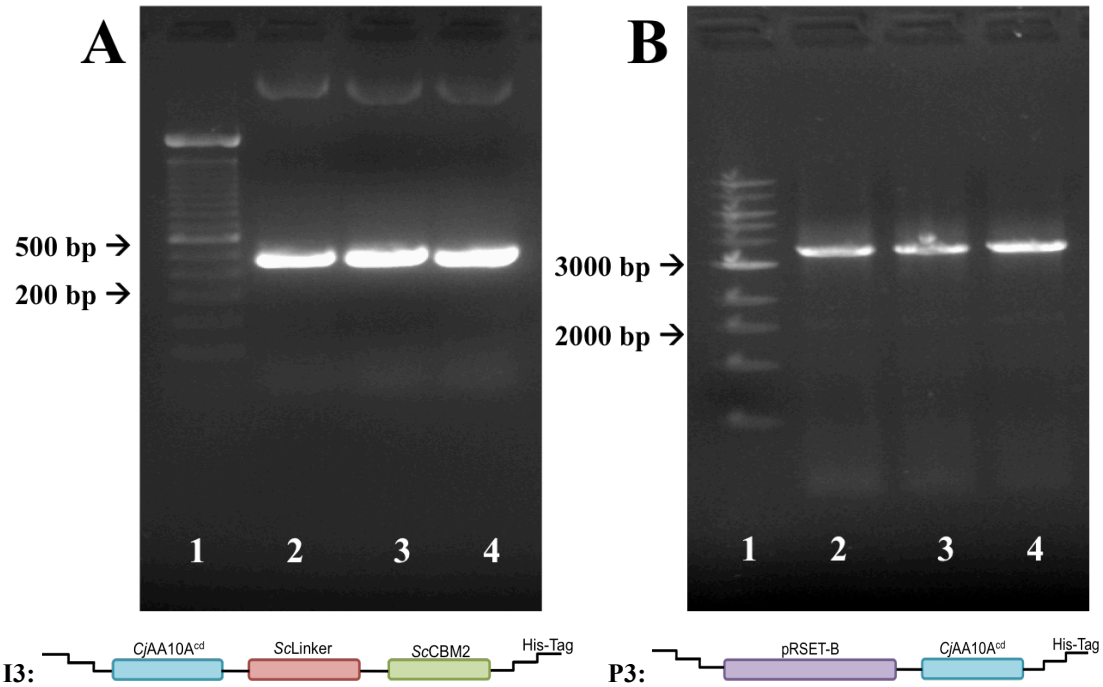


Figure 4.4. Agarose (2% (A) and 0.8% (B)) gel electrophoresis of insert (A) and plasmid (B) generated during PCR. In **panel A**, line 1 shows the GeneRuler 100 bp ladder, followed by the PCR products generated for the insert (447 bp) done with a variation of annealing temperatures (line 2 = 62˚C; line 3 = 65˚C; line 4 = 68˚C). **Panel B** contains the GeneRuler 1 kb ladder in line 1 followed by the PCR products generated for the plasmid (3500 bp) done with a variation of annealing temperatures (line 2 = 63˚C; line 3 = 66˚C; line 4 = 69˚C). The drawings below each gel indicate the expected corresponding protein architecture.

4.2.3. SITE-DIRECTED MUTAGENESIS

Following the design of *Cj*AA10A variants and validation of the fact that they could be expressed, purified and were still active (section 4.6), mutations were introduced in the *Cj*AA10A^{cd} domain of *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2-HisTag. The choice of mutations was guided by a correlation network showing the degree of co-evolved mutations in the regions between CBP21 (*Sm*AA10A), CelS2 (*Sc*AA10C) and *Cj*AA10A.

Using site-directed mutagenesis Q78F and Q78N were produced. Several attempts were conducted to generate T133Q and T133W mutations, but all with unsuccessfully results. Clones of T133Q and T133W were obtained, but sequencing results indicated that the mutations were lacking.

Mutated expression vectors were verified by DNA sequencing. The sequencing data was analyzed by pairwise sequence alignment (https://www.ebi.ac.uk/) and the results are listed in **Table 4.1**.

Table 4.1. Overview of results from DNA sequencing of mutated *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB5-HisTag encoding plasmid.

MUTATIONS	SUCCESSFUL (+) OR UNSUCCESSFUL (-)
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F	+
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N	+
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-T133Q	-
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-T133W	-

4.3. PROTEIN EXPRESSION AND PURIFICATION

4.3.1. PROTEIN EXPRESSION

All variants of *Cj*AA10A were successfully expressed. One set of SDS-PAGE analyses (**Figure 4.5**) is used in this section to represent expression of all *Cj*AA10A variants.

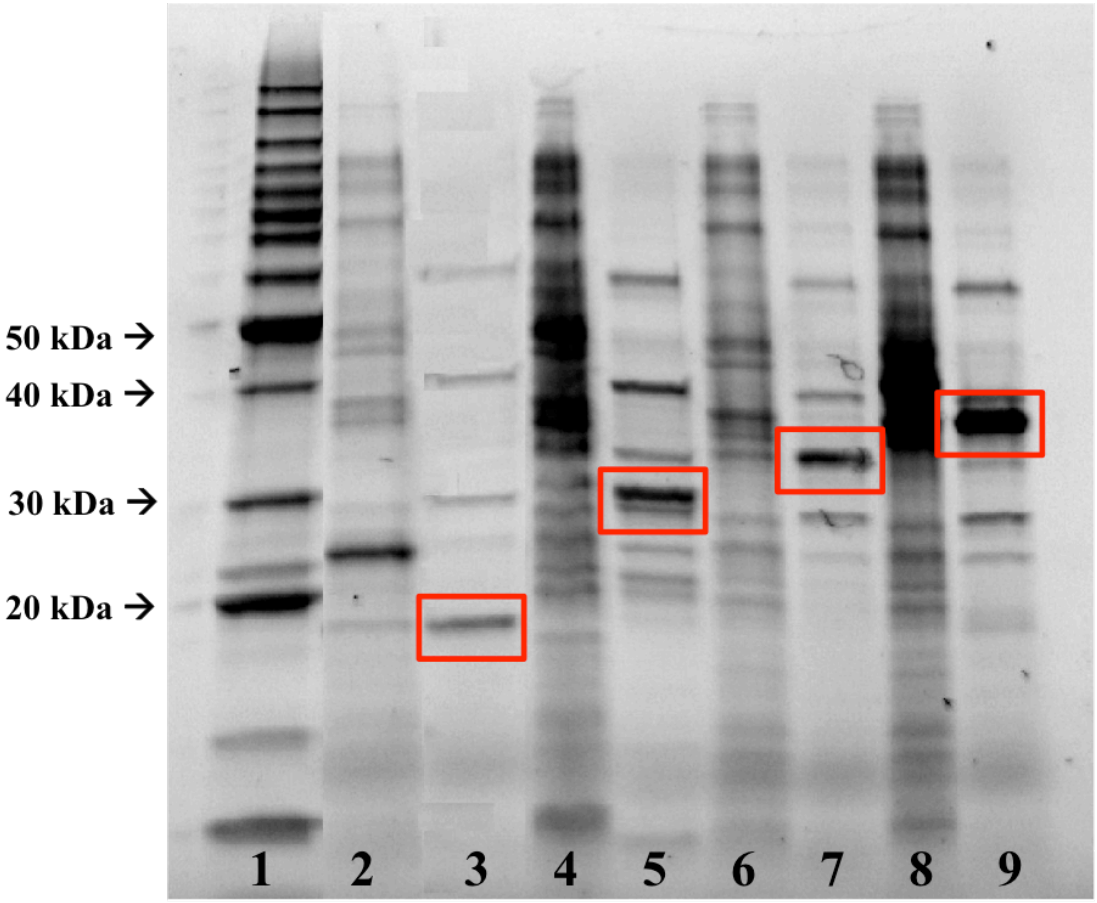


Figure 4.5. Expression of *Cj*AA10A variants. The figure shows SDS-PAGE analysis of *Cj*AA10A variants. Lane 1 is containing a BenchMark™ protein ladder with masses labeled for reference. Resuspended cell pellet and periplasmic extract are respectively shown for *Cj*AA10A^{cd} (lane 2 and 3), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (lane 4 and 5), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (lane 6 and 7) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (lane 8 and 9). Proteins of interest are indicated by a red square (See **Table 2.15** for enzyme properties).

4.3.2. PROTEIN PURIFICATION

Purification of *Cj*AA10A^{cd}, *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5, *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 was performed by ion exchange chromatography (IEC, section 3.4.5.) followed by size exclusion chromatography (SEC, section 3.4.8.). Purification of *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag and mutants was performed by purification of His-Tagged enzymes (section 3.4.9.) followed by ion exchange chromatography. One set of chromatograms is used in this section to represent purification of all *Cj*AA10A variants (**Figure 4.6** and **Figure 4.7**). Fractions containing pure LPMO were identified by SDS-PAGE and subsequently pooled. **Figure 4.8** shows SDS-PAGE analyses of LPMOs purified by both ion exchange chromatography and size-exclusion chromatography. **Figure 4.9** shows SDS-PAGE analyses of LPMOs purified by purification of His-Tagged enzymes and ion exchange chromatography. An SDS-PAGE analysis, including all purified constructs in this study, is shown in **Figure 4.10**.

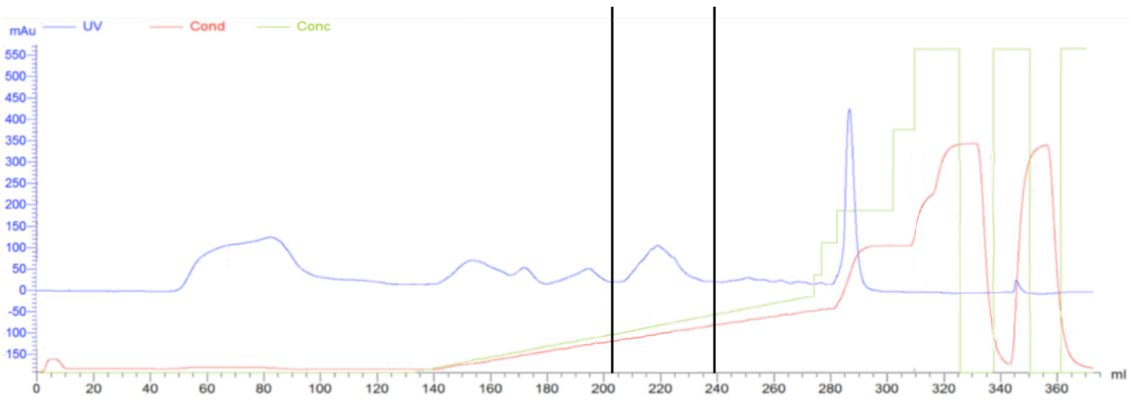


Figure 4.6. Chromatogram for purifying *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 on an ion exchanger. A broad peak containing contamination and undesired proteins arises at the start of the chromatogram when the periplasmic extract was applied into the column. The peak, marked between two black bars, contains *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2, as verified by gel electrophoresis. All variants were purified in the same manner.

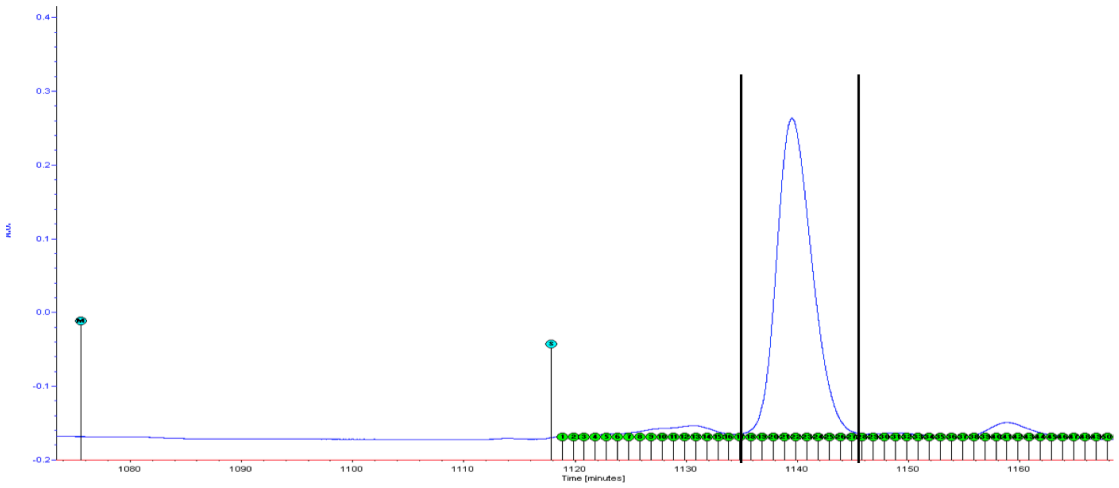


Figure 4.7. Chromatogram for isolating *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 in size exclusion chromatography. The isolated *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2, marked with two black bars, is eluted as one narrow, tall peak. The injection time is indicated by an "M" mark at approximately 1076 min. The purity of collected fractions was verified by gel electrophoresis. All variants were purified in the same manner.

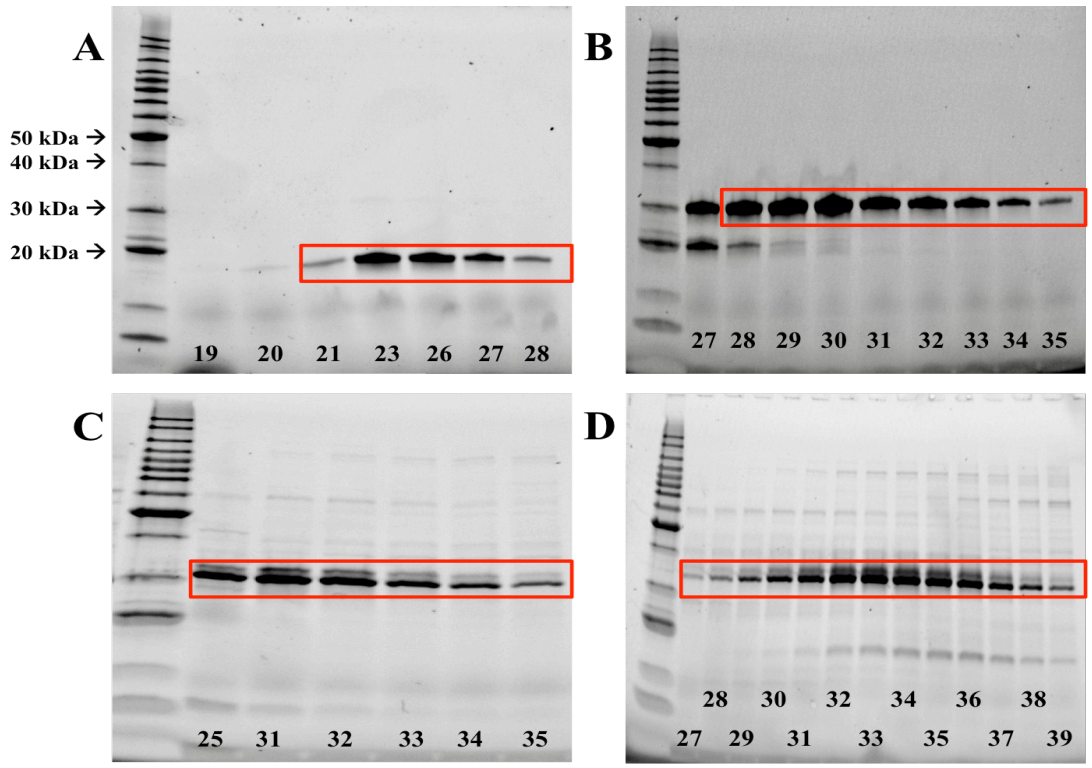


Figure 4.8. SDS-PAGE analysis of proteins after purification by ion exchange chromatography (IEC) and size-exclusion chromatography (SEC). The left lane of all four gels contains a BenchMark^{TM} protein ladder with masses labeled on the first gel for reference. All sample lanes are labeled with the corresponding fraction number collected during purification of the proteins. Proteins of interest are indicated by a red square. **Panel A** shows *Cj*AA10A^{cd}, **panel B** shows *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5, **panel C** shows *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 and **panel D** shows *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2. Proteins of interest are indicated by a red square.

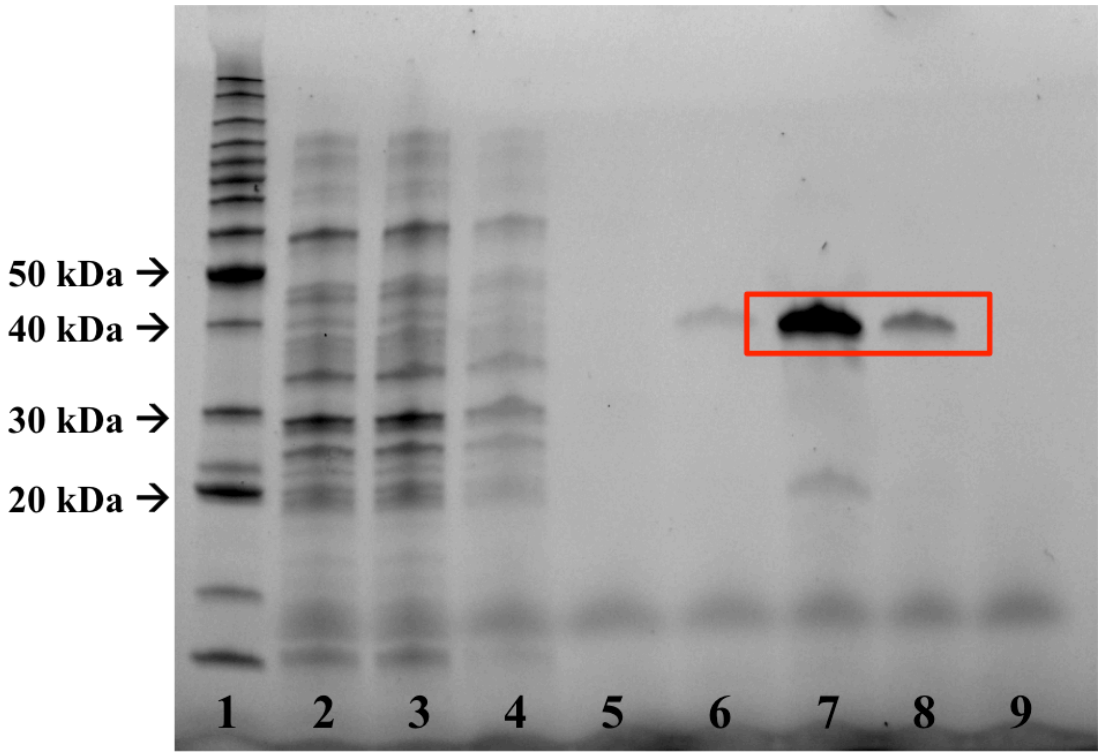


Figure 4.9. SDS-PAGE analysis of *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag purified by His-Tag purification. The left lane contains a BenchMark^{TM} protein ladder with masses labeled for reference. Line 2 shows the loading of protein. Line 3-5 shows three washing-steps followed by elution with 50 mM Imidazole in lane 6. Line 7-8 shows elution of *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag with 100 mM Imidazole. The last lane shows elution with 250 mM Imidazole. Proteins of interest are indicated by a red square.

After purification and up-concentration, the concentration was determined by measuring A_{280}. The total protein yield per liter culture is shown in **Table 4.2**.

Table 4.2. Measured final protein yields of *Cj*AA10A variants and mutants after purification.

ENZYME	mg/L YIELD
*Cj*AA10A^{cd}	~ 2
*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5	~ 3
*Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2	~ 13
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2	~ 13
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag	~ 7
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F	~ 4
*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N	~ 6

4.3.3. OVERALL PURIFICATION RESULTS

As described above, *Cj*AA10A variants were all purified in two steps. The purity of the isolated *Cj*AA10A variants after 2 months storage at 4 ˚C is shown in **Figure 4.10**. The SDS-analysis also shows the stability of the variants over time, where the two variants containing a *Cj* linker (construct 2 and construct 3) show signs of degradation.

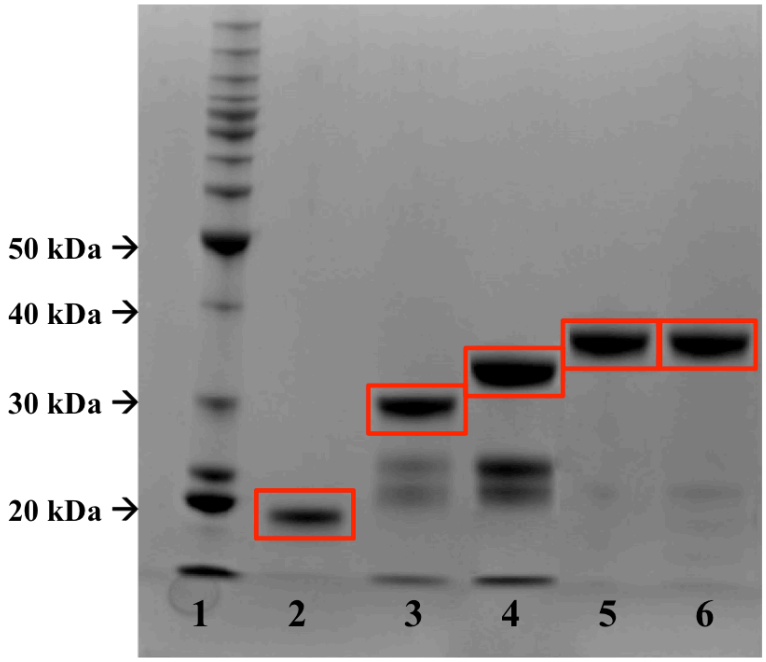


Figure 4.10. *Cj*AA10A variants expression and stability after 2 months. The figure shows SDS-PAGE gel electrophoresis containing a protein marker (BenchMark™ protein ladder, lane 1) and purified variants of *Cj*AA10A stored in sodium phosphate buffer (50 mM pH 7.0) at 4 ˚C for 2 monts. The gel shows *Cj*AA10A^{cd} (construct 1, lane 2), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2, lane 3), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3, lane 4), *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4, lane 5) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5, lane 6).

To compare the theoretical and the experimental molecular weight of the variants based on **Figure 4.10** the logarithm of kDa was plotted against the migration length in cm of the BenchMark™ protein ladder. The experimental molecular weight of constructs 1-5 is shown in **Figure 4.11** and the comparison between the theoretical and the experimental molecular weight is listed in **Table 4.3**.

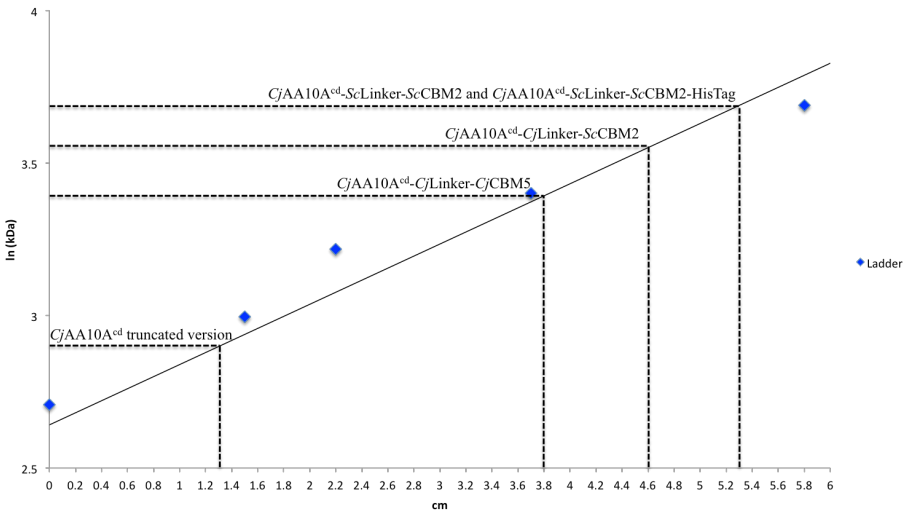


Figure 4.11. Determination of *Cj*AA10A variants molecular weight based on SDS-PAGE analysis. The figure shows the logarithm of kDa plotted against the migration length in cm of the BenchMark™ protein ladder (x = 0 cm was defined as 15 kDa). The migration distance of the different constructs was reported on the standard curve to determine their respective sizes.

Table 4.3. Comparison between theoretical and experimental molecular weights of *Cj*AA10A variants. The table shows the theoretical molecular weight[b] in kDa (based on protein sequence, see **Table 2.15**) compared to the observed size[a] in kDa from the SDS-PAGE analysis displayed in **Figure 4.10** and **Figure 4.11**.

	Construct 1	Construct 2	Construct 3	Construct 4	Construct 5
cm	1.3	3.8	4.6	5.3	5.3
ln (kDa)	2.9	3.4	3.55	3.7	3.7
Obs (kDa)[a]	18.17	29.96	34.81	40.45	40.45
Theo (kDa)[b]	20.34	29.79	33.39	34.04	34.94

4.4. POLYSACCHARIDE BINDING ASSAYS

An experiment was set up in order to investigate the binding affinity of *Cj*AA10A variants towards α-chitin, β-chitin and Avicel. The amount of unbound protein in soluble fractions of reaction mixtures was measured. **Figure 4.12** shows the binding affinity of four different constructs (constructs 1-4) to chitin and cellulose (Avicel), presented as percentage of free protein in the soluble fractions as measured by A_{280} over time. **Figure 4.13** shows the binding affinity for the variant with an additional poly-histidine tag (construct 5).

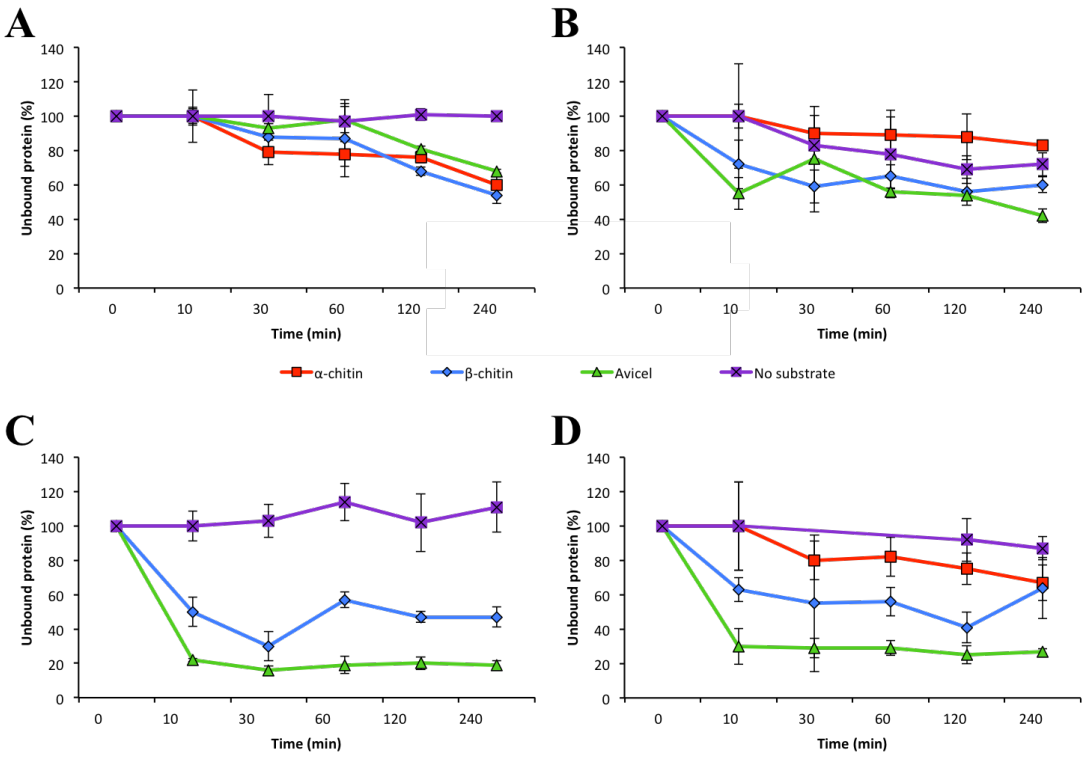


Figure 4.12. Binding of *Cj*AA10A variants to chitin and cellulose (Avicel). The graphs show the percentage of unbound protein (i.e. in solution) as a function of time for *Cj*AA10A^{cd} (construct 1, **panel A**), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2, **panel B**), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3, **panel C**) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4, **panel D**). n=3.

A difference in binding is observed for the different constructs on the various substrates (**Figure 4.12**). Construct 1 (*Cj*AA10A^{cd}) shows some binding to all three substrates after 240 min with ~ 60 % unbound protein. The truncated construct 2 (*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5) shows somewhat similar binding as construct 1, but binds less to α-chitin. Compared to the "natural" variants of *Cj*AA10A, constructs 1 and 2, the newly engineered constructs, with a cellulose binding CBM (construct 3) or

both *Sc*AA10C-derived linker and CBM (construct 4), show a much better binding to Avicel (cellulose), with ~ 20 % unbound protein.

Binding by the His-tagged version, construct 5 (**Figure 4.13**), is not as good as without the His-Tag with ~ 60-80 % unbound protein for all substrates.

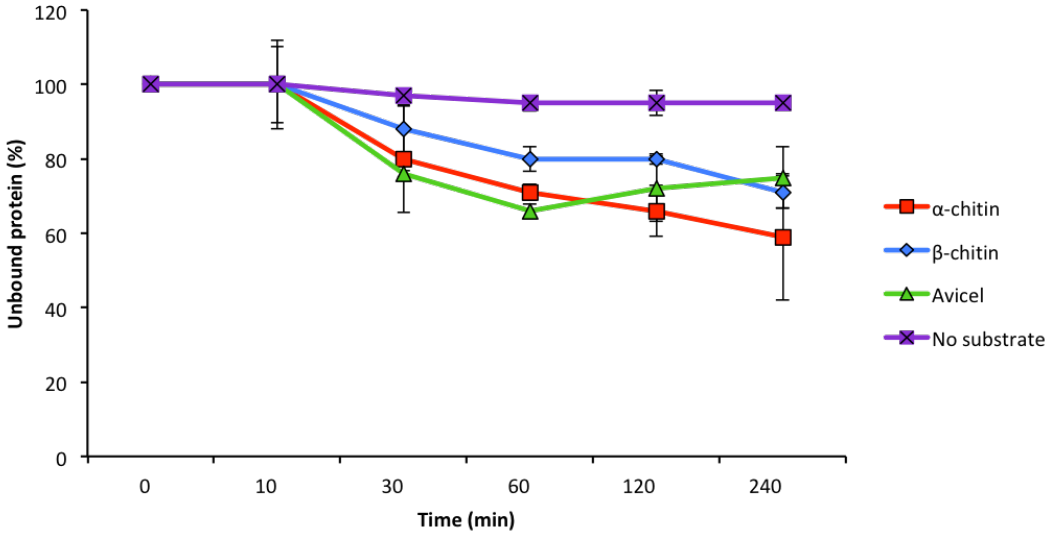


Figure 4.13. Binding of *Cj*AA10A variant to chitin and cellulose (Avicel). The graph shows the percentage of unbound protein (i.e. in solution) as a function of time for *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5). n=3.

4.5. H_2O_2 PRODUCTION EXPERIMENTS

When not bound to the substrate, LPMOs are well known for generating H_2O_2 (Kittl et al., 2012) by activation of O_2 in reducing conditions. *Cj*AA10A variants and mutants were assessed for the ability to produce H_2O_2 in the absence of substrate. The amount of H_2O_2 produced from *Cj*AA10A variants and mutants over time and the derived production rates are displayed in **Figure 4.14** and **Table 4.4**. The figure suggests that all variants except the Q78N mutant are H_2O_2 producers, but to different extent. The production of H_2O_2 seems to be lower for the constructs with both *Sc*AA10C-derived linker and cellulose-binding CBM (constructs 4, 5 and mutant Q78F) compared to the constructs containing *Cj*AA10A original linker (constructs 1-3). It is interesting to note that constructs 3 and 4, which both contain *Sc*AA10C-derived CBM but have different linkers behave very differently. The change in linker composition might alter the relative motion of both domains and, maybe, affect the catalytic domain.

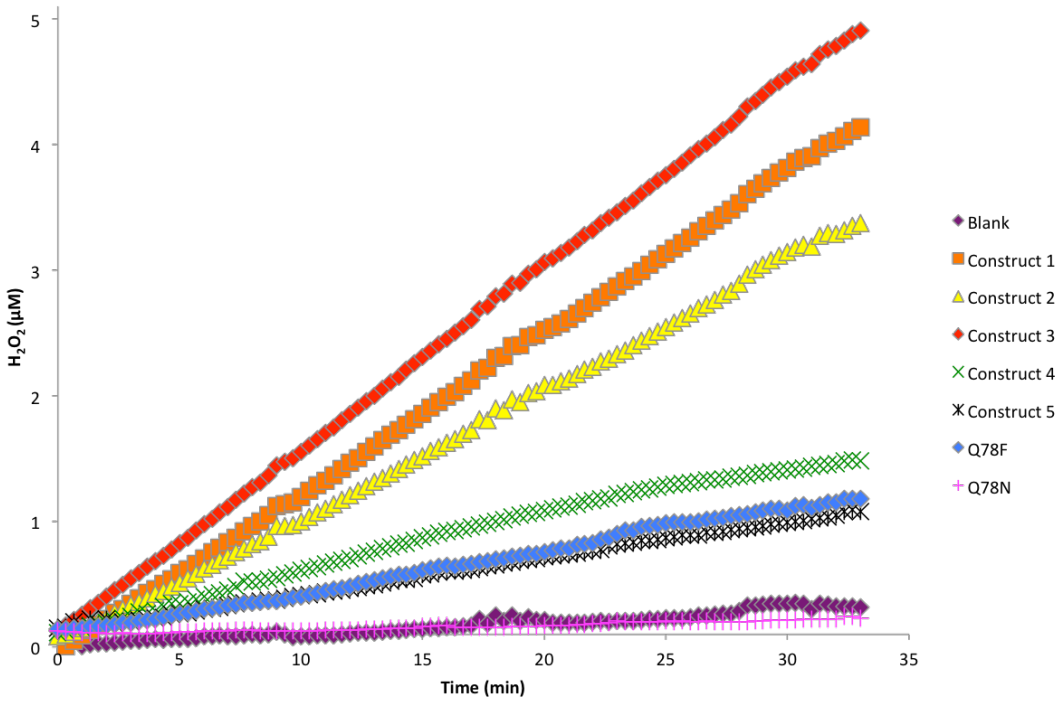


Figure 4.14. Time course of H_2O_2 production by *Cj*AA10A variants and mutants. The graph shows the H_2O_2 concentration (µM) as a function of time for each *Cj*AA10A variant and the mutants. Reactions were carried out with LPMO (1 µM), AscA (50 µM) and AmplexRed (100 µM)/HRP (5 U/mL) in sodium phosphate buffer (50 mM, pH 7.0), at 40 ˚C. The graph shows *Cj*AA10A^{cd} (construct 1), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3), *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5).

In addition, the mutants Q78F (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F) and Q78N (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N) are shown. For the sake of clarity, only one out three replicates is shown.

The H_2O_2 production rate for the different variants is displayed in **Table 4.4**.

Table 4.4. H_2O_2 production by *Cj*AA10A variants in the absence of substrate[a]. n/d = not detected. The table shows H_2O_2 production rate for *Cj*AA10A^{cd} (construct 1), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3), *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5). In addition, the mutants Q78F (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F) and Q78N (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N) are shown.

Enzyme	**Absolute H_2O_2 production rate ($\mu M.min^{-1}$)[b,c]**		**H_2O_2 production rate ($\mu M.min^{-1}$) after subtraction of LPMO-independent background**		**Activities expressed as percentage of construct 1**
	average	error[c]	average	error[d]	average
No LPMO	0.009	0.001	0.000	-	7 %
Construct 1	0.126	0.003	0.117	0.004	100 %
Construct 2	0.098	0.003	0.089	0.004	77.8 %
Construct 3	0.146	0.004	0.137	0.005	113 %
Construct 4	0.051	0.004	0.042	0.005	40.5 %
Construct 5	0.028	0.001	0.019	0.002	22 %
Q78F	0.032	0.002	0.023	0.003	25 %
Q78N	0.003	0.002	n/d	0.003	2 %

[a] Reactions were carried out in a 96-well microtiter plate incubated at room temperature. Reaction mixtures contained 1 µM of LPMO and 50 µM of ascorbic acid in 50 mM sodium phosphate buffer pH 7.0. Note the low concentration of ascorbic acid, which explains why levels of LPMO-dependent H_2O_2 generation were low.

[b] The rate is derived from the data points acquired during the 20 first minutes of the reaction.

[c] The error is the standard deviation derived from three replicates.

[d] The error represents the sum of errors calculated for the LPMO-free reaction and the one under consideration.

4.6. ANALYSIS OF LPMO ACTIVITY

4.6.1. INITIAL INVESTIGATION OF ACTIVITY OF *Cj*AA10A VARIANTS TOWARDS β-CHITIN, AVICEL AND PASC

The product profiles of *Cj*AA10A variants and mutants on β-chitin, Avicel and PASC are shown in **Figure 4.15**, **Figure 4.16** and **Figure 4.17**. Given the instability of construct 3 (i.e. containing *Cj*Linker) (**Figure 4.10**) and construct 4 (i.e. containing *Sc*Linker) being really stable, it was decided to discard construct 3 for further activity test. **Figure 4.15** show that constructs 1, 2, 4 and 5, and the mutants Q78F and Q78N are all active on chitin, while there is no activity detected on crystalline (i.e. Avicel) or amorphous (i.e. PASC) cellulose (**Figure 4.16** and **Figure 4.17**).

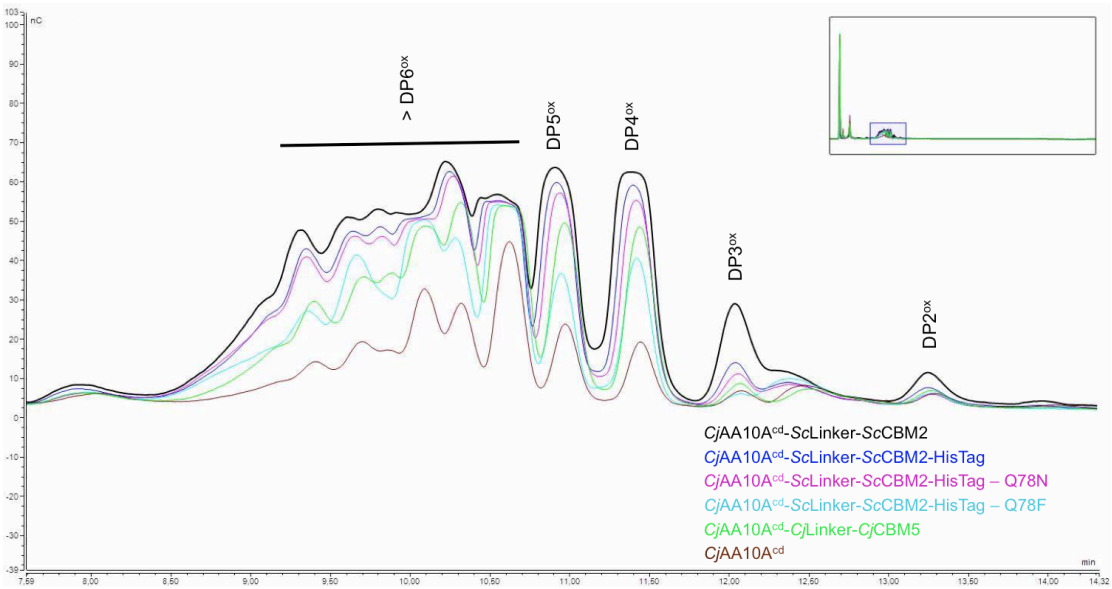


Figure 4.15. Qualitative analysis of oxidized products generated by *Cj* variants and mutants when acting on β-chitin. Samples were incubated overnight in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C and 1000 rpm before analysis by HPAEC-PAD. n=3.

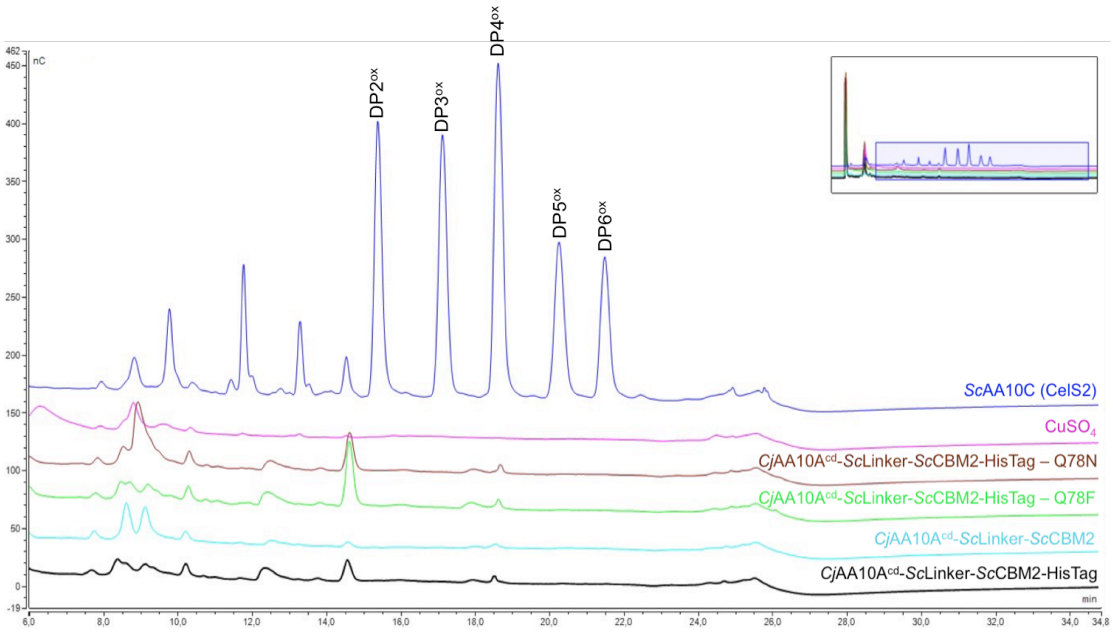


Figure 4.16. Qualitative analysis of oxidized products generated by *Sc*AA10C full length, *Cj* variants and $CuSO_4$ when acting on Avicel. Samples were incubated overnight in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C and 1000 rpm before analysis by HPAEC-PAD. n=3.

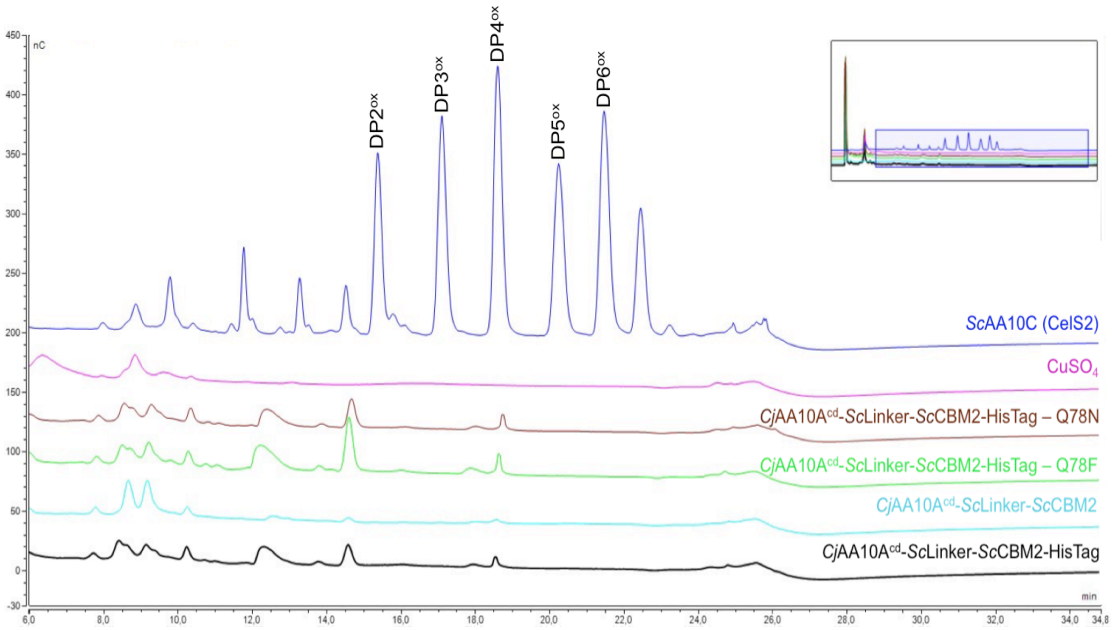


Figure 4.17. Qualitative analysis of oxidized products generated by *Sc*AA10C full length, *Cj* variants and $CuSO_4$ when acting on PASC. Samples were incubated overnight in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C and 1000 rpm before analysis by HPAEC-PAD. n=3.

4.6.2. ACTIVITY OF *Cj*AA10A VARIANTS AND MUTANTS ON β-CHITIN, AVICEL AND PASC

The overnight-incubated reactions containing β-chitin, Avicel and PASC were also analyzed using MALDI-ToF MS (**Figure 4.18**, **Figure 4.19** and **Figure 4.20**). MALDI-ToF MS analysis shows the products of the DP6 cluster, with sodium adducts of the lactone (Lac) and the aldonic acid (Ald) on β-chitin, Avicel and PASC. *Sc*AA10C (CelS2) is a positive control in **Figure 4.19** and **Figure 4.20** where none of the other constructs generated products on Avicel or PASC.

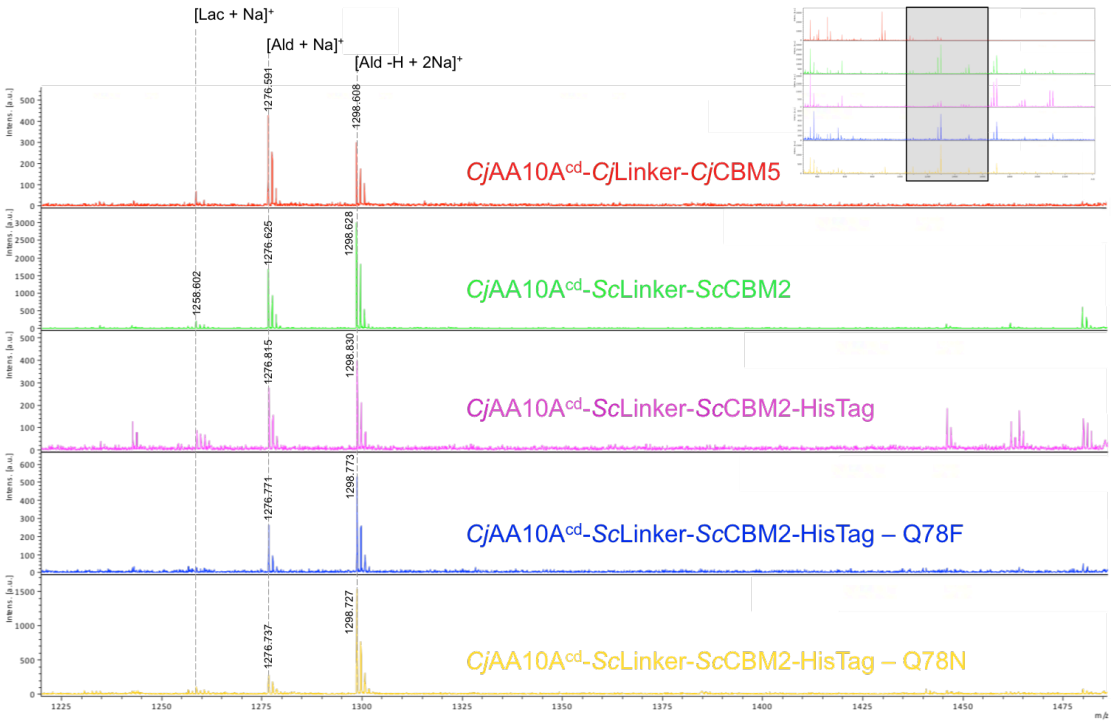


Figure 4.18. MALDI-ToF MS analysis of released oxidized cello-oligosaccharides from β-chitin by *Cj*AA10A variants after 24 h reactions. The figure shows an enlargement of the DP6 cluster (highlighted in the top right corner) of each LPMO, showing sodium adducts of the lactone (Lac) and the aldonic acid (Ald). See legend for **Figure 4.15** for more details. n=3.

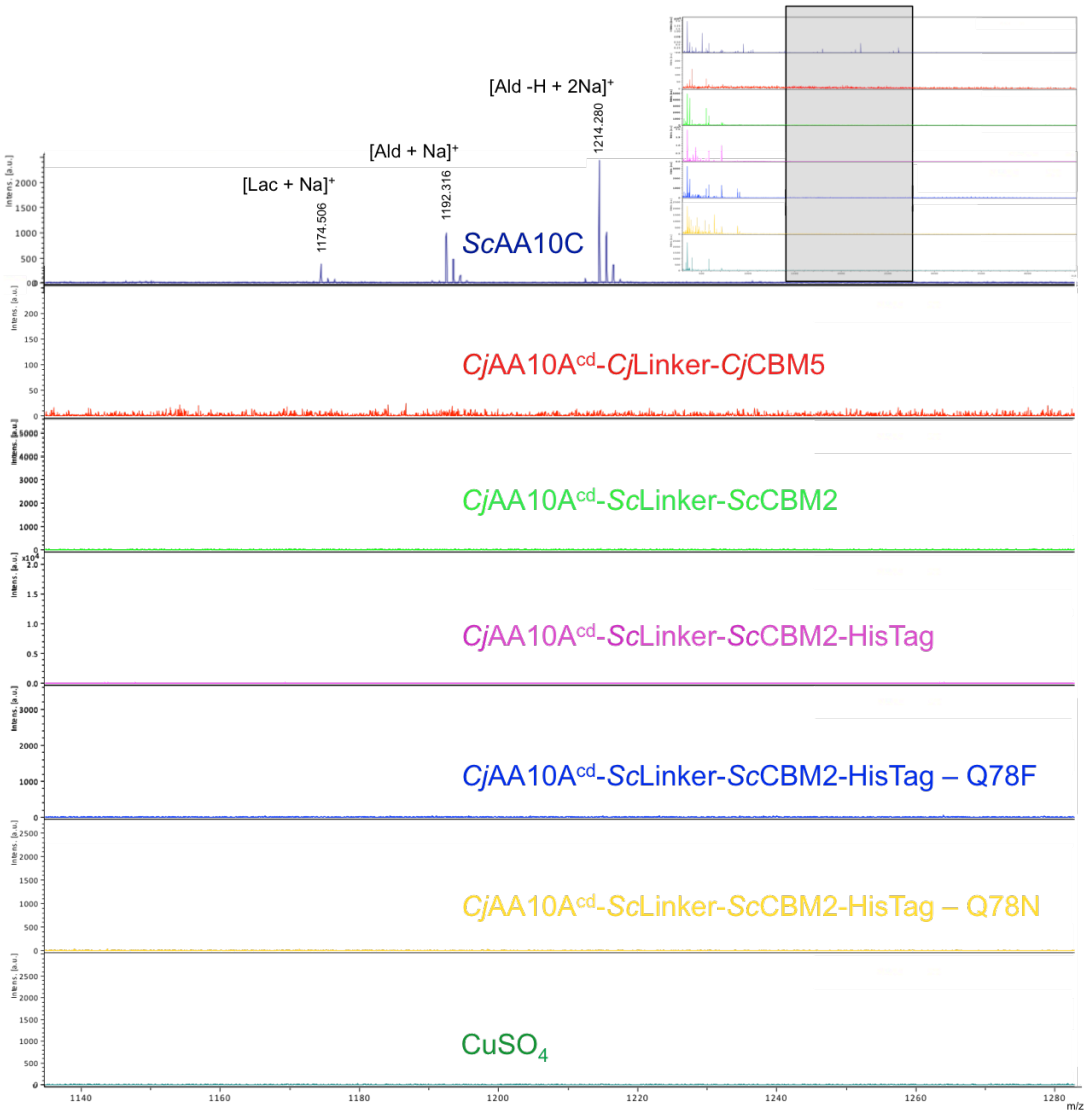


Figure 4.19. MALDI-ToF MS analysis of released oxidized cello-oligosaccharides from Avicel by *Sc*AA10C full length, *Cj*AA10A variants and $CuSO_4$ after 24 h reactions. The figure shows an enlargement of the DP6 cluster (highlighted in the top right corner) for each LPMO. See legend for **Figure 4.16** for more details. n=3.

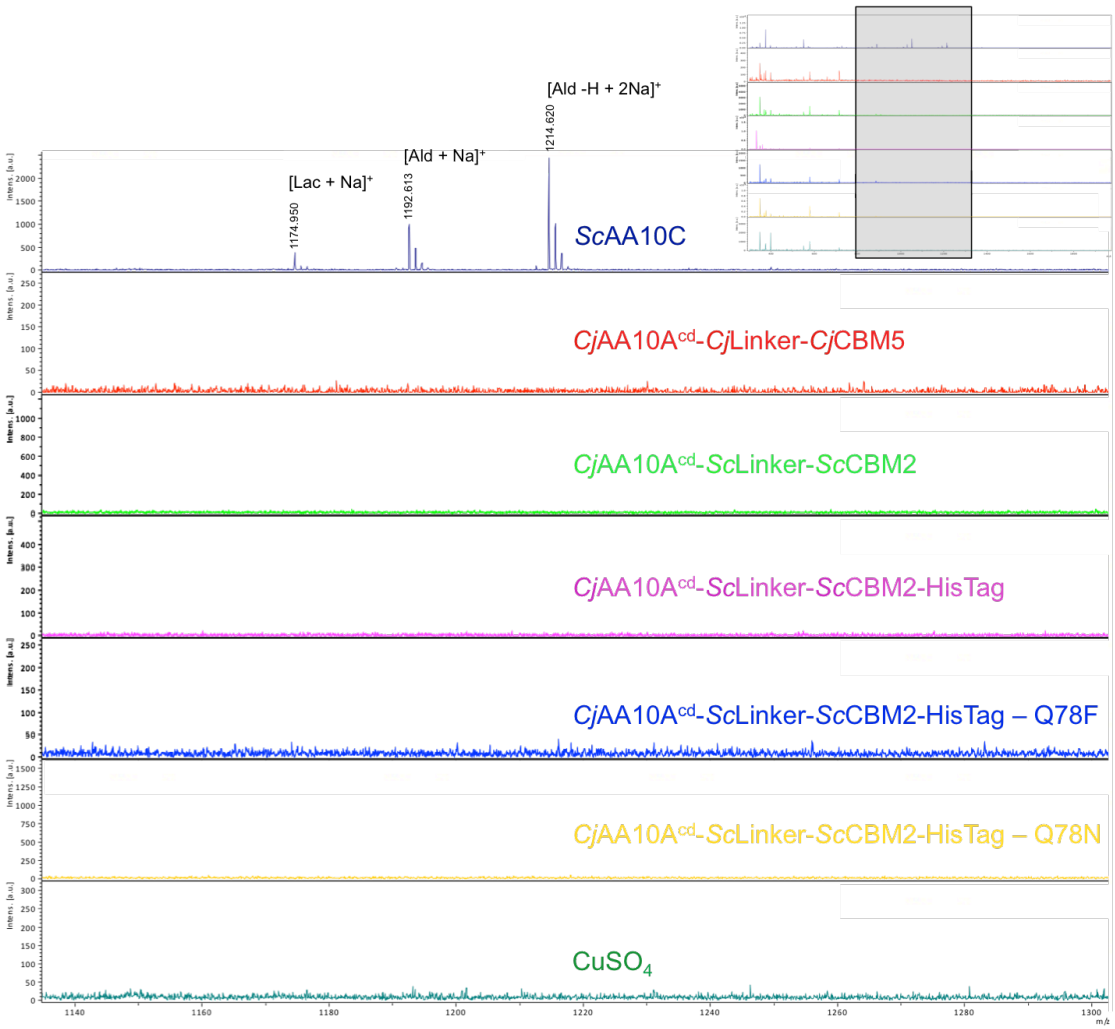


Figure 4.20. MALDI-ToF MS analysis of released oxidized cello-oligosaccharides from PASC by *Sc*AA10C full length, *Cj*AA10A variants and $CuSO_4$ after 24 h reactions. The figure shows an enlargement of the DP6 cluster (highlighted in the top right corner) of each LPMO, showing sodium adducts of the lactone (Lac) and the aldonic acid (Ald). See legend for **Figure 4.17** for more details. n=3.

4.6.3. ACTIVITY OF *Cj*AA10A VARIANTS AND MUTANTS ON α- AND β-CHITIN

Since the enzymes were shown to be active on chitin but not on cellulose (see above), their enzymatic activity was determined by performing a quantitative analysis of released oxidized chito-oligosaccharides using α-chitin and β-chitin as model substrates. The activity of *Cj*AA10A variants and mutants is shown in **Figure 4.21**, **Figure 4.22** and **Figure 4.23**.

The linearity of the graphs in **Figure 4.22** and **Figure 4.23** indicates that all variants are stable on β-chitin over time. On α -chitin (**Figure 4.21**), construct 2 (*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5) is the only variant that shows linearity over time. It appears that the modularity of constructs 1, 3 and 4 is affecting the release of oxidized chito-oligosaccharides on α-chitin.

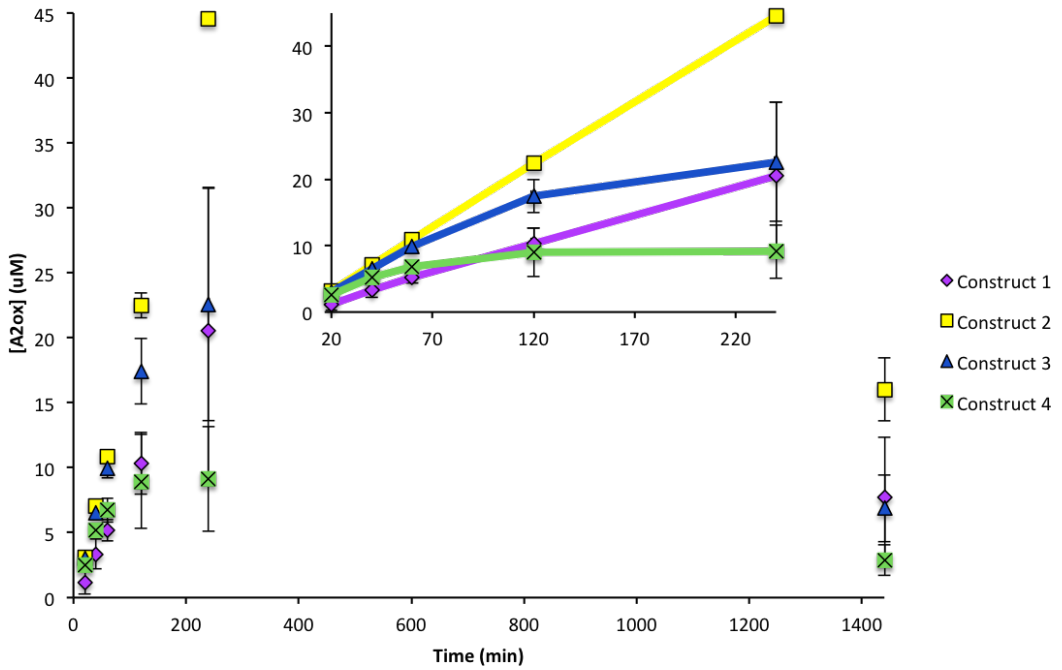


Figure 4.21. Activity of *Cj*AA10A variants on α-chitin. The graphs show the release of oxidized chito-oligosaccharide (expressed as A2^{ox}) from α-chitin by the different *Cj*AA10A constructs (1 μM) in presence of ascorbic acid (1 mM). Reactions were carried out in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C with shaking (1000 rpm). The graph shows *Cj*AA10A^{cd} (construct 1), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4). n=3.

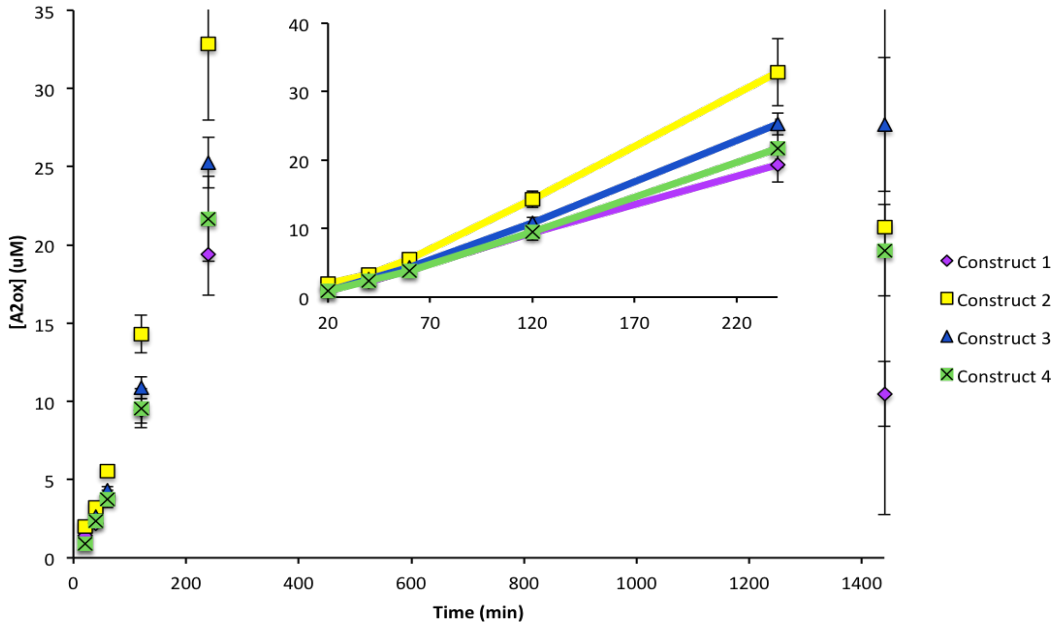


Figure 4.22. Activity of *Cj*AA10A variants on β-chitin. The graphs show the release of oxidized chito-oligosaccharide (expressed as $A2^{ox}$) from β-chitin by the different *Cj*AA10A constructs (1 µM) in presence of ascorbic acid (1 mM). Reactions were carried out in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C with shaking (1000 rpm). The graph shows *Cj*AA10A^{cd} (construct 1), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2), *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4). n=3.

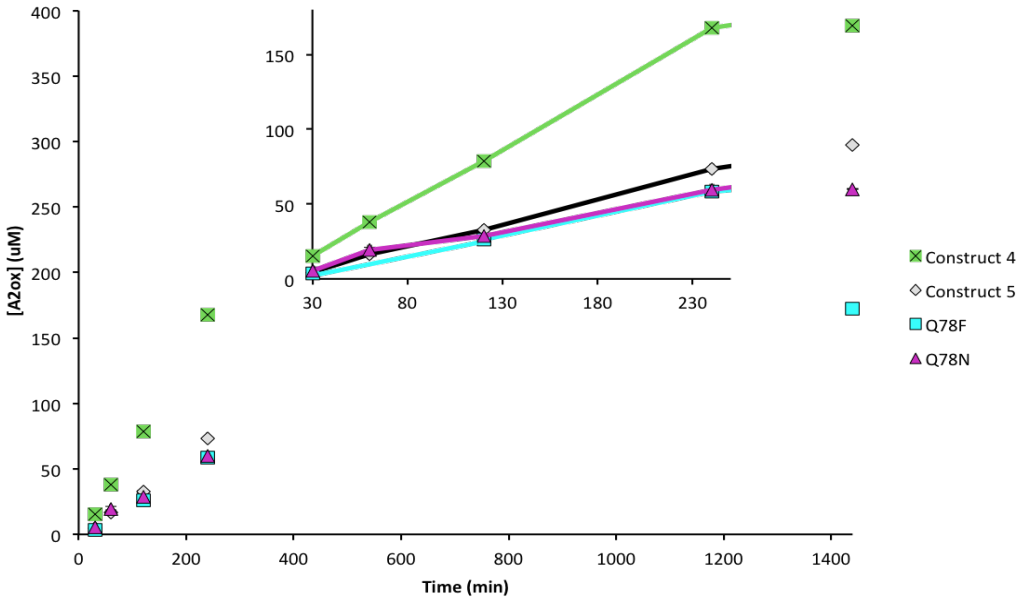


Figure 4.23. Activity of *Cj*AA10A variants and mutants on β-chitin. The graphs show the release of oxidized chito-oligosaccharide (expressed as $A2^{ox}$) from β-chitin by the different *Cj*AA10A constructs (1 µM) in presence of ascorbic acid (1 mM). Reactions were carried out in sodium phosphate buffer (50 mM, pH 7.0) at 40 ˚C with shaking (1000 rpm). The graph shows *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2 (construct 4) and *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag (construct 5). In addition, the mutants Q78F (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78F) and Q78N (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CBM2-HisTag-Q78N) are shown. n=3.

5. DISCUSSION

The essential boosting effect of LPMOs during polysaccharide degradation, have placed these enzymes at the center of worldwide research conducted on biomass conversion. Getting insights into the evolutionary divergence in substrate specificity amongst the AA10 family of LPMOs would allow for a to better understanding of the molecular basis of LPMO catalysis but also to improve the predictive capacity to link an LPMO sequence to its phenotype.

The main objective of this study was to engineer a platform enzyme, the chitin-active *Cj*AA10A with an exchange of the wild-type chitin-specific carbohydrate binding modules (CBMs) and linker for a cellulose-active LPMOs CBM and linker. The resulting hybrid platform enzyme was then subjected to site-directed mutagenesis by targeting the "hotspot" residues identified via a so-called "correlated mutation analysis" CMA.

Analysis of CBM distribution amongst AA10s LPMO provided us with a good starting hint. Indeed, while most chitin-active AA10s do not possess any CBM, or have chitin-specific CBMs for a few, it is clear that CelS2-like AA10s acquired cellulose-binding CBMs (CBM2 mainly). The importance of CBMs in LPMO operational stability and the presence of chitin-binding domains in native *Cj*AA10A justified why we first set off to engineer the architecture of the model enzyme.

After successfully cloning, expression and purification of *Cj*AA10A variants with swapped CBMs and linker, a 6x poly-histidine tag was added to the engineered platform enzyme, *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2. Several attempts were conducted before the successful addition of the His-Tag insertion. The first attempts to get a plasmid and insert with the desired change were employed by an "overhang" PCR strategy using the QuikChange II Site-Directed Mutagenesis Kit (section 3.2.1.), but the efficacy of this method proved to be low, as no clones were successfully created after homologues recombination. A new attempt was conducted using the iProof^{TM} High-Fidelity Master Mix Kit (section 3.3.2.) were the annealing temperature was slightly higher. This attempt resulted in an addition of a 5x poly-histidine tag on the platform enzyme used as a template (*Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2). As a final attempt to get the 6x His-Tag, a fresh batch of the template enzyme was used

alongside the iProof™ High-Fidelity Master Mix Kit with the same conditions as the previous attempt. Luckily, this final attempt allowed successful amplification and subsequent cloning and expression of the target *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2-HisTag enzyme.

However, this enzyme showed problems of proteolytic cleavage between the CBM and the catalytic domain, indicating presence of protease. To tackle this problem, the use of anti-protease during IMAC followed by IEC was necessary. In the end, the use of a His-Tag might not have been the best option considering that the resulting enzyme required two purification steps, ant thus no time was saved.

Experiments were set up to investigate the binding affinity of *Cj*AA10A variants towards α-chitin, β-chitin and Avicel. The percentage of free protein was determined by measuring the reduction in protein concentration (A_{280}) over time, in absence of an electron donor. A slight increase of the protein concentration measured over time showed that the LPMO domain (*Cj*AA10A^{cd}) binds weakly and slowly to α-chitin and β-chitin, but also weakly to cellulose (i.e. Avicel) (**Figure 4.12 A**). These results are consistent with previous experiments conducted on this domain (Forsberg et al., 2016). The truncated version (*Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5) showed binding to β-chitin and Avicel, but not to α-chitin (**Figure 4.12 B**). Previous studies of the CBM5 domain have shown strong binding to the two chitin substrates, and weak binding to the cellulose substrate (Forsberg et al., 2016). The study reports that the CBM73 domain shows a 90 % binding of both α-chitin and β-chitin within 15 minutes of incubation, suggesting that this domain is more specific than the CBM5 since no binding to cellulose was observed. The truncated version (i.e. lacking the CBM73) may show variously binding to α-chitin, β-chitin and Avicel due to the loss of the CBM domain believed to contribute more to chitin-specificity. Different batches of substrate may also have lead to rather different binding behaviors.

The importance of the CBM2 domain for binding to cellulose was demonstrated by the constructs with the CBM5 chitin-specific domain swapped with the CBM2 cellulose-specific domain (**Figure 4.12 C** and **Figure 4.12 D**). The results indicated stronger binding affinity to Avicel than β-chitin and more so to α-chitin, which is consistent with previous work done on a cellulose-active LPMO with the CBM2 domain (*Ma*AA10B) (Forsberg et al., 2018).

Binding experiments were also conducted on the poly-histidine tagged variant (i.e. construct 5) (**Figure 4.13**). These results were expected to look similar to the results from the identical construct lacking the His-Tag (i.e. construct 4). Surprisingly, this construct showed similar binding affinity to α-chitin, but less affinity to both β-chitin and Avicel.
In retrospect, one can argue that the additional His-Tag on the platform enzyme was unnecessary considering the lack of binding affinity to cellulose, alongside with the fact that two purifications steps was necessary for the His-Tagged protein, and thus no time was saved here. These proteins were pure from the His-Tag purification and the IEC purification was an added step to avoid degrading of the proteins.
With the knowledge acquired during the present project a better alternative would be to use a non-His-Tagget platform enzyme (construct 4) considering the altered binding to substrate and issues with purification.

It is commonly known that LPMOs will generate H_2O_2 when provided with a reductant and oxygen in absence of substrate (Isaksen et al., 2014; Kittl et al., 2012), and that they may become inactive due to auto-oxidation (Bissaro et al., 2017). Substrate binding prevents auto-oxidation, and it is thus not surprising that the constructs with linker and CBM from *Sc*AA10C were less stable than the *Cj*AA10A^{cd} (construct 1), *Cj*AA10A^{cd}-*Cj*Linker-*Cj*CBM5 (construct 2) and *Cj*AA10A^{cd}-*Cj*Linker-*Sc*CBM2 (construct 3).
All variants, except the Q78N mutant, produced hydrogen peroxide in various amounts (**Figure 4.14**). The results show that more modifications on the constructs (i.e. further from the *Cj*AA10A WT) decrease their ability to generate H_2O_2. The structural variation between construct 4, 5 and the Q78F mutant are low, thus the variation in H_2O_2 production are also shown to be low compared to the other constructs.

Initial activity assays demonstrated that all *Cj*AA10A variants were active on β-chitin, to different extent (**Figure 4.15**). The change in linker and CBM from chitin-specificity to cellulose-specificity did not alter the activity towards β-chitin significantly. The activity was also investigated on PASC and Avicel (**Figure 4.16** and **Figure 4.17**), here with *Sc*AA10C as a positive control and $CuSO_4$ as a negative control. As expected, the *Cj*AA10A variants did not generate any products when

acting on PASC or Avicel. This highlights the fact that engineering of the active site, and not only CBM swapping, is required to evolve the oxidative activity of the enzyme.

Loss of binding to the substrate may reduce the activity of the enzyme, as seen in previous studies on LPMOs (Crouch et al., 2016; Forsberg et al., 2014a; Forsberg et al., 2016). Indeed, these observations correspond to the results presented in the present study, where the percentage of unbound protein in solution is higher on α-chitin than on β-chitin. This reflects the activity of *Cj*AA10A variants towards α-chitin and β-chitin, where the *Cj*AA10A^{cd} (construct 1) and the constructs with *Sc*Linker/*Cj*Linker and *Sc*CMB2 (construct 3 and construct 4) show a lower amount of released oxidized chito-oligosaccharides when acting on α-chitin compared to β-chitin (**Figure 4.21** and **Figure 4.22**).

To investigate if the addition of a 6x poly-histidine had an effect on the activity towards β-chitin, the construct with and without the His-Tag was compared in a qualitative analysis (**Figure 4.23**). Even though the addition of a His-Tag should not affect the enzyme in a noticeable way, the release of oxidized chito-oligosaccharide is lower for the His-Tagged enzyme and mutants thereof than for the platform enzyme without the His-Tag addition. Although the reason for this was not further investigated, one can argue that the mutants should have been performed on the platform enzyme (construct 4) rather than the His-Tagged one (construct 5).

Engineering of a stable enzyme with cellulose binding capacity but remaining chitin oxidative specificity (e.g. construct 4) constituted a major achievement of the project, necessary for further evolution work.

In addition, the modeled structure *Cj*AA10A^{cd}-*Sc*Linker-*Sc*CMB2-HisTag was used in combination with literature studies (Forsberg et al., 2016; Forsberg et al., 2018) to select candidate residues for mutation intended to evolve *Cj*AA10A from a chitin-specific to a cellulose-specific enzyme. The mutagenesis work carried out here resulted in to mutants of the construct 5, Q78F and Q78N.

Creating the desired mutants proved to be difficult as several attempts were conducted with no success. Time constraints did not allow for the generating of more mutants. As indicated by a network of co-evolved residues, the simultaneous introduction of several mutations is likely necessary to observe a shift in substrate specificity. This is probably why the single mutations studied here were not enough.

5.1. CONCLUDING REMARKS AND PERSPECTIVES

This present study has provided some insight into the evolutionary divergence in substrate specificity amongst the family 10 of the AAs. Several variants of the *Cj*AA10A has been subjected to experiments conducted with the purpose of obtaining a better understanding of the molecular basis of LPMO catalysis and also to improve the predictive capacity to link an LPMO sequence to its phenotype. The addition of a His-Tag to the platform enzyme was believed to make further work easier, but this was not the case. The His-Tag is an insertion that should neither impact the catalytic domain itself nor the binding of the CBM (since located at the opposite side of the substrate binding surface of the CBM). Even though it seemed highly unlikely, the His-Tag might somehow interfere with substrate binding. However, this is not entirely reflected by activity tests conducted.

The present work shows that the His-Tagged construct may not be the best template for mutagenesis in the end given the altered binding to substrate and issues with purification. In future, mutations should rather be introduced on the non-His-Tagged version.

Nevertheless, we have achieved the engineering, expression and characterization of several hybrid enzymes showing that strong binding capacity to cellulose could be conferred to a chitin-active LPMO without altering too drastically its oxidative activity on the original substrate (i.e. chitin). The platform enzyme engineered during this project constitutes an excellent candidate for further mutagenesis work in order to unravel the evolutionary road between chitin and cellulose-active LPMOs.

6. REFERENCES

Abuja, P. M., Schmuck, M., Pilz, I., Tomme, P., Claeyssens, M. & Esterbauer, H. (1988). Structural and Functional Domains of Cellobiohydrolase-I from Trichoderma-Reesei - a Small-Angle X-Ray-Scattering Study of the Intact Enzyme and Its Core. *European Biophysics Journal with Biophysics Letters*, 15 (6): 339-342.

Abuja, P. M., Pilz, I., Tomme, P. & Claeyssens, M. (1989). Structural-Changes in Cellobiohydrolase I Upon Binding of a Macromolecular Ligand as Evident by Saxs Investigations. *Biochemical and Biophysical Research Communications*, 165 (2): 615-623.

Agger, J. W., Isaksen, T., Varnai, A., Vidal-Melgosa, S., Willats, W. G. T., Ludwig, R., Horn, S. J., Eijsink, V. G. H. & Westereng, B. (2014). Discovery of Lpmo Activity on Hemicelluloses Shows the Importance of Oxidative Processes in Plant Cell Wall Degradation. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (17): 6287-6292.

Amore, A., Knott, B. C., Supekar, N. T., Shajahan, A., Azadi, P., Zhao, P., Wells, L., Linger, J. G., Hobdey, S. E., Vander Wall, T. A., et al. (2017). Distinct Roles of N- and O-Glycans in Cellulase Activity and Stability. *Proceedings of the National Academy of Sciences of the United States of America*, 114 (52): 13667-13672.

Arfi, Y., Shamshoum, M., Rogachev, I., Peleg, Y. & Bayer, E. A. (2014). Integration of Bacterial Lytic Polysaccharide Monooxygenases into Designer Cellulosomes Promotes Enhanced Cellulose Degradation. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (25): 9109-9114.

Beckham, G. T., Matthews, J. F., Peters, B., Bomble, Y. J., Himmel, M. E. & Crowley, M. F. (2011). Molecular-Level Origins of Biomass Recalcitrance: Decrystallization Free Energies for Four Common Cellulose Polymorphs. *Journal of Physical Chemistry B*, 115 (14): 4118-4127.

Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. (2012). Oxidative Cleavage of Cellulose by Fungal Copper-Dependent Polysaccharide Monooxygenases. *Journal of the American Chemical Society*, 134 (2): 890-892.

Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. (2015). Cellulose Degradation by Polysaccharide Monooxygenases. In Kornberg, R. D. (ed.) Annual Review of Biochemistry, vol. 84 *Annual Review of Biochemistry, Vol 84*, pp. 923-946.

Beier, S. & Bertilsson, S. (2013). Bacterial Chitin Degradation-Mechanisms and Ecophysiological Strategies. *Frontiers in Microbiology*, 4.

Bertini, L., Breglia, R., Lambrughi, M., Fantucci, P., De Gioia, L., Borsari, M., Sola, M., Bortolotti, C. A. & Bruschi, M. (2018). Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. *Inorganic Chemistry*, 57 (1): 86-97.

Bissaro, B., Rohr, A. K., Muller, G., Chylenski, P., Skaugen, M., Forsberg, Z., Horn, S. J., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2017). Oxidative Cleavage of Polysaccharides by Monocopper Enzymes Depends on H2o2. *Nature Chemical Biology*, 13 (10): 1123-+.

Bissaro, B., Isaksen, I., Vaaje-Kolstad, G., Eijsink, V. G. & Rohr, A. K. (2018). How a Lytic Polysaccharide Monooxygenase Binds Crystalline Chitin. *Biochemistry*, 57 (12): 1893–1906.

Blake, A. W., McCartney, L., Flint, J. E., Bolam, D. N., Boraston, A. B., Gilbert, H. J. & Knox, J. P. (2006). Understanding the Biological Rationale for the Diversity of Cellulose-Directed Carbohydrate-Binding Modules in Prokaryotic Enzymes. *Journal of Biological Chemistry*, 281 (39): 29321-29329.

Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P. & Gilbert, H. J. (1998). Pseudomonas Cellulose-Binding Domains Mediate Their Effects by Increasing Enzyme Substrate Proximity. *Biochemical Journal*, 331: 775-781.

Book, A. J., Yennamalli, R. M., Takasuka, T. E., Currie, C. R., Phillips, G. N. & Fox, B. G. (2014). Evolution of Substrate Specificity in Bacterial Aa10 Lytic Polysaccharide Monooxygenases. *Biotechnology for Biofuels*, 7.

Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). Carbohydrate-Binding Modules: Fine-Tuning Polysaccharide Recognition. *Biochemical Journal*, 382: 769-781.

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*, 72 (1-2): 248-254.

Brown, R. M. (2004). Cellulose Structure and Biosynthesis: What Is in Store for the 21st Century? *Journal of Polymer Science Part a-Polymer Chemistry*, 42 (3): 487-495.

Busk, P. K. & Lange, L. (2015). Classification of Fungal and Bacterial Lytic Polysaccharide Monooxygenases. *Bmc Genomics*, 16.

Cannella, D., Mollers, K. B., Frigaard, N. U., Jensen, P. E., Bjerrum, M. J., Johansen, K. S. & Felby, C. (2016). Light-Driven Oxidation of Polysaccharides by Photosynthetic Pigments and a Metalloenzyme. *Nature Communications*, 7.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active Enzymes Database (Cazy): An Expert Resource for Glycogenomics. *Nucleic Acids Research*, 37: D233-D238.

Chen, X. Y., Zaro, J. L. & Shen, W. C. (2013). Fusion Protein Linkers: Property, Design and Functionality. *Advanced Drug Delivery Reviews*, 65 (10): 1357-1369.

Cocinero, E. J., Gamblin, D. P., Davis, B. G. & Simons, J. P. (2009). The Building Blocks of Cellulose: The Intrinsic Conformational Structures of Cellobiose, Its Epimer, Lactose, and Their Singly Hydrated Complexes. *Journal of the American Chemical Society*, 131 (31): 11117-11123.

Couturier, M., Ladeveze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., Villares, A., Cathala, B., Chaspoul, F., Frandsen, K. E., et al. (2018). Lytic Xylan Oxidases from Wood-Decay Fungi Unlock Biomass Degradation. *Nature Chemical Biology*, 14 (3): 306-+.

Cragg, S. M., Beckham, G. T., Bruce, N. C., Bugg, T. D. H., Distel, D. L., Dupree, P., Etxabe, A. G., Goodell, B. S., Jellison, J., McGeehan, J. E., et al. (2015). Lignocellulose Degradation Mechanisms across the Tree of Life. *Current Opinion in Chemical Biology*, 29: 108-119.

Crasto, C. J. & Feng, J. A. (2000). Linker: A Program to Generate Linker Sequences for Fusion Proteins. *Protein Engineering*, 13 (5): 309-312.

Crouch, L. I., Labourel, A., Walton, P. H., Davies, G. J. & Gilbert, H. J. (2016). The Contribution of Non-Catalytic Carbohydrate Binding Modules to the Activity

of Lytic Polysaccharide Monooxygenases. *Journal of Biological Chemistry*, 291 (14): 7439-49.

Cuskin, F., Flint, J. E., Gloster, T. M., Morland, C., Basle, A., Henrissat, B., Coutinho, P. M., Strazzulli, A., Solovyova, A. S., Davies, G. J., et al. (2012). How Nature Can Exploit Nonspecific Catalytic and Carbohydrate Binding Modules to Create Enzymatic Specificity. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (51): 20889-20894.

Duchesne, L. C. & Larson, D. W. (1989). Cellulose and the Evolution of Plant Life. *Bioscience*, 39 (4): 238-241.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A. C., Stenstrom, Y., MacKenzie, A., Sorlie, M., Horn, S. J. & Eijsink, V. G. H. (2011). Cleavage of Cellulose by a Cbm33 Protein. *Protein Science*, 20 (9): 1479-1483.

Forsberg, Z., Mackenzie, A. K., Sorlie, M., Rohr, A. K., Helland, R., Arvai, A. S., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2014a). Structural and Functional Characterization of a Conserved Pair of Bacterial Cellulose-Oxidizing Lytic Polysaccharide Monooxygenases. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (23): 8446-8451.

Forsberg, Z., Rohr, A. K., Mekasha, S., Andersson, K. K., Eijsink, V. G. H., Vaaje-Kolstad, G. & Sorlie, M. (2014b). Comparative Study of Two Chitin-Active and Two Cellulose-Active Aa10-Type Lytic Polysaccharide Monooxygenases. *Biochemistry*, 53 (10): 1647-1656.

Forsberg, Z., Nelson, C. E., Dalhus, B., Mekasha, S., Loose, J. S. M., Crouch, L. I., Rohr, A. K., Gardner, J. G., Eijsink, V. G. H. & Vaaje-Kolstad, G. (2016). Structural and Functional Analysis of a Lytic Polysaccharide Monooxygenase Important for Efficient Utilization of Chitin in Cellvibrio Japonicus. *Journal of Biological Chemistry*, 291 (14): 7300-7312.

Forsberg, Z., Bissaro, B., Gullesen, J., Dalhus, B., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2018). Structural Determinants of Bacterial Lytic Polysaccharide Monooxygenase Functionality. *Journal of Biological Chemistry*, 293 (4): 1397-1412.

Frandsen, K. E. H., Simmons, T. J., Dupree, P., Poulsen, J. C. N., Hemsworth, G. R., Ciano, L., Johnston, E. M., Tovborg, M., Johansen, K. S., von Freiesleben, P., et al. (2016). The Molecular Basis of Polysaccharide Cleavage by Lytic Polysaccharide Monooxygenases. *Nature Chemical Biology*, 12 (4): 298-+.

Frommhagen, M., Sforza, S., Westphal, A. H., Visser, J., Hinz, S. W. A., Koetsier, M. J., van Berkel, W. J. H., Gruppen, H. & Kabel, M. A. (2015). Discovery of the Combined Oxidative Cleavage of Plant Xylan and Cellulose by a New Fungal Polysaccharide Monooxygenase. *Biotechnology for Biofuels*, 8.

Fuchs, R. L., McPherson, S. A. & Drahos, D. J. (1986). Cloning of a Serratia-Marcescens Gene Encoding Chitinase. *Applied and Environmental Microbiology*, 51 (3): 504-509.

Faaij, A. P. C. (2006). Bio-Energy in Europe: Changing Technology Choices. *Energy Policy*, 34 (3): 322-342.

Gao, D. H., Chundawat, S. P. S., Sethi, A., Balan, V., Gnanakaran, S. & Dale, B. E. (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.

Gardner, J. G. & Keating, D. H. (2010). Requirement of the Type Ii Secretion System for Utilization of Cellulosic Substrates by Cellvibrio Japonicus. *Applied and Environmental Microbiology*, 76 (15): 5079-5087.

Gardner, K. H. & Blackwell, J. (1975). Refinement of Structure of Beta-Chitin. *Biopolymers*, 14 (8): 1581-1595.

Garland, N. L., Papageorgopoulos, D. C., Stanford, J. M. & Elsevier. (2012). Hydrogen and Fuel Cell Technology: Progress, Challenges, and Future Directions. In Energy Procedia, vol. 28 *Fuel Cells 2012 Science & Technology - a Grove Fuel Cell Event*, pp. 2-11.

Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M. H., Vaaje-Kolstad, G., Lundberg, D., Royant, A., Stahlberg, J., Eijsink, V. G. H., et al. (2014). Structural and Electronic Snapshots During the Transition from a Cu(Ii) to Cu(I) Metal Center of a Lytic Polysaccharide Monooxygenase by X-Ray Photoreduction. *Journal of Biological Chemistry*, 289 (27): 18782-18792.

Hallac, B. B. & Ragauskas, A. J. (2011). Analyzing Cellulose Degree of Polymerization and Its Relevancy to Cellulosic Ethanol. *Biofuels Bioproducts & Biorefining-Biofpr*, 5 (2): 215-225.

Hansson, H., Karkehabadi, S., Mikkelsen, N., Douglas, N. R., Kim, S., Lam, A., Kaper, T., Kelemen, B., Meier, K. K., Jones, S. M., et al. (2017). High-Resolution Structure of a Lytic Polysaccharide Monooxygenase from Hypocrea Jecorina Reveals a Predicted Linker as an Integral Part of the Catalytic Domain. *Journal of Biological Chemistry*, 292 (46): 19099-19109.

Harris, P. V., Welner, D., McFarland, K. C., Re, E., Poulsen, J. C. N., Brown, K., Salbo, R., Ding, H. S., Vlasenko, E., Merino, S., et al. (2010). Stimulation of Lignocellulosic Biomass Hydrolysis by Proteins of Glycoside Hydrolase Family 61: Structure and Function of a Large, Enigmatic Family. *Biochemistry*, 49 (15): 3305-3316.

Healthcare, G. (2018). *Ion Exchange Chromatography Principles and Methods*. Available at: https://www.gelifesciences.com/en/bs/solutions/protein-research/knowledge-center/protein-handbooks (accessed: 19/4-18).

Hein, L. & Leemans, R. (2012). The Impact of First-Generation Biofuels on the Depletion of the Global Phosphorus Reserve. *Ambio*, 41 (4): 341-349.

Hemsworth, G. R., Davies, G. J. & Walton, P. H. (2013). Recent Insights into Copper-Containing Lytic Polysaccharide Mono-Oxygenases. *Current Opinion in Structural Biology*, 23 (5): 660-668.

Hemsworth, G. R., Henrissat, B., Davies, G. J. & Walton, P. H. (2014). Discovery and Characterization of a New Family of Lytic Polysaccharide Monooxygenases. *Nature Chemical Biology*, 10 (2): 122-126.

Hemsworth, G. R., Dejean, G., Davies, G. J. & Brumer, H. (2016). Learning from Microbial Strategies for Polysaccharide Degradation. *Biochemical Society Transactions*, 44: 94-108.

Henrissat, B., Driguez, H., Viet, C. & Schulein, M. (1985). Synergism of Cellulases from Trichoderma-Reesei in the Degradation of Cellulose. *Bio-Technology*, 3 (8): 722-726.

Hervé, C., Rogowski, A., Blake, A. W., Marcus, S. E., Gilbert, H. J. & Knox, J. P. (2010). Carbohydrate-Binding Modules Promote the Enzymatic Deconstruction of Intact Plant Cell Walls by Targeting and Proximity Effects. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (34): 15293-15298.

Hill, P. J. & Stewart, G. S. A. B. (1992). The Polymerase Chain Reaction in Molecular and Micro-Biology. *Biotechnology and Genetic Engineering Reviews*, 10 (1): 343-378.

Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. (2007). Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science*, 315 (5813): 804-807.

Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K. M. & Eijsink, V. G. H. (2006). Costs and Benefits of Processivity in Enzymatic Degradation of Recalcitrant Polysaccharides. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (48): 18089-18094.

Horn, S. J., Vaaje-Kolstad, G., Westereng, B. & Eijsink, V. G. H. (2012). Novel Enzymes for the Degradation of Cellulose. *Biotechnology for Biofuels*, 5.

Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K. M. & Eijsink, V. G. (2016). Costs and Benefits of Processivity in Enzymatic Degradation of Recalcitrant Polysaccharides. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (48): 18089-18094.

Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttila, M., Ando, T. & Samejima, M. (2011). Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science*, 333 (6047): 1279-1282.

Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., Ludwig, R., Haltrich, D., Eijsink, V. G. H. & Horn, S. J. (2014). A C4-Oxidizing Lytic Polysaccharide Monooxygenase Cleaving Both Cellulose and Cello-Oligosaccharides. *Journal of Biological Chemistry*, 289 (5): 2632-2642.

Johansen, K. S. (2016). Discovery and Industrial Applications of Lytic Polysaccharide Mono-Oxygenases. *Biochemical Society Transactions*, 44: 143-149.

Khoushab, F. & Yamabhai, M. (2010). Chitin Research Revisited. *Marine Drugs*, 8 (7): 1988-2012.

Kim, S., Stahlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. (2014). Quantum Mechanical Calculations Suggest That Lytic Polysaccharide Monooxygenases Use a Copper-Oxyl, Oxygen-Rebound Mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (1): 149-154.

Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. & Ludwig, R. (2012). Production of Four Neurospora Crassa Lytic Polysaccharide Monooxygenases in Pichia Pastoris Monitored by a Fluorimetric Assay. *Biotechnology for Biofuels*, 5.

Kjaergaard, C. H., Qayyum, M. F., Wong, S. D., Xu, F., Hemsworth, G. R., Walton, D. J., Young, N. A., Davies, G. J., Walton, P. H., Johansen, K. S., et al. (2014). Spectroscopic and Computational Insight into the Activation of O-2 by the Mononuclear Cu Center in Polysaccharide Monooxygenases. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (24): 8797-8802.

Kumirska, J., Czerwicka, M., Kaczynski, Z., Bychowska, A., Brzozowski, K., Thoming, J. & Stepnowski, P. (2010). Application of Spectroscopic Methods for Structural Analysis of Chitin and Chitosan. *Marine Drugs*, 8 (5): 1567-1636.

Kurita, K., Sugita, K., Kodaira, N., Hirakawa, M. & Yang, J. (2005). Preparation and Evaluation of Trimethylsilylated Chitin as a Versatile Precursor for Facile Chemical Modifications. *Biomacromolecules*, 6 (3): 1414-1418.

Leggio, L., Simmons, T. J., Poulsen, J. C. N., Frandsen, K. E. H., Hemsworth, G. R., Stringer, M. A., von Freiesleben, P., Tovborg, M., Johansen, K. S., De Maria,

L., et al. (2015). Structure and Boosting Activity of a Starch-Degrading Lytic Polysaccharide Monooxygenase. *Nature Communications*, 6.

Lehtiö, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M. & Teeri, T. T. (2003). The Binding Specificity and Affinity Determinants of Family 1 and Family 3 Cellulose Binding Modules. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (2): 484-489.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. & Henrissat, B. (2013). Expansion of the Enzymatic Repertoire of the Cazy Database to Integrate Auxiliary Redox Enzymes. *Biotechnology for Biofuels*, 6.

Ljungdahl, L. G. & Eriksson, K. E. (1985). Ecology of Microbial Cellulose Degradation. *Advances in Microbial Ecology*, 8: 237-299.

Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. & Henrissat, B. (2014). The Carbohydrate-Active Enzymes Database (Cazy) in 2013. *Nucleic Acids Research*, 42 (D1): D490-D495.

Loose, J. S., Forsberg, Z., Fraaije, M. W., Eijsink, V. G. & Vaaje-Kolstad, G. (2014). A Rapid Quantitative Activity Assay Shows That the Vibrio Cholerae Colonization Factor Gbpa Is an Active Lytic Polysaccharide Monooxygenase. *FEBS Lett*, 588 (18): 3435-40.

Manoil, C. & Beckwith, J. (1986). A Genetic Approach to Analyzing Membrane-Protein Topology. *Science*, 233 (4771): 1403-1408.

Medronho, B., Romano, A., Miguel, M. G., Stigsson, L. & Lindman, B. (2012). Rationalizing Cellulose (in)Solubility: Reviewing Basic Physicochemical Aspects and Role of Hydrophobic Interactions. *Cellulose*, 19 (3): 581-587.

Merino, S. T. & Cherry, J. (2007). Progress and Challenges in Enzyme Development for Biomass Utilization. In Olsson, L. (ed.) Advances in Biochemical Engineering-Biotechnology, vol. 108 *Biofuels*, pp. 95-120.

Minke, R. & Blackwell, J. (1978). Structure of Alpha-Chitin. *Journal of Molecular Biology*, 120 (2): 167-181.

Montanier, C., van Bueren, A. L., Dumon, C., Flint, J. E., Correia, M. A., Prates, J. A., Firbank, S. J., Lewis, R. J., Grondin, G. G., Ghinet, M. G., et al. (2009). Evidence That Family 35 Carbohydrate Binding Modules Display Conserved Specificity but Divergent Function. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (9): 3065-3070.

Morag, E., Lapidot, A., Govorko, D., Lamed, R., Wilchek, M., Bayer, E. A. & Shoham, Y. (1995). Expression, Purification, and Characterization of the Cellulose-Binding Domain of the Scaffoldin Subunit from the Cellulosome of Clostridium-Thermocellum. *Applied and Environmental Microbiology*, 61 (5): 1980-1986.

O'Sullivan, A. C. (1997). Cellulose: The Structure Slowly Unravels. *Cellulose*, 4 (3): 173-207.

OMEGA BIO-TEK. (2018). *E.Z.N.A.® Plasmid Mini Kit I, V(Capped) Spin*. Available at: http://omegabiotek.com/store/product/hp-plasmid-mini-kit-1-vcapped-spin/ (accessed: 19/4-18).

Payne, C. M., Resch, M. G., Chen, L. Q., Crowley, M. F., Himmel, M. E., Taylor, L. E., Sandgren, M., Stahlberg, J., Stals, I., Tan, Z. P., et al. (2013). Glycosylated Linkers in Multimodular Lignocellulose-Degrading Enzymes Dynamically Bind to Cellulose. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (36): 14646-14651.

Payne, C. M., Knott, B. C., Mayes, H. B., Hansson, H., Himmel, M. E., Sandgren, M., Stahlberg, J. & Beckham, G. T. (2015). Fungal Cellulases. *Chemical Reviews*, 115 (3): 1308-1448.

Pell, G., Szabo, L., Charnock, S. J., Xie, H. F., Gloster, T. M., Davies, G. J. & Gilbert, H. J. (2004). Structural and Biochemical Analysis of Cellvibrio Japonicus Xylanase 10c - How Variation in Substrate-Binding Cleft Influences the Catalytic Profile of Family Gh-10 Xylanases. *Journal of Biological Chemistry*, 279 (12): 11777-11788.

Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. (2011). Cellobiose Dehydrogenase and a Copper-Dependent Polysaccharide Monooxygenase Potentiate Cellulose Degradation by Neurospora Crassa. *Acs Chemical Biology*, 6 (12): 1399-1406.

Quinlan, J. R., Sweeney, M. D., Leggio, L., Otten, H., Poulsen, J. C. N., Johansen, K. S., Krogh, K. B. R. M., Jorgensen, C. I., Tovborg, M., Anthonsen, A., et al. (2011). Insights into the Oxidative Degradation of Cellulose by a Copper Metalloenzyme That Exploits Biomass Components. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (37): 15079-15084.

Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C. N., Johansen, K. S., Krogh, K., Jorgensen, C. I., Tovborg, M., Anthonsen, A., et al. (2011). Insights into the Oxidative Degradation of Cellulose by a Copper Metalloenzyme That Exploits Biomass Components. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (37): 15079-15084.

Reese, E. T., Siu, R. G. H. & Levinson, H. S. (1950). The Biological Degradation of Soluble Cellulose Derivatives and Its Relationship to the Mechanism of Cellulose Hydrolysis. *Journal of Bacteriology*, 59 (4): 485-497.

Resch, M. G., Donohoe, B. S., Baker, J. O., Decker, S. R., Bayer, E. A., Beckham, G. T. & Himmel, M. E. (2013). Fungal Cellulases and Complexed Cellulosomal Enzymes Exhibit Synergistic Mechanisms in Cellulose Deconstruction. *Energy & Environmental Science*, 6 (6): 1858-1867.

Ryan, L., Convery, F. & Ferreira, S. (2006). Stimulating the Use of Biofuels in the European Union: Implications for Climate Change Policy. *Energy Policy*, 34 (17): 3184-3194.

Rye, C. S. & Withers, S. G. (2000). Glycosidase Mechanisms. *Current Opinion in Chemical Biology*, 4 (5): 573-580.

Sabbadin, F., Hemsworth, G. R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R. D. S., Sweeney, M. D., Besser, K., Elias, L., et al. (2018). An Ancient Family of Lytic Polysaccharide Monooxygenases with Roles in Arthropod Development and Biomass Digestion. *Nature Communications*, 9.

Sammond, D. W., Payne, C. M., Brunecky, R., Himmel, M. E., Crowley, M. F. & Beckham, G. T. (2012). Cellulase Linkers Are Optimized Based on Domain Type and Function: Insights from Sequence Analysis, Biophysical Measurements, and Molecular Simulation. *Plos One*, 7 (11).

Schmuck, M., Pilz, I., Hayn, M. & Esterbauer, H. (1986). Investigation of Cellobiohydrolase from Trichoderma-Reesei by Small-Angle X-Ray-Scattering. *Biotechnology Letters*, 8 (6): 397-402.

Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. (2017). The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. *Acs Chemical Biology*, 12 (4): 1095-1103.

Srisodsuk, M., Reinikainen, T., Penttila, M. & Teeri, T. T. (1993). Role of the Interdomain Linker Peptide of Trichoderma-Reesei Cellobiohydrolase-I in Its Interaction with Crystalline Cellulose. *Journal of Biological Chemistry*, 268 (28): 20756-20761.

Suzuki, K., Sugawara, N., Suzuki, M., Uchiyama, T., Katouno, F., Nikaidou, N. & Watanabe, T. (2002). Chitinases a, B, and C1 of Serratia Marcescens 2170 Produced by Recombinant Escherichia Coli: Enzymatic Properties and Synergism on Chitin Degradation. *Bioscience Biotechnology and Biochemistry*, 66 (5): 1075-1083.

Swatloski, R. P., Spear, S. K., Holbrey, J. D. & Rogers, R. D. (2002). Dissolution of Cellose with Ionic Liquids. *Journal of the American Chemical Society*, 124 (18): 4974-4975.

Taylor, L. E., Knott, B. C., Baker, J. O., Alahuhta, P. M., Hobdey, S. E., Linger, J. G., Lunin, V. V., Amore, A., Subramanian, V., Podkaminer, K., et al. (2018). Engineering Enhanced Cellobiohydrolase Activity. *Nature Communications*, 9.

Technologies, A. (2018). *Quikchange Ii Site-Directed Mutagenesis Kit*. Available at: https://www.agilent.com/cs/library/usermanuals/Public/200523.pdf (accessed: 19/4-18).

Tharanathan, R. N. & Kittur, F. S. (2003). Chitin - the Undisputed Biomolecule of Great Potential. *Critical Reviews in Food Science and Nutrition*, 43 (1): 61-87.

Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y. & Steitz, T. A. (1996). Crystal Structure of a Bacterial Family-Iii Cellulose-Binding Domain: A General Mechanism for Attachment to Cellulose. *Embo Journal*, 15 (21): 5739-5751.

Tuveng, T. R., Arntzen, M. O., Bengtsson, O., Gardner, J. G., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2016). Proteomic Investigation of the Secretome of Cellvibrio Japonicus During Growth on Chitin. *Proteomics*, 16 (13): 1904-1914.

Uversky, V. N. & Dunker, A. K. (2012). Multiparametric Analysis of Intrinsically Disordered Proteins: Looking at Intrinsic Disorder through Compound Eyes. *Analytical Chemistry*, 84 (5): 2096-2104.

Várnai, A., Siika-aho, M. & Viikari, L. (2013). Carbohydrate-Binding Modules (Cbms) Revisited: Reduced Amount of Water Counterbalances the Need for Cbms. *Biotechnology for Biofuels*, 6.

Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. (2014a). Determinants of Regioselective Hydroxylation in the Fungal Polysaccharide Monooxygenases. *Journal of the American Chemical Society*, 136 (2): 562-565.

Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R. & Marletta, M. A. (2014b). A Family of Starch-Active Polysaccharide Monooxygenases. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (38): 13822-13827.

Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M. F., Synstad, B. & Eijsink, V. G. H. (2005). The Non-Catalytic Chitin-Binding Protein Cbp21 from Serratia Marcescens Is Essential for Chitin Degradation. *Journal of Biological Chemistry*, 280 (31): 28492-28497.

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z. L., Zhai, H., Sorlie, M. & Eijsink, V. G. H. (2010). An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science*, 330 (6001): 219-222.

Vaaje-Kolstad, G., Bohle, L. A., Gaseidnes, S., Dalhus, B., Bjoras, M., Mathiesen, G. & Eijsink, V. G. H. (2012). Characterization of the Chitinolytic Machinery of Enterococcus Faecalis V583 and High-Resolution Structure of Its Oxidative Cbm33 Enzyme. *Journal of Molecular Biology*, 416 (2): 239-254.

Vaaje-Kolstad, G., Horn, S. J., Sorlie, M. & Eijsink, V. G. H. (2013). The Chitinolytic Machinery of Serratiamarcescens - a Model System for Enzymatic Degradation of Recalcitrant Polysaccharides. *Febs Journal*, 280 (13): 3028-3049.

Vaaje-Kolstad, G., Forsberg, Z., Loose, J. S. M., Bissaro, B. & Eijsink, V. G. H. (2017). Structural Diversity of Lytic Polysaccharide Monooxygenases. *Current Opinion in Structural Biology*, 44: 67-76.

Walton, P. H. & Davies, G. J. (2016). On the Catalytic Mechanisms of Lytic Polysaccharide Monooxygenases. *Current Opinion in Chemical Biology*, 31: 195-207.

Wang, B. J., Johnston, E. M., Li, P. F., Shaik, S., Davies, G. J., Walton, P. H. & Rovira, C. (2018). Qm/Mm Studies into the H2o2-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. *ACs Catalysis*, 8 (2): 1346-1351.

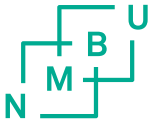
Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G. H., Igarashi, K., Samejima, M., Stahlberg, J., Horn, S. J. & Sandgren, M. (2011). The Putative Endoglucanase Pcgh61d from Phanerochaete Chrysosporium Is a Metal-Dependent Oxidative Enzyme That Cleaves Cellulose. *Plos One*, 6 (11).

Younes, I. & Rinaudo, M. (2015). Chitin and Chitosan Preparation from Marine Sources. Structure, Properties and Applications. *Marine Drugs*, 13 (3): 1133-1174.

Zhao, Y., Park, R. D. & Muzzarelli, R. A. A. (2010). Chitin Deacetylases: Properties and Applications. *Marine Drugs*, 8 (1): 24-46.

Aachmann, F. L., Sorlie, M., Skjak-Braek, G., Eijsink, V. G. H. & Vaaje-Kolstad, G. (2012). Nmr Structure of a Lytic Polysaccharide Monooxygenase Provides Insight into Copper Binding, Protein Dynamics, and Substrate Interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (46): 18779-18784.

Aam, B. B., Heggset, E. B., Norberg, A. L., Sorlie, M., Varum, K. M. & Eijsink, V. G. H. (2010). Production of Chitooligosaccharides and Their Potential Applications in Medicine. *Marine Drugs*, 8 (5): 1482-1517.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway