Norwegian University
of Life Sciences

Master's Thesis 2016 60 ECTS
Department of Chemistry, Biotechnology and Food Science (IKBM)

Site-directed mutagenesis of a lytic polysaccharide monooxygenase from *Micromonospora aurantiaca*

Jonathan Gullesen

Site-directed mutagenesis of a lytic polysaccharide monooxygenase from *Micromonospora aurantiaca*

Masters Thesis

Jonathan Gullesen

Protein Engineering and Proteomics Group

Department of Chemistry, Biotechnology and Food Science

The Norwegian University of Life Sciences

2016

Acknowledgements

The present work was carried out at the Department of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences with Dr. Gustav Vaaje-Kolstad and Dr. Zarah Forsberg as supervisors.

I would like to express my very great appreciation to my supervisor, Dr. Zarah Forsberg for her excellent guidance throughout this work. Her inspiring thoughts, encouragement and support has been of great importance to me. I also wish to thank her for reading and commenting on this work.

I would also like to thank Dr. Gustav Vaaje-Kolstad for his ideas, guidance and comments during the writing process of this thesis. I would like to express my deep gratitude to Professor Vincent Eijsink for giving me the opportunity to write my thesis in the Protein Engineering and Proteomics (PEP) group.

I wish to thank the people of the PEP group for their insightful conversations and a great time.

Last but not least, I am forever thankful for the love and support of my dearest girlfriend and our lovely daughter. Their patience and support throughout this work have been of essential importance. I wish to thank my extended family their encouragement and support.

Ås, May 18th 2016

Jonathan Gullesen

Abstract

Cellulose and chitin are the two most abundant biopolymers in nature and they are valuable biomaterials that supplements fossil resources in the production of fuels, materials and chemicals. Due to the recalcitrant nature of these polysaccharides, efficient conversion into soluble sugars is of major importance for a sustainable bio-economy in the future. Traditionally, enzymatic degradation of cellulose and chitin were thought to rely on the synergistic action of hydrolytic enzymes, but the recent discovery of lytic polysaccharide monooxygenases (LPMOs) shows that these oxidative enzymes are important contributors to the depolymerization process. Cellulose-active AA10 LPMOs cleaves glycosidic bonds in the cellulose backbone by either specific C1 or C1/C4 oxidation. This oxidative regioselectivity is thought to be a result of different positioning of the LPMO on the cellulose, but the mechanisms behind the differences in regioselectivity is unknown. Understanding the mechanism of oxidative regioselectivity in LPMOs is important for both fundamental and applied reasons where optimization of enzyme cocktails is a major issue for biorefining of biomass.

This study identifies a pair of amino acids in the sequence of LPMO10B from the bacterium *Micromonospora aurantiaca* that are likely to play an important role in the specificity of the protein's oxidative regioselectivity. Highly conserved positions in the sequences of specific C1 oxidizing cellulose-active AA10 proteins were targeted for site-directed mutagenesis, and the C4 oxidizing activity of a *Ma*LPMO10B variant, carrying two mutated residues, was almost completely lost. The finding suggests that both residues are highly important for the LPMO to carry out C4 oxidizing activity on cellulosic substrates.

Characterization of two *M. aurantiaca* LPMOs were conducted: *Ma*LPMO10B and *Ma*LPMO10D. Both contain a CBM2 domain and are active on both cellulose (C1/C4 oxidizers) and chitin. The structure of *Ma*LPMO10B was determined by X-ray crystallography to 1.08 Å. Structural comparison to other C1/C4 oxidizing cellulose-active AA10 proteins revealed a prolonged loop, comprising eight additional residues, near the substrate-binding surface of *Ma*LPMO10B.

From the research that has been conducted, it is possible to conclude that residues W82 and N85 play an important role in the determination of oxidative regioselectivity in *Ma*LPMO10B. Further

research will be necessary to identify additional residues to possibly eliminate the C4 oxidizing activity of the enzyme completely.

Sammendrag

Cellulose og kitin er de to mest forekommende biopolymerene i naturen, og de er verdifulle biomaterialer som supplerer fossile ressurser i produksjonen av drivstoff, materialer og kjemikaler. Effektiv konvertering av disse vanskelig nedbrytbare polysakkaridene til løselige sukkerarter er av stor viktighet for en fremtidig bærekraftig bioøkonomi. Det tradisjonelle synet på enzymatisk nedbryting av cellulose og kitin har innebåret synergistisk handling mellom hydrolytiske enzymer, men den nylige oppdagelsen av lytiske polysakkaridmonooksygenaser (LPMOer) viser at disse oksidative enzymene er viktige bidragsytere til depolymeriseringsprosessen. Cellulose-aktive AA10 LPMOer kløyver glykosidiske bånd i cellulose-kjedene ved å enten oksidere kun C1-karbonet eller både C1 og C4 karbonet. Denne oksidative regioselektiviteten mistenkes å være et resultat av ulik posisjonering av LPMOen på cellulose kjeden, men mekanismene bak dette er ukjent. Å forstå mekanismen bak oksidativ regioselektivitet i LPMOer er viktig for både fundamentale og anvendte årsaker der optimalisering av enzym-cocktails er av stor viktighet for bioraffinering av biomasse.

Denne studien identifiserer et par aminosyrer i sekvensen til LPMO10B fra bakterien *Micromonospora aurantiaca* som er antatt å spille en viktig rolle i spesifiseringen av proteinets oksidative regioselektivitet. Høyt konserverte posisjoner i sekvensen til spesifikt C1-oksiderende cellulose-aktive AA10 proteiner var målrettet for seterettet mutagenese, og den C4-oksiderende aktiviteten til en *Ma*LPMO10B-variant, som inneholdt to muterte residuer, var nesten helt fjernet. Dette funnet antyder at begge residuene er meget viktige for at LPMOen kan utføre C4-oksiderende aktivitet på cellulose-substrater.

Karakterisering av to LPMOer fra *M. aurantiaca* ble utført: *Ma*LPMO10B og *Ma*LPMO10D. Begge inneholdt et CBM2-domene, og begge var aktive på både cellulose (C1/C4-oksiderende) og kitin. Strukturen til *Ma*LPMO10B ble bestemt ved røntgen krystallografi til 1.08 Å. Strukturell sammenlikning med andre C1/C4-oksiderende cellulose-aktive AA10 proteiner avslører av *Ma*LPMO10B inneholder en forlenget loop, bestående av åtte ekstra residuer, nær den substratbindende overflaten.

Fra forskningen som har blitt utført i denne studien kan det trolig konkluderes at residuene W82 og N85 spiller en viktig rolle i bestemmelsen av oksidativ regioselektivitet hos *Ma*LPMO10B.

Videre forskning vil være nødvendig for å identifisere ytterligere residuer for å muligens eliminere den C4-oksiderende aktiviteten til enzymet fullstendig.

Abbreviations

A280	Absorbance of ultraviolet light at 280 nm.
aa	Amino acids
AA	Auxiliary activities
bp	Base pair
CAZY	Carbohydrate-Active Enzymes
CBM	Carbohydrate-Binding Module
CBM2a	Family 2 Carbohydrate-Binding Module
dH_2O	Sterile water (Milli-Q)
DP	Degree of polymerization
HPLC	High-Performance Liquid Chromatography
kb	Kilobases
kDa	Kilo Dalton
LB	Luria Bertani broth
LPMO	Lytic Polysaccharide Monooxygenase
MALDI-TOF	Matrix-Assisted Lased Desorption/Ionization Time of Flight
*Ma*LPMO10B, -D	LPMO10 proteins from *Micromonospora aurantiaca*, full-length protein
*Ma*LPMO10B^{cd}, -D^{cd}	LPMO10 proteins from *Micromonospora aurantiaca*, catalytic domain (truncated protein)
*Ma*LPMO10D^{sl}	LPMO10D protein from *Micromonospora aurantiaca*, shortened linker region
MS	Mass spectrometry

PASC	Phosphoric acid swollen cellulose
PCR	Polymerase chain reaction
PDB	Protein data bank
rpm	Rotations per minute
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
UV	Ultraviolet
v/v	Volume/volume
w/v	Weight/volume

Table of contents

1 Introduction

1.1 Carbohydrates

Carbohydrates represent the most abundant organic compounds in nature, and are major constituents of plants, animals and microorganisms. Carbohydrates have the general formula $C_x(H_2O)_y$, and can be classified as polyhydroxy aldehydes and ketones. The basic units of carbohydrates are monosaccharides; simple sugars (e.g. glucose, fructose), usually composed of 3 to 9 carbon atoms. Linking of two monosaccharides through a glycosidic bond makes a disaccharide (e.g. sucrose, lactose), and a polymer of 3 to 10 monosaccharides is called an oligosaccharide. Polysaccharides are composed of long chains of monosaccharides and they range in structures from linear to highly branched. Carbohydrates play a diverse but important role in living organisms (Figure 1.1). For example, mono- and disaccharides acts as fuel for metabolism and as building blocks for nucleic acids; oligosaccharides are involved in cell recognition and cell adhesion, whereas polysaccharides function as food reserves (e.g. starch and glycogen) and structural elements (e.g. cellulose and chitin).

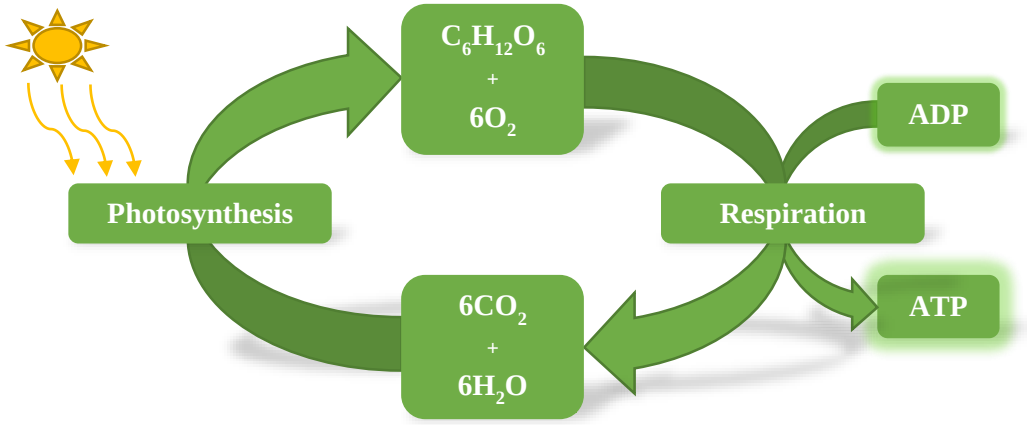


Figure 1.1. Carbon cycle. *In the process of photosynthesis, plants collect energy from the sun and use carbon dioxide and water to produce carbohydrates (e.g. cellulose and starch). Microorganisms and animals digest plants, and through the process of oxidative metabolism, carbohydrates and molecular oxygen (O_2) are converted into energy (adenosine triphosphate (ATP)), releasing carbon dioxide and water in the process. The cycle is repeated with plants reusing carbon dioxide and water during photosynthesis.*

The glycosidic bond, linking two monosaccharide units, is formed between the hemiacetal group in one saccharide and the hydroxyl group of another saccharide. Formation of the glycosidic bond is catalyzed by a group of enzymes called glycosyltransferases that synthesizes covalent bonds between saccharides that either obtain an alpha (α) or beta (β) configuration of the anomeric

carbon. The stereochemistry of the glycosidic bond is determined by the reaction mechanism of the glycosyltransferases, which either can be retaining (giving an (α)-configuration) or inverting (giving a (β)-configuration). Glycosidic bonds may be formed between various carbon atoms of two monosaccharides, but the most common linkages are 1→1, 1→2, 1→4 and 1→6. Replacement of hydroxyl groups by other functional groups, or attachment of substituents, results in modified monosaccharides. The vast number of different monosaccharides, and the many possible modifications and glycosidic linkages, results in a huge diversity among carbohydrates in nature.

The two most abundant polysaccharides in nature are cellulose and chitin. As valuable renewable biomaterials, they have both been subject to extensive research in recent years. One of the major aims of this research has been to improve and optimize their degradation, since the recalcitrant nature of these polysaccharides makes depolymerization challenging. The soluble sugars obtained by degradation of cellulose and chitin can be utilized in the production of fuels, chemicals and other materials (Hemsworth et al., 2015). Today, mainly first-generation feedstocks (e.g. sugarcane, wheat and corn) are used for the production of biofuels, organic chemicals and biomaterials (Hein & Leemans, 2012); however, due to environmental issues (e.g. palm oil production) and the reliance of food sources, a transition to using primarily second-generation feedstock (e.g. lignocellulosic biomass and chitin) is of great importance.

1.1.1 Cellulose

Cellulose is the main constituent of the plant cell wall, making it the most abundant polymer on earth(Beeson et al., 2015). It is a linear polymer of D-glucopyranose units connected by glycosidic β1→4 linkages (Figure 1.2). Each glucose unit is rotated 180° with respect to its neighbor, allowing hydrogen bonding between parallel chains. The strong inter-chain interactions generate a planar structure, and are responsible for the insolubility of cellulose in water (Medronho et al., 2012).

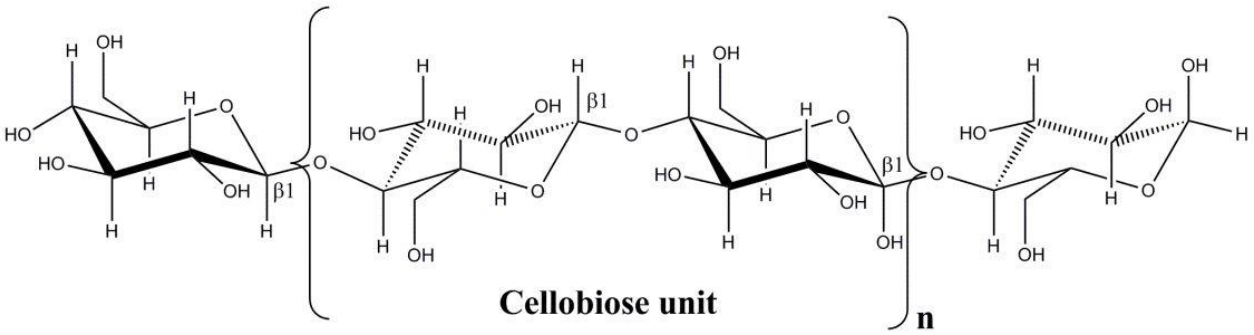


Figure 1.2. The structure of cellulose. *Cellulose is composed of several β1→4 linked glucose units, each rotated 180° with respect to its neighbor. The repeating unit in cellulose is cellobiose, which is composed of two glucose saccharides joined together. Figure source: http://www.intechopen.com/books/liquid-gaseous-and-solid-biofuels-conversion-techniques/hydrotreating-catalytic-processes-for-oxygen-removal-in-the-upgrading-of-bio-oils-and-bio-chemicals*

In higher plants, cellulose is synthesized by rosette-like cellulose synthase complexes localized at the plasma membrane. The rosettes are composed of six subunits arranged in a hexagonal structure; each subunit is presumed to consists of six cellulose synthase (CESA) proteins (S. Li et al., 2014). Each CESA protein polymerizes a β-(1,4)-glucan chain (a polymer of β-D-glucose units), and synthesized chains from each CESA protein associates to form elementary microfibrils that aggregate into larger sized crystalline cellulose microfibrils (Doblin et al., 2002). The length of a cellulose microfibril is usually described by the degree of polymerization (DP; i.e. the number of monosaccharide units in the chain), which is estimated to range from some hundreds to several thousand glucose units, depending on the source (e.g. plant cell wall component and plant species) of the cellulose microfibrils (Hallac & Ragauskas, 2011). Glucan chains shorter than DP9 are soluble in water solution; however, any longer chains will have greater affinity for other glucan chains than for the surrounding water, and will thus be insoluble (Brown, 2004). Inter- and intramolecular hydrogen bonding of hydroxyl groups present on elementary microfibrils result in cellulose microfibrils with various ordered crystallinity (Park et al., 2010). The core of cellulose microfibrils are highly crystalline, whereas regions with less intramolecular bonds are referred to as amorphous and are located on or close to the microfibril surface (Ding & Himmel, 2006).

In nature, two different crystalline forms (allomorphs) of cellulose are produced: cellulose Iα, mainly synthesized by bacteria; and cellulose Iβ, the main cellulose form produced in higher plants (Brown, 2004). Referred to as native cellulose, both forms have similar conformations, but

the stacking arrangement of hydrogen-bonded layers are different (Matthews et al., 2012). By pretreatment of cellulose I, five other allomorphs of crystalline cellulose can be prepared: celluloses II, III_I, III_{II}, IV_I and IV_{II} (Park et al., 2010).

1.1.2 Chitin

Chitin is a structural polysaccharide that is found in several kingdoms of life. It is a major component of the exoskeletons of arthropods, as a structural component in some molluscs, as well as constituents of the cell wall of fungi, algae and yeast. Chitin is a homopolymer of *N*-acetyl-D-glucosamine (GlcNAc) units, and like cellulose, the monosaccharide units are linked through β1→4 glycosidic bonds with each unit rotated 180° with respect to its neighbor (Figure 1.3). Two main crystalline allomorphs of chitin are found in nature, α- and β-chitin; a third less abundant variant, γ-chitin, is a combination of the two. The most abundant allomorph, α-chitin, is usually found in the exoskeletons of crustaceans (e.g. shrimps and crabs). β-chitin can be isolated from squid pens, whereas γ-chitin can be found in insects, fungi and the stomach of the *Loligo* squid (Alvarez, 2014; Kumirska et al., 2010). As with cellulose, chitin is insoluble in water due to strong intermolecular hydrogen bonding (Khoushab & Yamabhai, 2010); however, the bonding patterns of α- and β-chitin are different, resulting in different solubility in water. β-chitin, composed of intrasheet hydrogen-bonding by parallel chains, has higher solubility and swelling properties than α-chitin which has an antiparallel chain structure with stronger intermolecular bonding (Kumirska et al., 2010; Kurita et al., 2005). Deacetylation of chitin, by chemical or enzymatic treatment, can lead to conversion of chitin into a random copolymer of β-(1,4)-D-glucosamine (deacetylated) units and β-(1,4)-*N*-acetyl-D-glucosamine (acetylated) units. If the degree of acetylation is lower than 50 %, the polysaccharide is soluble in water and is referred to as chitosan (Younes & Rinaudo, 2015).

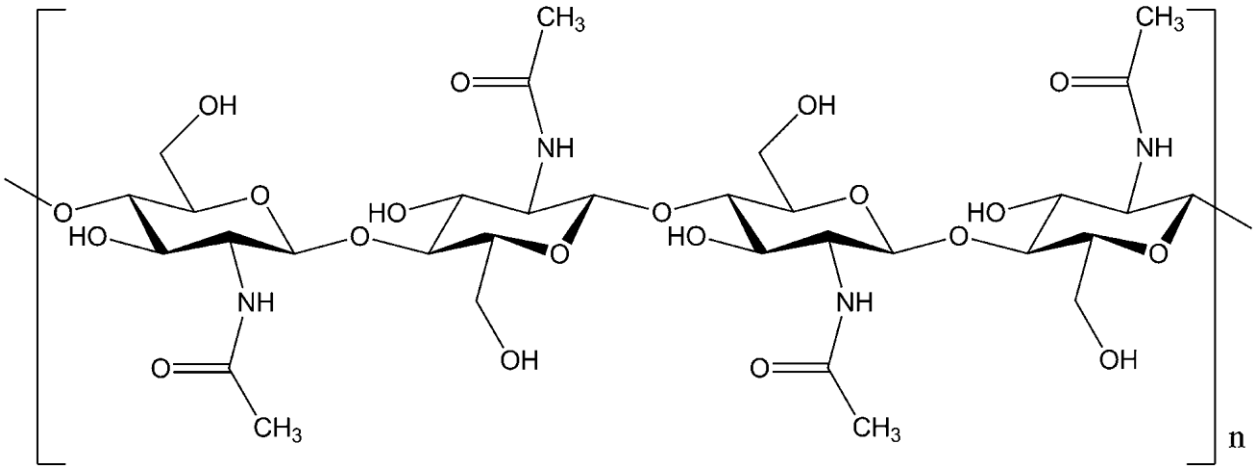


Figure 1.3. Structure of chitin. *Chitin is composed of several β1→4 linked N-acetyl-D-glucosamine (GlcNAc) monosaccharides. Because every other monosaccharide is rotated 180° relative to each other, the repeating structural unit of chitin is the disaccharide N,N'-diacetylchitobiose ($(GlcNAc)_2$). Figure source: Chen et al. (2014)*

1.2 Microbial degradation of structural carbohydrates

The function of structural polysaccharides (i.e. cellulose, hemicellulose and chitin) is mainly to protect the organisms by providing mechanical strength and chemical resistance. The strong intermolecular hydrogen bonding of these polysaccharides helps to form crystalline and insoluble structures that are highly recalcitrant and therefore resistant to degradation. However, several microorganisms have developed efficient enzymatic machineries that allow a relatively efficient conversion of these polysaccharides into soluble sugars, i.e. nutrients for growth. Bacteria, fungi and protozoa degrade polysaccharides using enzyme systems that include several types of carbohydrate-active enzymes (CAZymes). Aerobic cellulolytic and chitinolytic microorganisms secrete a vast array of free enzymes that act synergistically to degrade the biomass (Resch et al., 2013). In contrast, anaerobic bacteria utilize large macromolecular multienzyme complexes called cellulosomes that often are anchored to the cell wall of the bacterium. The main focus of the present thesis is on enzymes from aerobic systems, thus the following sections will primarily describe the secreted, free enzyme systems involved in cellulose and chitin degradation.

1.3 Enzymatic degradation of structural carbohydrates

1.3.1 Carbohydrate-active enzymes

Enzymes that degrade, modify or synthesize glycosidic bonds of carbohydrates are referred to as Carbohydrate-Active Enzymes (CAZymes). The Carbohydrate-Active enZYmes Database (CAZy; www.cazy.org) describes the families of both catalytic and carbohydrate-binding modules, and the classification system is based on amino acid sequence and structure similarities, rather than enzyme specificities (Cantarel et al., 2009). As of date (May, 2016), five different enzyme classes and one class of associated modules are described in the CAZy database. The largest class, the glycoside hydrolases (GHs), are subdivided into 135 families that are responsible for breakage of glycosidic bonds through hydrolysis. GHs covers most of the enzymes involved in degradation of cellulose and chitin as described in the next sections. The other CAZy classes are glycosyltransferases (synthesis of glycosidic bonds), polysaccharide lysases (non-hydrolytic cleavage of glycosidic bonds), carbohydrate esterases (hydrolysis of carbohydrate esters) and auxiliary activities (redox enzymes that act in connection with CAZymes). The latter class is the latest addition to the CAZy database (Levasseur et al., 2013), and it includes enzymes involved in lignin degradation and the families of the newly discovered lytic polysaccharide monooxygenases [LPMOs; discussed in detail in later sections; (Gustav Vaaje-Kolstad et al., 2010)] among others. The last CAZy class covered by the CAZy database is carbohydrate-binding modules (CBMs). These are non-catalytic binding modules that usually are appended to a catalytic domain of a carbohydrate-active enzyme. CBMs bind to carbohydrate substrates and thus bring the catalytic domains close to the substrate and thereby promotes catalytic activity (Boraston et al., 2004).

In addition to the CAZy database, an alternative classification system is provided by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology in which enzymes are given an Enzyme Commission number (EC number) based on the reactions they catalyze. The numerical classification divides enzymes into six groups based on the type of reaction that is catalyzed. The class of enzymes that is most relevant to carbohydrate degradation is glycoside hydrolases, that are given the enzyme commission number EC 3.2.1.x, where x represents the respective substrate.

1.3.2 Enzymatic degradation of cellulose

Due to the recalcitrant nature of cellulose, cellulolytic microorganisms utilize enzyme systems consisting of several different types of hydrolytic and oxidative enzymes that act in synergy in the enzymatic degradation of lignocellulosic substrates (Figure 1.4). The traditional view of enzymatic degradation of cellulose involved three different types of enzymes: endo-β-1,4-glucanases, exo-β-1,4-glucanases and β-glucosidases (Horn et al., 2012). Endo-β-1,4-glucanases (EC 3.2.1.4), also known as endoglucanases (EGs), are cellulases that can be found in 15 of the glycoside hydrolase (GH) families in the CAZy database. Endoglucanases randomly hydrolyze glycosidic bonds in the cellulose chain, and may act in a processive (bound to a CBM module) or a non-processive manner. Various endoglucanases may have different preferences for varying types of cellulose microfibrils; the CBM module of endoglucanases are thought to have high affinity for crystalline regions of cellulose (Orłowski et al., 2015), whereas the catalytic domain has a high affinity for amorphous cellulose regions.

Hydrolysis of internal glycosidic bonds in the cellulose chain results in generation of new reducing and non-reducing chain ends, which can be accessed by a group of exo-acting cellobiohydrolases. Exo-β-1,4-glucanases attacks the reducing or the non-reducing ends of the cellulose chains and releases mainly the disaccharide cellobiose. Exo-β-1,4-glucanases acting on the reducing ends of cellulose polymers (EC 3.2.1.176) can be found in GH families 7, 9 and 48, whereas those acting on the non-reducing ends (EC 3.2.1.91) are described in GH families 5,6 and 9. As with endoglucanases, many exo-β-1,4-glucanases are processive (Horn et al., 2012), meaning they can catalyze consecutive hydrolytic reactions without releasing from the cellulose chains.

Exo-β-1,4-glucanases have been described to be involved in two different types of synergism: synergy between endo-β-1,4-glucanases and exo-β-1,4-glucanases, referred to as endo-exo synergism (Jalak et al., 2012); and synergy between two exo-β-1,4-glucanases acting on the reducing end and the non-reducing end, referred to as exo-exo synergism (Igarashi et al., 2011).

Hydrolysis of the glycosidic bond of cellobiose and longer solubilized cello-oligomers is performed by β-glucosidases (EC 3.2.1.21). β-glucosidases, found in GH families 1, 3, 5, 9, 30 and 116, converts solubilized products released of exo-β-1,4-glucanases into single glucose units.

All cellulose active enzymes described so far cleave glycosidic bonds through hydrolysis; however, in recent years it has been shown that a group of enzymes, known as lytic polysaccharide monooxygenases (LPMOs), catalyze oxidative cleavage of insoluble polysaccharides (i.e. cellulose and chitin) (Forsberg et al., 2011; Gustav Vaaje-Kolstad et al., 2010). Cleavage of cellulose chains by LPMOs results in two products, in which one contains an oxidized chain end. It is thought that LPMO action results in an increase in available chain ends in regions that are otherwise inaccessible to cellulases (Eibinger et al., 2014). The synergistic action of LPMOs and cellulases improves saccharification yields which is of great importance for the present and future bio-economy (Horn et al., 2012).

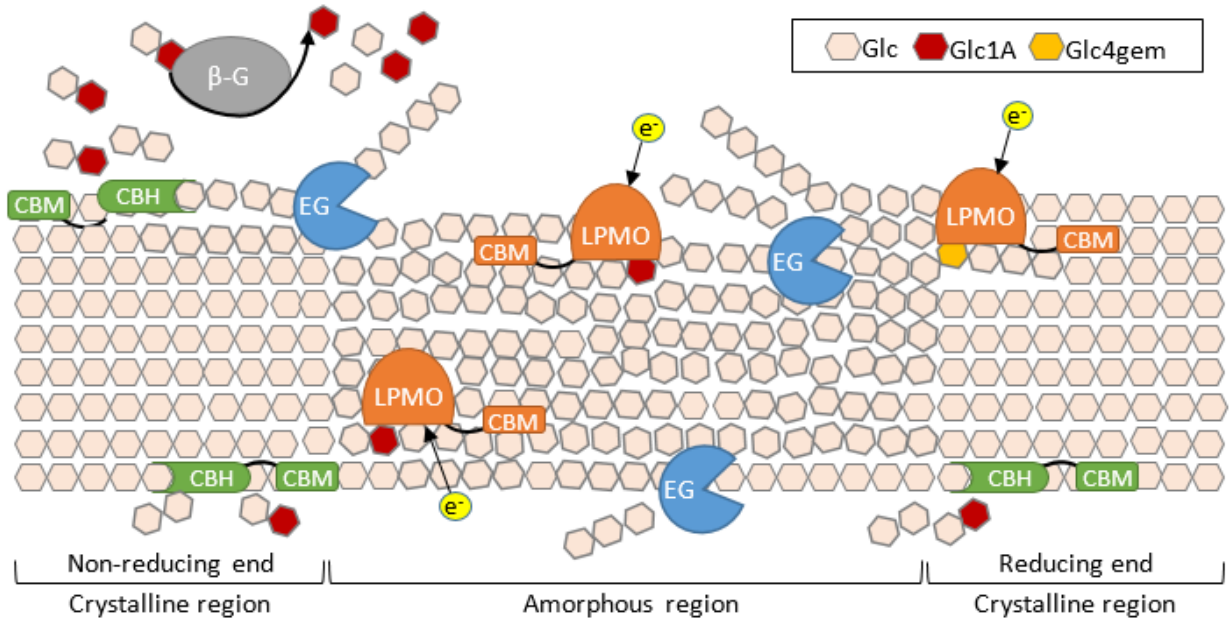


Figure 1.4. Schematic illustration of the synergistic action of hydrolytic and oxidative enzymes in cellulose degradation*. Abbreviations: LPMO, lytic polysaccharide monooxygenase; CBM, carbohydrate-binding module; EG, endoglucanase; CHB, cellobiohydrolase; β-G, β-glucosidase. LPMOs (orange) cleave the cellulose chains at either C1 and/or C4 position which leads to products containing a gluconic acid (Glc1A) or a geminal diol (Glc4gem) chain end. The activity of LPMOs are dependent on electrons (yellow circles). Products released by LPMOs are hydrolyzed by endoglucanases (blue), and the resulting products containing either a reducing or a non-reducing end serves as substrates for cellobiohydrolases (green). Cellobiohydrolases are exo enzymes that cuts of dimers in a processive manner. The dimers (cellobiose) are further hydrolyzed into single glucose units by the β-glucosidases (grey).*

1.3.3 Enzymatic degradation of chitin

Microbial degradation of chitin can occur via two different pathways. A chitinolytic process involves hydrolysis of β1→4 glycosidic bonds in the chitin chains by chitinases, followed by further degradation by a chitobiase to monomeric *N*-acetylglucosamine (Beier & Bertilsson, 2013). Alternatively, deacetylation of chitin to chitosan and subsequent hydrolysis by chitosanases and glucosaminidases results in monomeric glucosamine (Yong et al., 2010).

In the chitinolytic pathway, various chitinases (EC 3.2.1.14) act synergistically to hydrolyze the insoluble chitin chains (Suzuki et al., 2002; G. Vaaje-Kolstad et al., 2013). Chitinases belong to GH families 18, 19, 23 and 48, and degrade chitin by randomly attacking a point along the chains (endochitinase) or by attacking either the reducing or the non-reducing end of the chitin chain (exochitinase) (Svein J. Horn et al., 2006). Hydrolysis of chitin chains by chitinases generates soluble chitin oligomers (mainly *N*,*N*'-diacetylchitobiose ($[GlcNAc]_2$)) that are subsequently cleaved into *N*-acetylglucosamine (GlcNAc) by β-N-acetylhexosaminidases (EC 3.2.1.52). β-*N*-acetylhexosaminidases are found in GH families 3, 20, 84 and 116, and they cleave off GlcNAc units from the terminal non-reducing end of the chitin oligomer through hydrolysis (Beier & Bertilsson, 2013).

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

1.4 Lytic polysaccharide monooxygenases (LPMOs)

1.4.1 The discovery of LPMOs

As mentioned in earlier sections, the traditional view of enzymatic degradation of cellulose and chitin involved the action of mainly hydrolytic enzymes (e.g. cellulases and chitinases), a view that persisted up to the early years of the 21st century. However, as early as 1950, Reese *et al*. proposed that biological solubilization (i.e. degradation) of cellulose would require at least two steps, in which the first step would be performed by an activity that disrupted the cellulose structure (today identified as the LPMOs), whereas the second step involves hydrolysis of the β-(1,4)-glucan chain (Reese et al., 1950). In 1997, Watanabe *et al*. described a 21 kDa protein that

was secreted from *Serratia marcescens* when grown on chitin (Watanabe et al., 1997). The protein, at the time classified as a family 33 carbohydrate-binding module (CBM33), was shown to have chitin binding properties and was named chitin binding protein 21 (CBP21). In 2005, the structure of CBP21 was solved by Vaaje-Kolstad *et al*., and later that year it was shown that the presence of CBP21 boosted the hydrolysis of crystalline chitin by chitinases (Gustav Vaaje-Kolstad et al., 2005b; Gustav Vaaje-Kolstad et al., 2005a); however, the catalytic activity of CBP21 was still not uncovered. At about the same time, studies that involved a GH family 61 protein named Cel61A were published. The protein was shown to be a weak endoglucanase that could adsorb to cellulose substrates, and there were indications that Cel61A could have a function in hydrolysis of polysaccharides (Karlsson et al., 2001). Over the next couple of years, the proteins classified in the GH61 family where described as cellulase boosting enzymes that acts synergistically with cellulases in degradation of cellulose in the presence of divalent metal ions (Harris et al., 2010).

The major breakthroughs regarding catalytic activity of CBM33 and GH61 proteins occurred in 2010 and 2011. First, Vaaje-Kolstad *et al*. showed that CBP21 in fact is an oxidative enzyme that cleaves glycosidic bonds in crystalline chitin (Vaaje-Kolstad et al., 2010). The activity was shown to be dependent on the presence of divalent metal ions and that the presence of electron donors (e.g. ascorbic acid) greatly enhanced the activity. The next year, Quinlan *et al*. demonstrated that GH61 actually were copper-dependent oxidases, and that the cellulose degrading activity was boosted by the addition of gallic acid (as an electron donor) (Quinlan et al., 2011). Following the discovery of the catalytic activity of CBP21 on chitin, Forsberg *et al*. showed that CelS2, a CBM33 protein from *Streptomyces coelicolor*, is capable of oxidative cleavage of cellulose (Forsberg et al., 2011).

The term lytic polysaccharide monooxygenases (LPMOs) was first coined by Horn *et al*. in 2012, and at the time it covered both CBM33 and GH61 proteins (Horn et al., 2012). In 2013, LPMOs were reclassified in the CAZy database as auxiliary activity families 9 (formerly GH61) and 10 (formerly CBM33) (Levasseur et al., 2013). In recent years, the class of auxiliary activities have been expanded with the addition of two more families: AA11, comprising fungal chitin-active LPMOs (Hemsworth et al., 2014); and AA13 which includes fungal starch-active LPMOs (Leggio et al., 2015; Vu et al., 2014a).

1.4.2 Classification

LPMO encoding genes are widely found in the genomes of biomass-degrading microorganisms (Levasseur et al., 2013). In the CAZy database, LPMOs are classified as auxiliary activities 9, 10, 11 and 13. Families AA9, 11 and 13 consists of fungal proteins only, whereas family AA10 consists of proteins from several domains of life, although mainly bacterial AA10 proteins have been characterized so far. There are great variations in the number of LPMOs that are encoded in the genomes of biomass-degrading organisms. Fungal genomes usually encode several LPMOs, with as many as 40 AA9 genes which has been reported in the saprophytic fungus *Cheatomium globosum* (Busk & Lange, 2015). In contrast, bacterial genomes usually contain one or two AA10 genes, although a few bacterial genomes from the phylum of Actinobacteria have been shown to harbor up to seven LPMO genes.

LPMOs target a variety of substrates, and some are able to bind to and degrade more than one substrate. Family AA10 proteins, which are the focus of this study, are separated phylogenetically (Figure 1.5) into distinct groups that target either chitin (Nakagawa et al., 2015; Gustav Vaaje-Kolstad et al., 2012; Gustav Vaaje-Kolstad et al., 2010) or cellulose (Forsberg et al., 2011), and some have been described to target both substrates (Forsberg et al., 2014b). Family AA9 are exclusively cellulose-active, although recent studies have showed that an LPMO from *Neurospora crassa* (*Nc*LPMO10C) is active on soluble cello-oligosaccharides as well as hemicelluloses such as xyloglucan and glucomannan. (Agger et al., 2014; Isaksen et al., 2014). Members of the families AA11 and AA13 are active solely on chitin and starch, respectively (Hemsworth et al., 2014; Vu et al., 2014a).

The ability of LPMOs to bind efficiently to various substrates is often promoted by the presence of one or more carbohydrate-binding modules (CBMs). CBMs are non-catalytic modules that facilitate substrate binding by leading an enzymes catalytic domain into close proximity to the substrate surface. Enhanced and prolonged contact with the substrate leads to an increase of enzyme concentration on the substrate surface which results in more efficient degradation of the polysaccharide (Bolam et al., 1998). In LPMOs, one or more CBMs are attached to the catalytic domain via a flexible linker. The loss of a CBM module from two AA10 LPMOs have been shown to cause a reduction in enzyme activity towards Avicel and phosphoric acid swollen cellulose [PASC; (Arfi et al., 2014; Crouch et al., 2016; Forsberg et al., 2014b)]. CBMs are

currently classified into 74 families according to sequence similarities in the CAZy database. The majority of LPMOs appear as single domain proteins; however, a study by Book *et al.* showed that 31 % of AA9 proteins includes a cellulose-binding CBM1. It was also shown that 31 % of AA10 sequences included cellulose-binding domains CBM2 and CBM3 or chitin-binding domains CBM5 and CBM12 (Book et al., 2014).

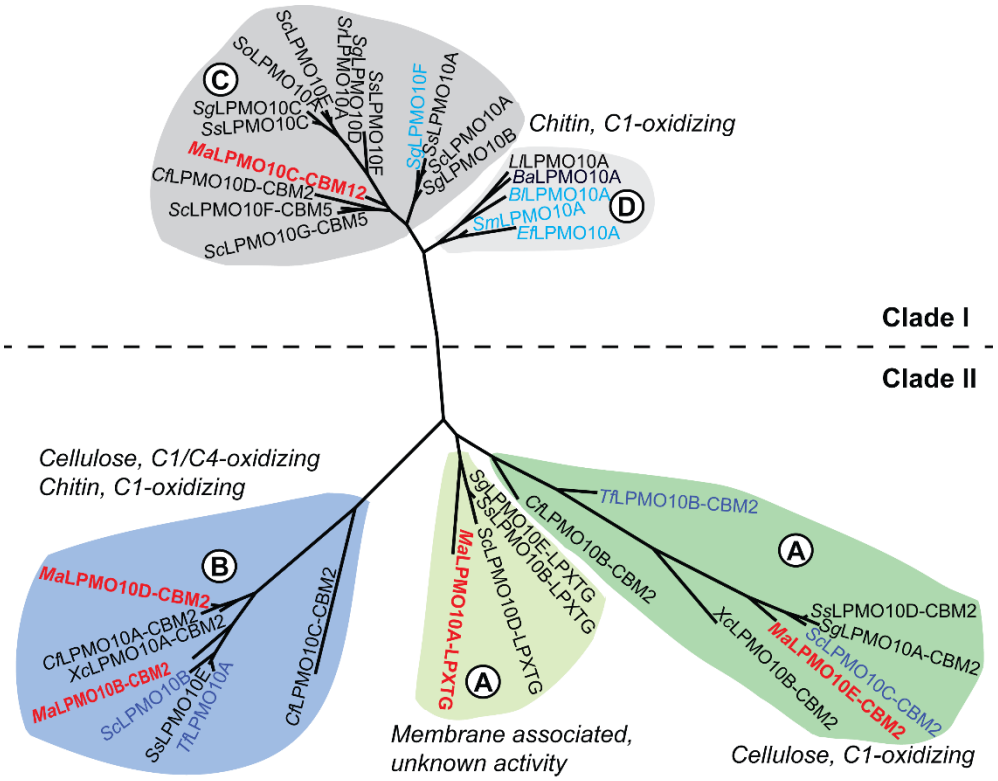


Figure 1.5. Phylogenetic three of selected family AA10 LPMOs. *Clade I is divided into subclade C and D and includes chitin-active LPMOs. Clade II contains LPMOs that are known to be cellulose active (subclade A) or both cellulose and chitin active (subclade B). Furthermore, subclade A contains C1-oxidizing LPMOs and membrane associated LPMOs (with an LPXTG motif) with unknown activity. Besides the dual substrate specificity in subclade B, those enzymes have been demonstrated to oxidize both the C1 and the C4 carbon on cellulosic substrates. LPMOs labeled red are encoded in the genome of the bacterium Micromonospora aurantiaca 27029, and the two MaLPMO10s in subclade B are the focus of this study. Figure source: Modified from Forsberg et al. (2016).*

1.4.3 Structure of LPMOs

A study comparing the sequences of AA9 and AA10 proteins in the CAZy database demonstrated that no sequences across these two families had significant homology [E-value $<1e^{-5}$ (Book et al., 2014)]. Similar studies have also demonstrated no sequence homology between these two families and the recently introduced AA11 and AA13 families (Hemsworth et al., 2014; Leggio

et al., 2015). Although sequence similarity is low, LPMOs across the various AA families have high structural similarity with several common structural features. The first crystal structure of an LPMO to be determined was that of the chitin-active CBP21 from *S. marcescens* by Vaaje-Kolstad *et al.* in 2005. The structure revealed a fibronectin type III (FnIII)-like domain consisting of a β-sandwich comprising a three-stranded and a four stranded β-sheet. Moreover, it was found that the putative substrate-binding surface of the protein was flat and that it contained several conserved hydrophilic residues (Gustav Vaaje-Kolstad et al., 2005a). These findings stood in contrast to the substrate-binding surface of other chitin active enzymes that commonly contains a substrate-binding surface with a groove, cleft or tunnel-like architecture lined with several aromatic residues. In 2010 it was shown that CBP21 actually was an enzyme, that due to its flat substrate-binding surface preferred binding to the flat highly ordered surface of crystalline chitin (Gustav Vaaje-Kolstad et al., 2010). Two years later, the structure of CBP21 was confirmed by NMR spectroscopy as a compact and rigid molecule (Aachmann et al., 2012). Since 2005, structures of 24 LPMOs have been solved, comprising all of the four AA families. In general, the overall fold of LPMOs are similar to that of CBP21 (Gustav Vaaje-Kolstad et al., 2005a), involving a β-sandwich with a varying number of β-strands (usually 8-10; Figure 1.6). The β-sandwich of cellulose-active LPMOs contain several conserved aromatic residues that have been suggested to play a role in electron transfer to the copper ion (Forsberg et al., 2014b). The β-strands are connected by a series of loops, in which longer loops may contain short helices (Beeson et al., 2015). While the β-sandwich is highly conserved, a loop referred to as loop L2 (termed motif 1 in AA10 proteins; colored magenta in figure 1.6) constitutes a highly diverse region of LPMOs (Hemsworth et al., 2015; X. Li et al., 2012). The L2 loop/motif 1 varies in length and conformation, and has been suggested as important for substrate binding and specificity (Book et al., 2014; Forsberg et al., 2014b; Forsberg et al., 2016).

As for CBP21, the substrate-binding surface of most LPMOs are flat [one exception exist; the starch-active *Ao*AA13 has a shallow groove along the surface (Leggio et al., 2015)], and contains the copper-active site, which will come in contact with the substrate upon binding. Due to the insolubility of crystalline polysaccharide substrates, structural studies of LPMOs bound to substrates have not been possible; however, it is presumed that binding is dependent on the presence of one or more aromatic or polar residues on the flat substrate-binding surface. Aromatic residues are assumed to play an important role in binding of CBMs to various

substrates (Boraston et al., 2004), and are also present (usually two or three) on the binding surface of most AA9 proteins (Wu et al., 2013). The binding surface of family AA10 proteins usually contain only one aromatic residue (Forsberg et al., 2014b), and are presumed to depend partly on the formation of hydrogen bonds from the polar surface residues to the polysaccharide surface. Neither AA11 nor AA13 proteins contain aromatic binding surface residues, but where AA11 depend on hydrogen bonding of polar residues, the AA13 proteins are presumed to be entirely dependent on a starch-binding CBM20 module for substrate binding (Hemsworth et al., 2014; Leggio et al., 2015).

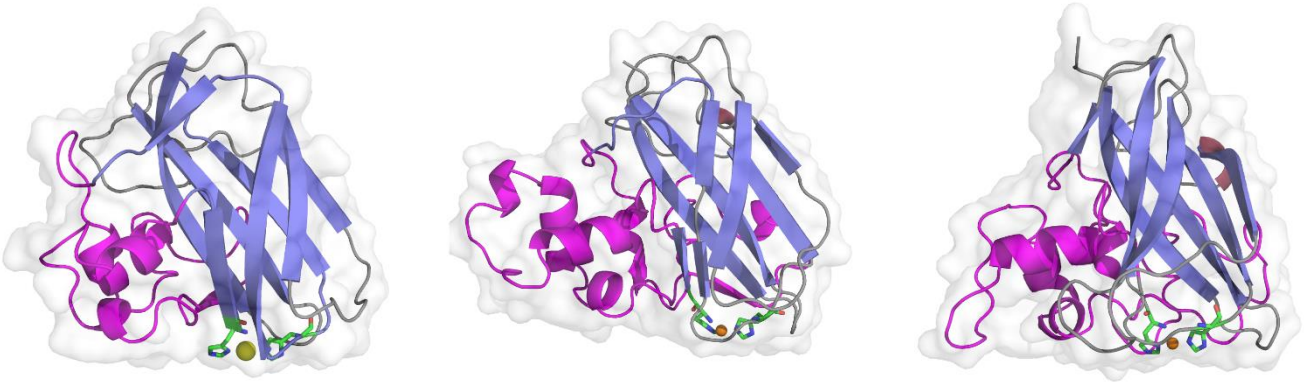


Figure 1.6. Cartoon representation of three AA10 proteins. *The structures of SmLPMO10A (Chitin-active C1 oxidizer), ScLPMO10B (Cellulose and-chitin active C1/C4 oxidizer) and ScLPMO10C (Cellulose-active C1 oxidizer) are shown as cartoon representations, with the histidine brace (discussed below) represented as sticks. The metal ions are shown as orange (copper) or yellow (sodium) spheres. The β-strands are colored blue, and the short loops between β-strands are colored grey. Motif 1 (L2 loop), including a varied number of helices, are colored in magenta. Notice the variable sizes of the motif, which in the chitin-active SmLPMO10A is smaller than in those of the cellulose-active ScLPMO10B and ScLPMO10C. Figures were made with PyMOL (Delano & Lam, 2005).*

LPMO activity is dependent on a copper ion bound to two fully conserved histidines in a T-shaped configuration referred to as the histidine brace (Figure 1.7). After some initial confusion regarding the metal-binding site (Harris et al., 2010; Gustav Vaaje-Kolstad et al., 2010), later studies demonstrated that the active site contained a monomeric type 2 copper ion (Phillips et al., 2011; Quinlan et al., 2011; Gustav Vaaje-Kolstad et al., 2012). The first determined protein structures contained various metal ions in the active site (e.g. sodium ion in *Sm*LPMO10A in figure 1.7), but since 2011, all determined structures contain a copper ion. Apart from the N-terminal histidine and a second fully conserved histidine that together forms the histidine brace, additional residues in the copper coordinating sphere constitutes of an axially positioned

hydrophobic residue (tyrosine or phenylalanine; figure 1.7 on the protein side and an alanine residue occupies the axial solvent side (Forsberg et al., 2014b; Hemsworth et al., 2013).

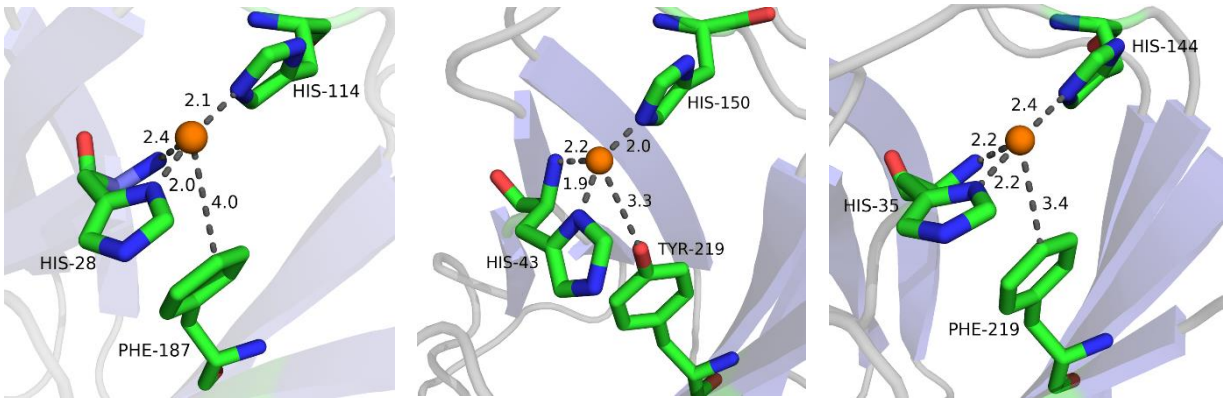


Figure 1.7. Active site of three AA10 type LPMOs. *The copper coordinating histidines and tyrosine/phenylalanine residues of SmLPMO10A (Chitin-active C1 oxidizer), ScLPMO10B (Cellulose and-chitin active C1/C4 oxidizer) and ScLPMO10C (Cellulose-active C1 oxidizer) are shown as stick representations, and the distances to the metal ion (orange sphere) are indicated by dashed lines. The active site of ScLPMO10B and ScLPMO10C contain a copper ion, whereas a sodium ion is present in the active site of SmLPMO10A. Figures were made with PyMOL (Delano & Lam, 2005).*

1.4.4 Activity of LPMOs

With the discovery of oxidative cleavage of chitin by CBP21 (*Sm*LPMO10A) in 2010, it was demonstrated by isotope-labeling that the two oxygen atoms introduced at the chain ends of oxidized products comes from water and molecular oxygen (O_2). It was also shown that the activity was dependent on the presence of a divalent metal ion at the active site, and that activity was greatly enhanced by an external electron donor [i.e. ascorbic acid (Gustav Vaaje-Kolstad et al., 2010)]. The exact mechanism for electron delivery to the active site is still unclear, and several electron-transfer routes have been suggested. Binding of LPMO to a substrate most likely makes the active site inaccessible for external electron donors, and thus it is plausible that that long-distance electron delivery takes place. One possible electron delivery route is across a series of conserved tryptophan residues located in the β-sandwich close to the active site (Hemsworth et al., 2013). Electrons to the active site may be provided by the enzyme cellobiose dehydrogenase [CDH; (Langston et al., 2011; Phillips et al., 2011)] or by small molecule electron donors (e.g. ascorbic acid, gallic acid and reduced glutathione). In a recent study by Westereng *et al*., it was

demonstrated that lignin may function as an electron reservoir and that electrons are provided to the enzyme by long-distance electron transfer (Westereng et al., 2015).

The catalytic mechanism of LPMOs is not yet fully understood. A study by Philips *et al.* suggested (Figure 1.8) that an electron donor (in this case the enzyme CDH) reduces the active site Cu(II) to a Cu(I) which then binds molecular oxygen. Internal electron-transfer results in formation of a copper superoxo intermediate which abstracts a hydrogen atom from either the C1 or C4 carbon on the carbohydrate. A second electron is added to the reaction from the electron donor which leads to cleavage of the Cu-bound hydroperoxide. The substrate radical couples with the copper oxo species, hydroxylates the substrate, which results in cleavage of the glycosidic bond and release of the adjacent glucan unit (Phillips et al., 2011).

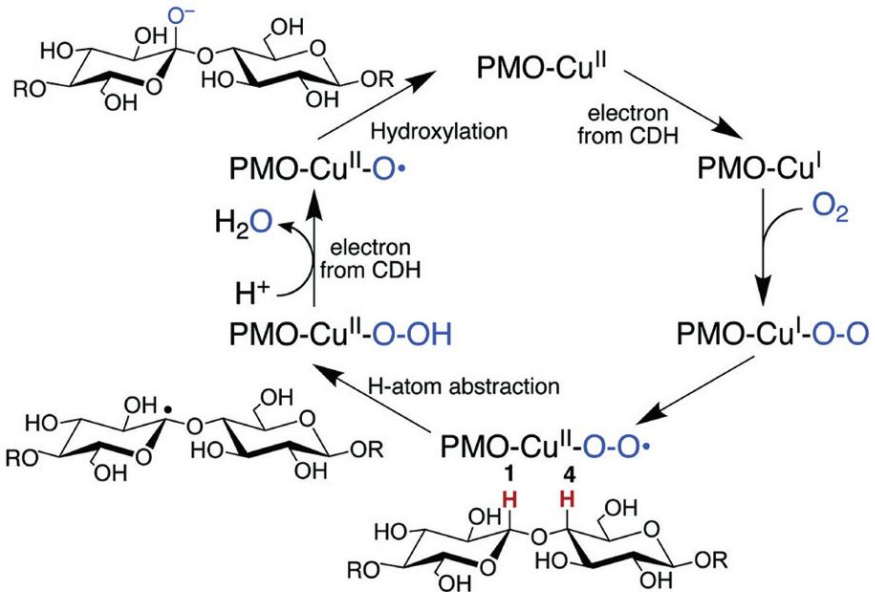


Figure 1.8. LPMO reaction mechanism as proposed by Phillips et al. *CDH (electron donor) reduces copper, which then activates molecular oxygen (O_2). One of the oxygen atoms abstracts a hydrogen atom from either the C1 or the C4 carbon of the substrate. CDH donates a second electron to the LPMO which leads to cleavage of the Cu-bound hydroperoxide. The remaining oxygen atom bound to the copper of the LPMO couples with the substrate radical which leads to hydroxylation of the substrate. An elimination reaction leads to cleavage of the glycosidic bond. The PMO in the figure represents an LPMO. Figure source: Modified from Phillips et al. (2011).*

Oxidation of cellulose or chitin substrates by LPMOs results in the release of two products in which one has an oxidized chain end. Oxidation of a carbohydrate by an LPMO can occur at two different positions in the monosaccharide unit; the C1 or the C4. The preference for cleavage of

specific sites in a molecule is referred to as regioselectivity. Oxidative regioselectivity in LPMOs was first described for AA9 proteins, which were early demonstrated to generate several different oxidized products (Phillips et al., 2011; Quinlan et al., 2011). Based on the regioselectivity of substrate oxidation, AA9 proteins can be divided into three groups; 1) oxidation of C1 only; 2) oxidation of C4 only; and 3) oxidation of both C1 and C4. Oxidative cleavage of cellulose by AA10 proteins was initially known to only occur at C1; however, a study by Forsberg *et al.* in 2014 demonstrated for the first time that two cellulose active AA10 proteins could oxidize both C1 and C4. The study also demonstrated a synergistic effect when a C1 oxidizer (*Sc*LPMO10C) was mixed with a C1/C4 oxidizer [(*Sc*LPMO10B (Forsberg et al., 2014b)]. Proteins of AA11 and AA13 are currently known to oxidize C1 only (Hemsworth et al., 2015). The products formed by oxidation of C1 or C4 is illustrated in figure 1.9. Oxidation at C1 results in formation of soluble oligosaccharides with a δ-1,5-lactone at the reducing end, which spontaneously becomes hydrolyzed into the more stable aldonic acid form. C4 oxidation (at the non-reducing end) generates a 4-ketoaldose which becomes hydrolyzed into its geminal diol form. Double-oxidized products can be formed when an oligosaccharide is cleaved at both ends by a C1/C4 oxidizing LPMO (not included in figure 1.9).

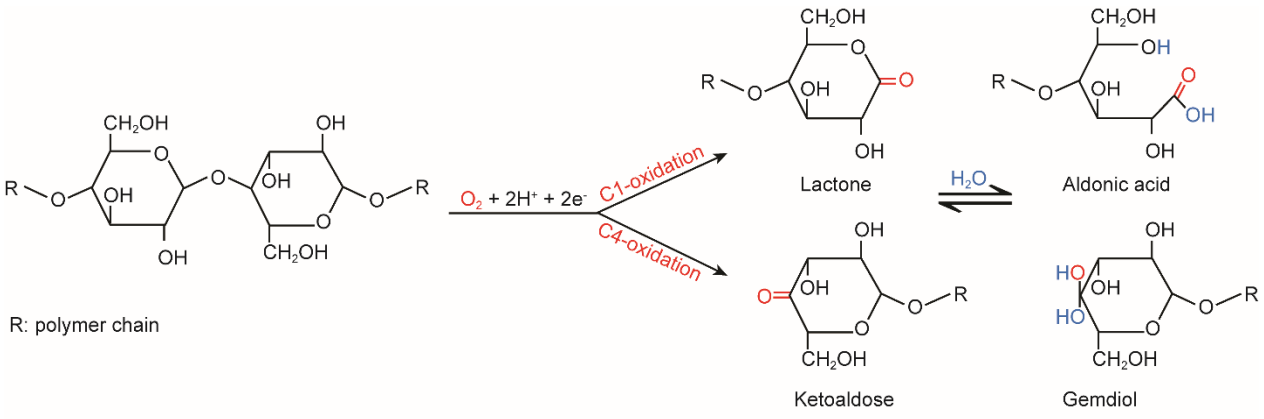


Figure 1.9. Oxidized products formed from oxidation of the C1 (top) or C4 (bottom) carbon of a cellulose chain. *Oxidation at C1 results in formation of a lactone (δ-1,5-lactone of Glc_n) that spontaneously hydrolyses to an aldonic acid (Glc_nGlc1A). Oxidation at C4 leads to formation of a 4-ketoaldose ($Glc4KGlc_n$) that also becomes hydrolyzed into its geminal diol form. Figure source: Modified from Loose et al. (2014).*

Since the discovery of LPMOs and its potential major role in the enzymatic conversion of biomass, extensive research has been conducted to uncover the molecular mechanisms of LPMO

activity. The mechanisms and structural properties behind substrate specificity and oxidative regioselectivity is still not fully understood. Several proposals have been made regarding the structural architecture behind oxidative regioselectivity, and studies have been conducted trying to connect regioselectivity and phylogeny (X. Li et al., 2012; Vu et al., 2014b). Oxidative regioselectivity is currently assumed to be due to the positioning of the LPMO on the polysaccharide substrate (Beeson et al., 2015). With the majority of the catalytic domain (i.e. β-sandwich) being highly conserved and rigid, the loop L2/motif 1 region (colored magenta in figure 1.6) is less conserved and more structurally diverse. This region, in addition to residues proximal to the active site, have been suggested as important for oxidative regioselectivity in AA10 proteins (Forsberg et al., 2014b). Figure 10 shows a comparison of the active site of a C1 oxidizer (*Sc*LPMO10C) and a C1/C4 oxidizer (*Sc*LPMO10B), and illustrates the differences in active site positioned residues and the distances between them. The active-site structure of *Sc*LPMO10C resembles that of chitin-active bacterial LPMOs, whereas the active-site of *Sc*LPMO10B is similar to cellulose-active AA9 LPMOs [e.g. *Ta*LPMO9A (Forsberg et al., 2014b)].

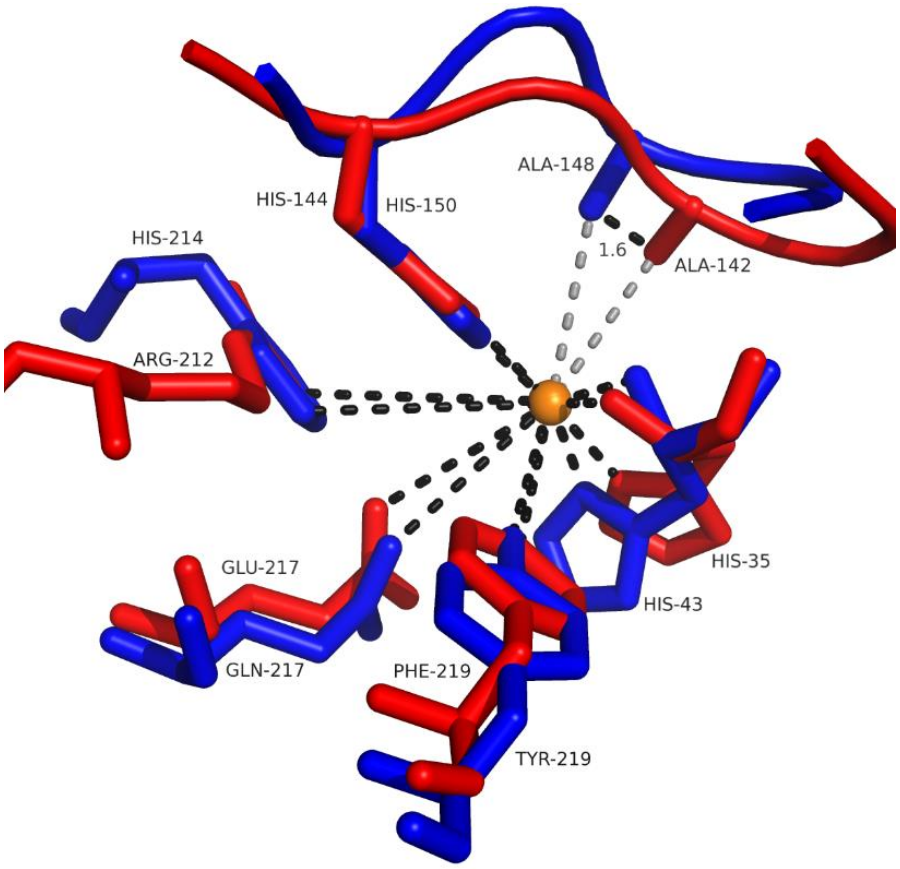


Figure 1.10. Structural comparison of the active site of ScLPMO10C (C1 oxidizer, red color) and ScLPMO10B (C1/C4 oxidizer, blue color). Distances to the copper ion (orange sphere) are indicated by dashed lines. Notice the difference in the position of the axial alanine. Excluding the conserved histidine restudies, the rest of the residues varies between the two LPMOs. Figure made with PyMOL (Delano & Lam, 2005)

1.6 Aim of the study

The prelude to this thesis was the discovery of C4 oxidation by cellulose-active AA10 proteins by Forsberg *et al.* in 2014, in which two LPMOs, both lacking a CBM, were characterized as C1/C4 oxidizing AA10 proteins. At the time it was suggested that determination of oxidative regioselectivity in cellulose-active AA10 proteins might be caused by different positioning of the active site alanine. In this study, two LPMOs from the bacterium *Micromonospora aurantiaca* were characterized. Both LPMOs contained a CBM, which led to easier and more accurate detection of activity levels, and they are both C1/C4 cellulose-active.

The main objective of this study was to gain more insight into the function and mechanism of cellulose-active AA10 proteins. The strategy was to target highly conserved residues in the sequences of strict C1 oxidizing cellulose-active AA10 proteins for site-directed mutagenesis. The targeted residues were mostly located near the copper active site or the substrate-binding surface of the LPMO.

The first part of this study focused on cloning, expression and purification of the two wild type LPMOs. Furthermore, the structure of *Ma*LPMO10B was determined by X-ray crystallography.

The second part of this study focused on analysis of *Ma*LPMO10B mutants. A total of 9 full-length *Ma*LPMO10B mutants were planned, in addition to one mutant of the truncated protein. The activity of the mutants was analyzed by MALDI-TOF MS and HPLC in order to identify residues that might be important for determination of C4 oxidation in the proteins.

2 Materials

2.1 Chemicals

Table 2.1. Chemicals

Chemical	**Supplier**
2,5-Dihydroxybenzoic acid (DHB)	Bruker
Acetonitrile (CH_3CN)	VWR
Agar-agar	Merck
Agarose, SeaKem®	Lonza
Ampicillin sodium salt	Sigma-Aldrich
Bacto™ Tryptone	Becton, Dickinson and Company
Bacto™ Yeast Extract	Becton, Dickinson and Company
Bis-Tris	Sigma-Aldrich
Bis-Tris propane	Sigma-Aldrich
CAPS	Sigma-Aldrich
Copper (II) sulfate ($CuSO_4$)	VWR
D(+)-Cellobiose, ≥ 98 %	Sigma-Aldrich
D(+)-Glucose monohydrate	VWR
Ethanol 96 %	Arcus
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Glycerol 85 %	Merck
Hydrochloric acid (HCl)	Merck
L-ascorbic acid	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich
Magnesium sulfate heptahydrate ($MgSO_4$) x 7H2O	Sigma-Aldrich

Poly(ethylene glycol) (PEG) 8000	Sigma-Aldrich
Potassium chloride (KCl)	Merck
Potassium phosphate dibasic (K_2HPO_4)	Sigma-Aldrich
Potassium phosphate monobasic (KH_2PO_4)	Sigma-Aldrich
Sodium acetate anhydrous (CH_3COONa)	Sigma-Aldrich
Sodium chloride (NaCl)	VWR
Sodium hydroxide (NaOH) 50 %	Sigma-Aldrich
Tris/Glycine/SDS Buffer 10x	Bio-Rad
Trizma® base	Sigma-Aldrich

2.2 Proteins and enzymes

Table 2.2. Proteins and enzymes

Protein/enzyme	**Supplier**
Cel5A from *Thermobifida fusca*	Made in-house
Cel7A from *Trichoderma reesei*	Made in-house
CDH from *Myriococcum thermophilum*	Made in-house
CHB from *Serratia marcescens*	Made in-house
NEBuffer 2 (10x)	NEB
Phusion HF DNA Polymerase	NEB
Red Taq 2x DNA Polymerase Master Mix	VWR
PfuUltra HF DNA Polymerase	Agilent Technologies
BenchMark™ Protein Ladder	Thermo Fisher

2.3 DNA

Table 2.3. DNA ladders and dNTP mixtures

DNA	Supplier
dNTP mix (QuikChange II XL Site-Directed Mutagenesis Kit)	Agilent Technologies
dNTP mix (Phusion® High-Fidelity PCR Kit)	NEB
GeneRuler 1 kb DNA Ladder	Thermo Fisher
1 kb DNA Ladder	NEB

2.4 Carbohydrate substrates

Table 2.4. 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

2.5 Primers

Table 2.5. Primers by name and sequence

Primer name	Primer sequence (5' → 3')
MaB1_pR_IF	GAAGGAGATATACATATGTCAACGCCGTATCGT
MaB1-2_pR_IR	CAGCCGGATCAAGCTTTTAACTGGTCGTGGT

MaB1N-2N_pR_IR	CAGCCGGATCAAGCTTTTAGCCAAAGTCAACATCG
MaB2_pR_IF	CGCAACAGGCGAATGCCCATGGCAGCGTGGTT
MaD_InF_F	CGCAACAGGCGAATGCCCATGGCAGTGTTACGAAT
MaD_InF_R	CAGCCGGATCAAGCTTTTAGCGGGCGGTGCAGGTC
MaD_InF_2R	CAGCCGGATCAAGCTTTTAGCCAAAGATAACGTCAGA
MaB_W82Y_N85F_S	ACGAAACAGGCCAAACCAGTTATACATTGCGTTCGGATCGGCTTGCCA
MaB_W82Y_N85F_AS	TGGCAAGCCGATCCGAACGCAATGTATAACTGGTTTGGCCTGTTTCGT
MaB_Q141W_S	GGCGCCGTGGCTCGCCCAATCGAAAAATTTAACACGGAAGTTGTTGG
MaB_Q141W_AS	CCAACAACTTCCGTGTTAAATTTTTCGATTGGGCGAGCCACGGCGCC
MaB_Q141G_S	GCCGTGGCTCGCCCAATCGAAAAATTTAACACGGAAGTTG
MaB_Q141G_AS	CAACTTCCGTGTTAAATTTTTCGATTGGGCGAGCCACGGC
MaB_F88Y_S	TGTGGAACTGGAATGGCCTGTATCGTGAAGGTG
MaB_F88Y_AS	CACCTTCACGATACAGGCCATTCCAGTTCCACA
MaB_F88L_S	GCCACACCTTCACGTAACAGGCCATTCCAGT
MaB_F88L_AS	ACTGGAATGGCCTGTTACGTGAAGGTGTGGC
MaB_W82Y_S	GAAACAGGCCATTCCAGTTATACATTGCGTTCGGATCGGC
MaB_W82Y_AS	GCCGATCCGAACGCAATGTATAACTGGAATGGCCTGTTTC
MaB_N85F_S	CCTTCACGAAACAGGCCAAACCAGTTCCACATTGCGTT
MaB_N85F_AS	AACGCAATGTGGAACTGGTTTGGCCTGTTTCGTGAAGG
MaB_loop_S	GTGCGCAGCGACTCTCAAGAAAACTTTTACCTGTGCAGCGATGTTGACTT
MaB_loop_AS	AAAGTTTTCTTGAGAGTCGCTGCGCACCCAGATGGTATACACGACATGGC

Table 2.6. Primers by name and description

Primer name	**Primer sequence (5' → 3')**
MaB1_pR_IF	*Malpmo*10B, with native signal sequence in pRSET-B, forward cloning primer
MaB1-2_pR_IR	*Malpmo*10B, reverse cloning primer, full-length protein
MaB1N-2N_pR_IR	*Malpmo*10B, reverse cloning primer, N-terminal domain
MaB2_pR_IF	*Malpmo*10B, with cbp21 signal sequence (not codon optimized) in pRSET-B, forward cloning primer
MaD_InF_F	*Malpmo*10D, with cbp21 signal sequence in pRSET-B, forward cloning primer
MaD_InF_R	*Malpmo*10D, reverse cloning primer, full-length protein
MaD_InF_2R	*Malpmo*10D, reverse cloning primer, N-terminal domain
MaB_W82Y_N85F_S	*Malpmo*10B W82Y/N85F, forward mutational primer
MaB_W82Y_N85F_AS	*Malpmo*10B W82Y/N85F, reverse mutational primer
MaB_Q141W_S	*Malpmo*10B Q141W, forward mutational primer
MaB_Q141W_AS	*Malpmo*10B W141W, reverse mutational primer
MaB_Q141G_S	*Malpmo*10B Q141G, forward mutational primer
MaB_Q141G_AS	*Malpmo*10B Q141G, reverse mutational primer
MaB_F88Y_S	*Malpmo*10B, F88Y, forward mutational
MaB_F88Y_AS	*Malpmo*10B, F88Y, reverse mutational

MaB_F88L_S	*Malpmo*10B, F88L, forward mutational
MaB_F88L_AS	*Malpmo*10B, F88L, reverse mutational
MaB_W82Y_S	*Malpmo*10B, W82Y, forward mutational
MaB_W82Y_AS	*Malpmo*10B, W82Y, reverse mutational
MaB_N85F_S	*Malpmo*10B, N85F, forward mutational
MaB_N85F_AS	*Malpmo*10B, N85F, reverse mutational
MaB_loop_S	*Malpmo*10B, 213QASHILDQSY → 213VRSDSQENF ("Loop"), forward mutational
MaB_loop_AS	*Malpmo*10B, 213QASHILDQSY → 213VRSDSQENF ("Loop"), reverse mutational

2.6 Bacterial strains

Table 2.7. Bacterial strains

Strain	**Source**
Escherichia coli OneShot® BL21 Star™ (DE3)	Thermo Fisher
Escherichia coli OneShot® TOP10	Thermo Fisher
Escherichia coli XL1 Blue Supercompetent	Agilent Technologies

2.7 Plasmids

Table 2.8. Plasmids

Plasmid	**Source and reference**
pRSET-B_*cels2-n2*	pRSET-B from Thermo Fisher, with ligated *cels2_n2* gene (Forsberg et al., 2014) (Vector map is shown in Appendix A)
pUC57	GenScript®

3 Methods

3.1 Cultivation of bacteria

3.1.1 Agar and cultivation media

Lysogenic broth (LB)

Liquid medium:

- 10 g Bacto™ Tryptone (Becton, Dickinson and Company)
- 5 g Bacto™ Yeast Extract (Becton, Dickinson and Company)
- 10 g NaCl

All solid ingredients were dissolved in 800 ml dH_2O. The volume was adjusted to 1 liter with dH_2O and the medium was autoclaved. Media used for growing transformants were added ampicillin to a final concentration of 50 µg/ml prior to cultivation.

Agar plates:

- 10 g Bacto™ Tryptone (Becton, Dickinson and Company)
- 5 g Bacto™ Yeast Extract (Becton, Dickinson and Company)
- 10 g NaCl
- 15 g Agar (Merck)

All solid ingredients were dissolved in 800 ml dH_2O. The volume was adjusted to 1 liter with dH_2O and the medium was autoclaved. The solution was then cooled down to approximately 55 °C before 100 µg/ml ampicillin was added while working in a laminar flow cabinet. The medium was subsequently poured into petri dishes and after further cooling for 20 minutes the agar plates were stored at 4 °C.

Terrific broth (TB)

10x TB salts:

- 23,13 g KH_2PO_4

- 164.32 g K_2HPO_4

Prior to autoclavation, the chemicals were mixed and dissolved in dH_2O to a final volume of 1 liter.

Liquid medium:

- 12 g Bacto™ Tryptone
- 24 g Bacto™ Yeast Extract
- 4 ml 85 % glycerol

The ingredients were dissolved in dH_2O to a final volume of 900 ml and autoclaved. Before use 100 ml of the 10x TB salts solution was added.

Super Optimal Broth (S.O.C.) medium

- 10 g Bacto™ Tryptone
- 2.5 g Bacto™ Yeast Extract
- 0.292 g NaCl
- 0.093 g KCl

The ingredients were dissolved in dH_2O to a final volume of 400 mL. After autoclaving the following filter-sterilized (0.2 µm) solutions were added to the medium:

- 5 ml $MgCl_2$ (1 M)
- 10 ml $MgSO_4$ (0.5 M)
- 9 ml glucose (1.11 M)

Finally, the volume was adjusted to 500 ml with dH_2O.

3.1.2 Antibiotics

Ampicillin is a β-lactam antibiotic and part of the penicillin group of antibiotics. Ampicillin inhibits bacterial growth by interrupting bacterial cell wall synthesis and thus weakening the bacterial cell wall leading to cell lysis. The inner membrane of the cell wall of both Gram-

positive and Gram-negative bacteria contains penicillin-binding proteins. By binding to and inactivating these proteins, ampicillin alters the cross-linkage of peptidoglycan chains responsible for maintaining the strength and rigidity of the bacterial cell wall.

Resistance to ampicillin requires the bacteria to produce the β-lactamase protein. Through hydrolysis, the enzyme cleaves the β-lactam ring of ampicillin, causing inactivation of the antibiotic. The use of plasmids harboring the β-lactamase (*bla*) gene allows the utilization of ampicillin as a selectable marker.

In this study, the expression plasmid pRSET-B contains the *bla* gene, allowing the selection of the plasmid in *Escherichia coli* transformants. Ampicillin was used in both agar and liquid media, with end concentrations of 100 µg/ml and 50 µg/ml, respectively (unless otherwise stated).

3.1.3 Cultivation of bacterial strains

All reagents and media used for cultivation were sterilized either by autoclaving or sterile-filtration (0.45 µm). All culturing work was performed under sterile conditions. To start a new culture, a single colony from an agar plate or a piece of glycerol stock (see section 3.2) was inoculated in 5 ml medium of choice in sterile culture tubes and incubated overnight at 30 °C or 37 °C with shaking at a variety of speeds as described in detail below. For media recipes, see section 3.1.1.

E. coli is a Gram-negative, rod-shaped and facultatively anaerobic bacterium normally found in the lower intestine of warm-blooded organisms. The bacterium is the most common prokaryotic model organism, and plays an important role in the fields of biotechnology and microbiology. *E. coli* is often used for the production of heterologous proteins, where genes are introduced into the microbe using plasmids, and thus allowing the production of recombinant proteins. Several features of the bacterium make it preferable as a model organism in molecular cloning. Most strains have a generation time of only 20 minutes and it can grow both with and without oxygen. The optimal growth temperature of *E. coli* is 37 °C, but various strains can grow at a wide range of temperatures. The genome is sequenced and well understood, and competence can easily be artificially induced. Several *E. coli* strains were used in this study:

- *E. coli* One Shot® TOP10 cells were used for cloning purposes, with subsequent replication of high-copy number plasmids. Cells carrying plasmids encoding ampicillin resistance were cultivated in liquid LB-medium supplemented with 50 µg/ml ampicillin with shaking, or on LB-agar plates supplemented with 100 µg/ml ampicillin at 37 °C.
- *E. coli* BL21 Star™ (DE3) cells were used for expression of recombinant protein. Cells carrying plasmids encoding ampicillin resistance were cultivated in LB-medium supplemented with 50 µg/ml ampicillin with shaking, or on LB-agar plates supplemented with 100 µg/ml ampicillin. For optimal expression of *Ma*LPMO10B^{cd} cells were grown in TB-medium supplemented with 50 µg/ml ampicillin at 37 °C overnight.
- *E. coli* XL1-Blue supercompetent cells were used for replication of high-copy number mutated plasmids. Cells carrying plasmids encoding ampicillin resistance were cultivated on LB-agar plates supplemented with 100 µg/ml ampicillin at 37 °C overnight.

3.2 Long-term storage of bacteria

Long-term storage of bacterial cultures was accomplished by adding sterile glycerol to the cells. Glycerol serves as a cryoprotectant that is essential to reduce the cell damage caused by the low freezing temperatures. Bacterial culture and autoclaved glycerol were carefully mixed in a cryogenic tube as follows:

- 1 ml overnight culture of bacteria
- 300 µl glycerol (85 % (v/v), autoclaved)

After mixing of the ingredients, the glycerol stock was stored at -80 °C. Moreover, inoculation of bacterial cultures from the frozen glycerol stocks were made by scraping off some of the frozen culture with a sterile toothpick, and then transferring the toothpick into a tube containing LB-medium and ampicillin.

3.3 Restriction digestion

Preparation of the expression vector used for cloning of LPMO genes and subsequent transformation into chemically competent One Shot® TOP10 *E. coli* cells was performed using

various restriction endonucleases. The expression vector, pRSET-B_*cels2-n2*, was digested using two different sets of restriction enzymes (see below) to remove and subsequently replace the *cels2-n2* gene with or without the signal sequence of *cbp21* (a sequence encoding a N-terminal signal peptide that enables secretion of the recombinant protein to the periplasm). The restriction endonucleases recognize specific sequences and produce double-stranded cuts in the DNA. Cleavage at the same position in both strands, form DNA molecules with blunt ends, while cleavage at different positions form sticky ends. Different restriction enzymes require different reaction conditions (such as temperature and buffers) for optimal digestion.

3.3.1 Plasmid preparation of pRSET-B by *Hind*III and *Nde*I

Digestion of the pRSET-B_*cels2-n2* expression vector using *Hind*III and *Nde*I restriction endonucleases removes the *cels2-n2* gene and the *cbp21* signal sequence (see Appendix A for cleavage sites). The provider (NEB) recommends a reaction temperature of 37 °C for both enzymes, as well as NEBuffer 2 for optimal activity. Cleavage of dsDNA by each of the enzymes results in sticky ends.

Materials

- Plasmid: pRSET-B_*cels2-n2*
- Restriction enzymes
    - *Hind*III, 20 U/µl (NEB)
    - *Nde*I, 20 U/µl (NEB)
- 10x NEBuffer 2(NEB)
- Deionized H_2O

Procedure

The pRSET-B_*cels2-n2* plasmid extract was diluted to a concentration of 60 µg/ml using dH_2O, and 25 µl of the diluted sample was mixed with 5 µl 10x NEBuffer 2, and 4 µl of the two restriction enzyme. The digestion reaction mixture was incubated for two hours at 37 °C in a water bath. The digested and thereby linearized plasmid was visualized on a 1 % agarose gel as described in section 3.4.4 and was subsequently extracted from the gel as described in section 3.4.5.

3.3.2 Plasmid preparation of pRSET-B by *Hind*III and *Bsm*I

Restriction digestion of pRSET-B_*cels2-n2* by *Hind*III and *Bsm*I results in removal of the *cels2-n2* gene only, while keeping the signal sequence of *cbp21* (see Appendix A for cleavage sites). The provider (NEB) recommends a reaction temperature of 37 °C and 65 °C for *Hind*III and *Bsm*I, respectively, as well the addition of NEBuffer 2 for optimal activity. Cleavage of dsDNA by each of the enzymes results in sticky ends.

Materials

- Plasmid: pRSET-B_*cels2-n2*
- Restriction enzymes
    - *Hind*III, 20 U/µl (NEB)
    - *Bsm*I, 20 U/µl (NEB)
- 10x NEBuffer 2 (NEB)
- Deionized H_2O

Procedure

The pRSET-B_*cels2-n2* plasmid extract was diluted to a concentration of 60 µg/ml using dH_2O, and 25 µl of the diluted sample was mixed with 5 µl 10x NEBuffer 2 and 4 µl *Hind*III restriction enzyme. The mixture was incubated in a water bath at 37 °C for two hours. Next, 6 µl *Bsm*I was added to the reaction, followed by incubation of the reaction for another two hours in a water bath adjusted to 65 °C. The digested and thereby linearized plasmid was visualized on a 1 % agarose gel as described in section 3.4.4. and was subsequently extracted from the gel as described in section 3.4.5.

3.4 Polymerase chain reaction

Polymerase chain reaction (PCR) is an *in vitro* technique used for amplification of specific nucleotide sequences. The reactions require both sequence specific oligonucleotide primers complementary to the sequence of interest, dNTPs, and a thermo-stable DNA polymerase. In this study PCR was used for the following purposes:

- Gene amplification

- Verification of transformed products
- Site-directed mutagenesis

3.4.1 Gene amplification

PCR was used for amplification of codon optimized (see section 4.1.3 and Appendix B) *Malpmo10B* and *Malpmo10D* genes from the bacterium *Micromonospora aurantiaca* ATCC 27029. Transcription and translation of heterologous proteins require DNA sequences to be correct after amplification, and thus high-fidelity DNA polymerases must be utilized. Phusion® High-Fidelity DNA Polymerase (NEB) works fast and generates long DNA molecules with high fidelity. The polymerase extends primers in 5'-3' direction, and its 5'-3' exonuclease activity generates DNA molecules with blunt ends.

Materials

- Phusion® High-Fidelity DNA Polymerase kit (NEB)
- Template DNA
- Primers
- Nuclease-free dH_2O

Procedure

PCR reactions of 50 µl were set up on ice in 0.2 ml PCR tubes according to table 3.1. Reactions were carried out in a SimpliAmp™ Thermal Cycler (Thermo Fisher) using the settings shown in table 3.2.

Table 3.1. PCR reaction setup for Malpmo10B and Malpmo10D genes using Phusion® High-Fidelity DNA Polymerase.

Reagents	**Volume (Final concentration)**
Nuclease-free water	To 50 µl
5x Phusion HF / GC* buffer	10 µl (1x)
10 mM dNTPs	1 µl (200 µM)

10 pmol/µl Forward primer	2.5 µl (0.5 µM)
10 pmol/µl Reverse primer	2.5 µl (0.5 µM)
Template DNA (pUC57)	1 µl (0.08 µg/ml)
DMSO*	1.5 µl (3 %)
Phusion DNA polymerase	0.5 µl (1.0 units per 50 µl PCR)

** Reagent only used in attempts to clone Malpmo10D (see section 4.2.1)*

Table 3.2. Thermocycling conditions for PCR using Phusion® High-Fidelity DNA Polymerase. In attempts to clone Malpmo10D, an annealing temperature of 65 °C was used (see section 4.2.1).

Step	**Temperature**	**Time (mm:ss)**	**Number of cycles**
Initial denaturation	98 °C	00:30	1
Denaturation	98 °C	00:10	35
Annealing	55 °C (65 °C)	00:20	
Elongation	72 °C	00:25	
Final elongation	72 °C	10:00	1

After amplification, the PCR products were analyzed by agarose gel electrophoresis as described in section 3.4.4.

3.4.2 Verification of transformed products

Verification of transformed *E. coli* One Shot® TOP10 harboring the pRSET-B vector with the LPMO gene of interest was performed using Red Taq DNA Polymerase Master Mix (VWR). The mix is ready-to-use and contains bacterially derived Taq polymerase, dNTPs and $MgCl_2$. Since the mix contains a loading dye, the reaction mix only requires the addition of primers and a template, and can then be directly loaded onto a 1 % agarose gel for analysis.

Materials

- Red Taq DNA Polymerase Master Mix (VWR)
- Primers (see section 2.5)
- DNA template
- Deionized H_2O

Procedure

PCRs (20 µl) were set up on ice in 0.2 ml PCR tubes according to table 3.3. DNA template was prepared as follows: from a petri dish containing transformed cells, a single bacterial colony was picked using a sterile toothpick. Before inoculation in LB medium, the toothpick tip was brought in contact with the bottom of the 0.2 ml PCR tube, leaving a small amount of the bacterial colony inside the tube. Cells were lysed by heating the PCR tube in a microwave oven at maximum power for two minutes, releasing DNA. After cooling the tubes on ice, the rest of the reagents were added. The colony PCR reactions were carried out in a VWR Collection Thermal Cycler DOPPIO using the settings shown in table 3.4.

Table 3.3. Colony PCR reaction setup for MaLPMO10B and MaLPMO10D using Red Taq DNA Polymerase Master Mix

Reagent	**Volume (Final concentration)**
Taq 2x Master Mix	10 µl (1x)
Forward sequencing primer for pRSET-B	1 µl (0.5 µM)
Reverse sequencing primer for pRSET-B	1 µl (0.5 µM)
dH_2O	8 µl

Control reactions was performed to verify successful insertion of *Ma*LPMO genes into the pRSET-B plasmid. The reactions were set up as follows:

- Positive control reactions contained Taq 2x Master Mix and the pRSET-B_*cels2-n2* plasmid
- Negative control reactions contained only Taq 2x Master Mix

Table 3.4. Thermocycling conditions for colony PCR using Red Taq DNA Polymerase Master Mix

Step	**Temperature**	**Time (mm:ss)**	**Number of cycles**
Initial denaturation	95 °C	02:00	1
Denaturation	95 °C	00:30	25
Annealing	58 °C	00:30	
Elongation	72 °C	00:30	
Final elongation	72 °C	05:00	1

Following amplification, the PCR products were analyzed by agarose gel electrophoresis (see section 3.4.4).

3.4.3 Site-directed mutagenesis

*Ma*LPMO10B single and double point mutants were generated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). The kit was used to insert point mutations in selected codons or to replace the nucleotide sequence encoding a conserved loop region of 9 amino acids, in connection to the active site, with the nucleotide sequence of CelS2 (*Sc*LPMO10C). The method is based on the use of *PfuUltra* High-Fidelity DNA polymerase and a PCR thermal cycler to replicate both strands of the plasmid with two synthetic primers that both contains the desired mutation. After denaturation of the template DNA, the mutagenic primers, both being complementary to the DNA template, anneal to the plasmid and the region targeted for mutation before the temperature is increased and the primers are extended with *PfuUltra* DNA polymerase. The newly amplified plasmid, containing several staggered nicks, is then

treated with the *Dpn*I endonuclease, specifically targeting deoxyadenosine methylase (dam) methylated DNA. The parental DNA template, originating from a dam methylated *E. coli* strain, is thereby digested, leaving the mutated plasmid intact. For transformation of vector DNA into XL1-Blue Supercompetent cells, see section 3.8.2.

Materials

- QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies)
    - 10x reaction buffer
    - dNTP mix
    - *PfuUltra* HF DNA polymerase
    - *Dpn*I
- DNA template
- Primers (see section 2.5)

Procedure

The parental DNA template containing the *Malpmo10B* gene was isolated from transformed One Shot® *E. coli* TOP10 cells using the NucleoSpin® Plasmid kit (see section 3.6.1). Plasmid DNA concentration was determined using BioPhotometer (Eppendorf) at 260 nm. PCR reactions (50 µl) were prepared on ice in 0.2 ml PCR tubes according to table 3.5. *PfuUltra* HF DNA polymerase was the last reagent added to the mixture.

Table 3.5. PCR reaction setup for MaLPMO10B using QuikChange II XL Site-Directed Mutagenesis Kit. For information about primers used, see section 2.5.

Reagents	**Volume**
10x reaction buffer	5 µl
DNA template (pRSET-B_*Malpmo10B)*	1 µl (209 ng)
Forward primer	1.25 µl (0.25 µM)
Reverse primer	1.25 µl (0.25 µM)
dNTP mix	1 µl

dH2O	40 µl
PfuUltra HF DNA polymerase	0.5 µl

Reactions were thermocycled in a SimpliAmp™ Thermal Cycler (Thermo Fisher) using the parameters outlined in table 3.6.

Table 3.6. Thermocycling conditions for PCR using QuikChange II XL Site-Directed Mutagenesis Kit

Step	**Temperature**	**Time (mm:ss)**	**Number of cycles**
Initial denaturation	95 °C	00:30	1
Denaturation	95 °C	00:30	16
Annealing	60 °C	01:00	
Elongation	68 °C	03:30*	

*1 minute/kb of plasmid length (pRSET-B *Malpmo*10B is approximately 3.9 kb)

Following thermocycling, 1 µl of the *Dpn* I restriction enzyme was added directly to the reaction. Mixing of the reaction was performed by gently pipetting the solution up and down several times, before the reaction mixture was spun down (short spin up to 6000 rpm) using Centrifuge 5418 R (Eppendorf). The reaction was incubated at 37 °C for 1 hour to allow the endonuclease to digest the parental methylated DNA. Subsequently, the mutated DNA was transformed into XL1-Blue Supercompetent cells according to section 3.8.2.

3.4.4 Agarose gel-electrophoresis

Gel electrophoresis was used to separate and analyze the final PCR products based on their size. Separation occurs by applying an electric field to move the negatively charged linear DNA molecules through a matrix of 1 % agarose. The negatively charged DNA molecules will migrate towards the positively charged anode. The shorter molecules will migrate farther than the longer molecules because larger molecules are slowed down through interaction with the pores of the

gel (the larger the fragments, the more interactions and slower migration speed). Staining of DNA was performed using peqGREEN dye (VWR), allowing the visualization of the DNA bands by UV light. The size of the DNA molecules was determined by comparison to a DNA ladder comprised of DNA fragments of known size.

Materials

The following ingredients were used to make one gel consisting of 1 % agarose in a 7 x 10 cm gel tray in the Mini-Sub Cell GT system (Bio-Rad):

- 0.6 g agarose
- 60 ml 1x Tris-acetate-EDTA (TAE) buffer
    - 40 mM Tris base
    - 20 mM acetic acid
    - 1 mM EDTA, pH 8.0
    - Dissolved in dH_2O, leading to a final pH of 8.5
- 3 µl peqGREEN

In this study, the following DNA ladders where used:

- GeneRuler 1 kb DNA Ladder (Thermo Fisher)
- 1 kb DNA Ladder (NEB)

FastDigest Green Buffer (Thermo Fisher) loading buffer was used for analysis of all PCR products except when the Red Taq Master Mix (VWR) was used.

Procedure

1 % agarose gels were made by dissolving the agarose powder in 1x TAE buffer in a 200 ml beaker. The beaker was heated in a microwave oven at maximum power until the solution started to boil. After removing the beaker from the microwave oven, the solution was cooled to about 55 °C before peqGREEN was added directly to the mixture. The gel solution was poured into a UV-transparent gel casting trey with an 8-well spacer in the Mini-Sub Cell GT system, and left to solidify for 30 minutes. When solid, the spacer was removed, and the gel, including the gel chamber, was transferred to an electrophoresis chamber. The gel was then completely covered by

filling the chamber with 1x TAE buffer. The respective DNA ladder, as well as samples, where mixed with 0.1 volumes of 10x FastDigest Green Buffer. The gel was run for 30 minutes at 90 volts using PowerPac™ Basic (Bio-Rad) power supply. DNA bands where visualized with a Benchtop UV Transilluminator, and gel imaging was performed using Gel Doc™ EZ Imager (Bio-Rad) with a UV Sample Tray.

3.4.5 Extraction of DNA fragments from agarose gels

Extraction of size verified PCR products from agarose gels was done using the NucleoSpin® Gel and PCR Clean-up kit, and was performed according to the instructions (Protocol 5.2 DNA extraction from agarose gels) provided by the supplier; however, with some minor changes.

Materials

- 1 % agarose gel, containing DNA fragments
- NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel)
    - Binding buffer NTI
    - Wash buffer NT3, with added ethanol
    - Elution buffer NE
    - NucleoSpin® Gel and PCR Clean-up columns
    - Collection tubes (2 ml)
- Scalpel

Procedure

All centrifugation steps were carried out at room temperature at 11,000 x g using a Centrifuge 5418 R (Eppendorf).

The DNA fragment of interest was excised from the agarose gel with a clean scalpel while working on a Benchtop UV Transilluminator, allowing the visualization of the various DNA bands. The gel piece containing the DNA band was transferred to a 1.5 ml Eppendorf tube, and the weight of the gel piece was determined. For every 100 mg of agarose gel, 200 µl NTI buffer was added, followed by incubation at 50 °C until the gel piece was completely dissolved. A NucleoSpin® Gel and PCR Clean-up column was placed into a 2 ml collection tube and the

sample was loaded onto the column and centrifuged for 30 seconds. The flow-through was discarded and the column was washed with 700 µl NT3 buffer by centrifugation for 30 seconds, and the resulting flow-through was discarded. To remove the NT3 buffer completely, the sample was centrifuged for an additional minute. The NucleoSpin® Gel and PCR Clean-up column was carefully removed from the collection tube and placed into a new 1.5 ml Eppendorf tube. 30 µl NE buffer was added to the sample, and the sample was incubated for 1 minute at room temperature before the DNA was eluted by centrifugation for 1 minute.

3.5 In-Fusion® cloning of LPMO genes into pRSET-B

The In-Fusion HD Cloning kit was used to insert DNA fragments into a linearized vector. The In-Fusion enzyme creates single-stranded 5' overhangs of 15 nucleotides at the ends of the DNA fragments and the linearized vector. Following incubation of the reaction mixture, these overhangs anneal, and the final circular construct have been generated (see figure 3.1 for an overview of the In-Fusion process). For cloning of LPMO genes into pRSET-B, DNA fragments were amplified according to section 3.4.1, and isolated from agarose gels as shown in section 3.4.5. The primers used for PCR amplification of the genes to be inserted, were designed with a 15-bp extension to overlap the restriction site and the corresponding part of the pRSET-B vector. The pRSET-B expression vector was linearized using two sets of restriction enzymes (see section 3.3), and contains ends identical to the primers used for amplification of the LPMO genes. By mixing the DNA fragments and the linearized vector with the In-Fusion HD Enzyme Premix, the insert and vector are fused together, without the use of restriction digestion, phosphatase treatment or ligation reactions.

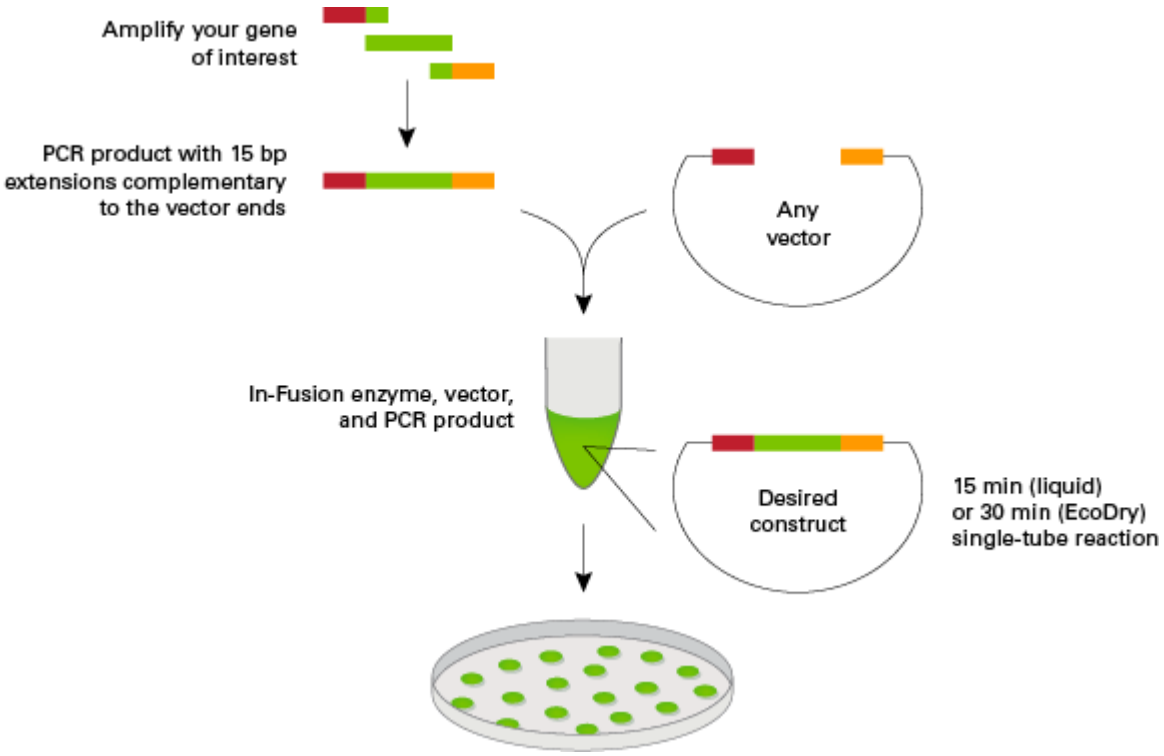


Figure 3.1. Overview of the In-Fusion HD Cloning protocol. *Amplification of the gene of interest by PCR using designed primers extending the gene with 15 base pairs at the termini, homologous to the linearized vector termini. Following 15 minutes of incubation, the In-Fusion enzyme have created single-stranded regions at the termini of the PCR product and the vector. The PCR product and the vector are annealed at the homologous regions and the resulting circular construct is transformed into bacterial expression strains for selective growth of transformants. The picture was taken from http://www.clontech.com/US/Products/Cloning_and_Competent_Cells/Cloning_Resources/FAQs/In-Fusion_Cloning*

Materials

- 5x In-Fusion HD Enzyme Premix
- Nuclease free dH_2O
- TE buffer, pH 8.0
    - 10 mM Tris-HCl
    - 1 mM EDTA
    - Dissolved in dH_2O and pH adjusted to 8.0 with HCl.

Procedure

Cloning of LPMO genes into the pRSET-B expression vector was performed using Protocol I, "In-Fusion Cloning Procedure w/Spin-Column Purification", of the In-Fusion® HD Cloning Kit User Manual. The amount of purified PCR fragment and linearized vector used during this procedure was calculated using the online In-Fusion® Molar Ratio Calculator (http://bioinfo.clontech.com/infusion/molarRatio.do). An insert to vector molar ratio of 2:1 was recommended and used for optimal cloning results.

Cloning reactions were set up according to table 3.7 and table 3.8 for *Malpmo10B* and *Malpmo10D*, respectively. After mixing of ingredients, the reactions were incubated for 15 minutes at 50 °C and subsequently cooled on ice. Before transformation, 90 µl TE buffer was added to the reactions. TE buffer was used as a storage solution for the cloned DNA, protecting it from degradation.

Table 3.7. Reaction setup for In-Fusion HD Cloning of Malpmo10B into pRSET-B.

Reagents	**Volume**
5x In-Fusion HD Enzyme Premix	2 µl (1x)
*Hind*III and *Nde*I linearized vector (pRSET-B)	2.8 µl (100 ng)
Purified PCR fragment (*Malpmo10B*)	2.7 µl (80.9 ng)
Deionized H_2O	2.5 µl

Table 3.8. Reaction setup for In-Fusion HD Cloning of Malpmo10D into pRSET-B.

Reagents	**Volume**
5x In-Fusion HD Enzyme Premix	2 µl (1x)

*Hind*III and *Bsm*I linearized vector (pRSET-B)	1.9 µl (100 ng)
Purified PCR fragment (*Malpmo10D*)	1.3 µl (71.4 ng)
Deionized H_2O	4.8 µl

3.6 Plasmid isolation from *E. coli*

3.6.1 Plasmid isolation using NucleoSpin® Plasmid kit

Following cloning of LPMO genes into the pRSET-B expression plasmid, transformed One Shot® TOP10 *E. coli* cells were inoculated in LB medium supplemented with 50 µg/ml ampicillin and incubated overnight at 37 °C in a shaking incubator. Extraction of plasmid DNA from the bacterial culture was performed using the NucleoSpin® Plasmid kit.

Materials

- NucleoSpin® Plasmid/Plasmid (NoLid) kit (Macherey-Nagel)
    - Resuspension buffer A1, with added RNase A
    - Lysis buffer A2
    - Neutralization buffer A3
    - Wash buffer AW
    - Wash buffer A4, with added ethanol
    - Elution buffer AE
    - NucleoSpin® Plasmid/Plasmid (NoLid) Columns
    - Collection tubes (2 ml)

Procedure

pRSET-B plasmids containing the *Malpmo10B* or the *Malpmo10D* genes, were isolated from One Shot® *E. coli* TOP10 cells following the protocol for isolation of high-copy plasmid DNA from *E. coli* supplied with the NucleoSpin® Plasmid/Plasmid (NoLid) kit. All reaction steps were carried out at room temperature and all centrifugations were done at 11,000 x g using a Centrifuge 5418 R (Eppendorf).

Plasmid isolation was performed using 2 ml of an overnight culture of One Shot® *E. coli* TOP10. The bacterial culture was transferred to a 2 ml microcentrifuge tube and harvested by centrifugation for 30 seconds. The supernatant was discarded and the pellet was resuspended in 250 µl resuspension buffer A1. The buffer helps the cells maintain an optimal pH for subsequent steps, and the added RNase I enzyme digests bacterial RNA after cell lysis. For bacterial cell lysis, 250 µl lysis buffer A2 was mixed with the solution by carefully inverting the tube 6-8 times. The lysis buffer contains sodium dodecyl sulfate (SDS), that disrupts the cell membrane and denatures proteins, but also sodium hydroxide (NaOH), which aids in breaking down the bacterial cell wall. NaOH is also responsible for denaturing double-stranded DNA in the cells, converting both genomic and plasmid DNA into single-stranded DNA. The sample was incubated at room temperature for no longer than 5 minutes, or until the lysate appeared clear. Cell lysis was stopped by adding 300 µl neutralization buffer A3, and the sample was mixed by inverting the tube 6-8 times until the solution appeared completely colorless. The addition of neutralization buffer resulted in renaturation of both genomic and plasmid DNA in the solution; however, only plasmid DNA renatures in its correct conformation due to its circular nature. The long regions of genomic DNA are unable to renature correctly, and ends up as precipitate. The tube was centrifuged for 5 minutes to separate the precipitated genomic DNA from the plasmid DNA. The clear supernatant, containing plasmid DNA, was loaded onto a NucleoSpin® Plasmid/Plasmid (NoLid) column placed in a 2 ml collection tube and centrifuged for 1 minute. Plasmid DNA was bound to the silica membrane of the column, and the flow-through in the collection tube was discarded. The silica membrane was washed by adding 500 µl wash buffer AW (preheated to 50 °C) to the column, and after 1 minute of centrifugation, the silica membrane was washed once again by adding 600 µl wash buffer A4 followed by 1 minute of centrifugation. The washing steps removes any impurities left on the membrane. Wash buffer AW contains guanidine hydrochloride, responsible for removing proteins, while the ethanol added to wash buffer A4 remove salts from the membrane. The flow-through was discarded and the silica membrane was dried by centrifuging the tube for 2 minutes. The NucleoSpin® column was placed in a new 1.5 ml microcentrifuge tube and 60 µl elution buffer AE was added to the column. Following 1 minute of incubation at room temperature, the plasmid was eluted into the new microcentrifuge tube by centrifugation for 1 minute. The sample was stored at -20 °C until further use. Elution

buffer AE is composed of a Tris solution with a slightly basic pH (8.5), ensuring the plasmid DNA remains stable during storage.

3.6.2 Plasmid isolation using PureYield™ Plasmid Miniprep System

Plasmid isolation of *Malpmo10D*^{sl} was performed using the PureYield™ Plasmid Miniprep System; however, the first step of the supplemented protocol (4.A. Centrifugation Protocol) was not done. Instead, the first steps of the NucleoSpin® Plasmid/Plasmid (NoLid) kit (Macherey-Nagel) protocol was performed, ending with the adding of resuspension buffer A1. The reason for this was a lack of lysis buffer A2 from the NucleoSpin® kit. The PureYield™ Plasmid Miniprep System is based on the same principles of plasmid isolation as the NucleoSpin® Plasmid/Plasmid (NoLid) kit.

Materials

- PureYield™ Plasmid Miniprep System (Promega)
    - Cell lysis buffer (CLC)
    - Neutralization solution (NSC)
    - Endotoxin removal wash (ERB)
    - Column wash solution (CWC), with added ethanol
    - Elution buffer (EBB)
    - PureYield™ Minicolumns
    - PureYield™ Collection tubes

Procedure

All reaction steps were carried out at room temperature and all centrifugations were done at 13,000 x g using a Centrifuge 5418 R (Eppendorf).

Plasmid isolation was performed using 2 ml of an overnight culture of One Shot® *E. coli* TOP10 harboring the pRSET-B_*Malpmo10D*^{sl} plasmid. The bacterial culture was transferred to a 2 ml microcentrifuge tube and harvested by centrifugation for 30 seconds. The supernatant was discarded and the pellet was resuspended in 250 µl resuspension buffer A1 (Macherey-Nagel).

For cell lysis, 100 µl cell lysis buffer CLC was added and mixed by carefully inverting the tube 6 times. Lysis was stopped by adding 350 µl cold (4 °C) neutralization solution NSC, and the solution was mixed by inverting the tube 6 times. After 5 minutes of centrifugation, the supernatant (900 µl) was transferred to a PureYield™ Minicolumn placed in a PureYield™ Collection tube and subsequently centrifuged for 15 seconds. The flow-through was discarded and 200 µl endotoxin removal wash ERB was added to the minicolumn. After 15 seconds of centrifugation, 400 µl column wash solution CWC was added, followed by centrifugation for 30 seconds. The minicolumn was transferred to a new 1.5 ml microcentrifuge tube and 30 µl elution buffer EBB was directly added to the minicolumn matrix. Following 1 minute of incubation at room temperature, the tube was centrifuged for 1 minute, eluting the plasmid DNA into the new microcentrifuge tube and stored at -20 °C.

3.7 DNA sequencing

Plasmid DNA isolated from One Shot® *E. coli* TOP10 cells (see section 3.6) were sequenced before transformation into *E. coli* BL21 Star expression strain. DNA sequencing was performed to check for undesired mutations in cloned plasmids from wild type LPMOs (see section 3.5), as well as verifying successful mutagenesis reactions according to section 3.4.3.

For each plasmid to be sequenced, reactions were set up for both forward and reverse pRSET-B primers. In each reaction, 500 ng plasmid DNA and 2.5 µl pRSET-B primer (10 pmol/µl stock) were mixed with dH_2O to a final volume of 11 µl. Plasmid DNA samples were sent to GATC Biotech (www.gatc-biotech.com) for sequencing. All data analysis of sequences was performed using the GENtle software (Manske 2003).

3.8 Transformation of *E. coli*

3.8.1 Transformation of One Shot® Competent Cells

Materials

- S.O.C. medium
- Lysogenic broth (LB) medium

- One Shot® TOP10 chemically competent cells (Invitrogen)
- One Shot® BL21 Star (DE3) chemically competent cells (Invitrogen)
- Plasmid DNA

Procedure

One Shot® TOP10 chemically competent cells were used for plasmid amplification of pRSET-B plasmids containing *Malpmo10* genes (see section 3.5). One Shot® BL21 Star (DE3) chemically competent cells were used for periplasmic expression of wild type *Ma*LPMO10 variants, as well as *Ma*LPMO10B and *Ma*LPMO10B^{cd} mutants. The protocols for transformation of One Shot® chemically competent cells supplied with both TOP10 and BL21 Star (DE3) were similar in regards to transformation procedures. Both SOC and LB media were used for transformation purposes in this study. During incubation with shaking, several different speeds (rpm) were used, ranging from 140 to 180 rpm (see table 3.9).

Chemically competent cells were thawed on ice, and for each transformation, 25 µl competent cells were mixed with plasmid DNA from either In-Fusion PCRs (2.5 µl of the reaction mixture after completed plasmid amplification; see section 3.5) or NucleoSpin purified plasmid (2.5 µl; see section 3.6) in 13 ml polypropylene tubes (Sarstedt). The reaction was incubated on ice for 30 minutes, allowing plasmid DNA to come into close contact with the competent cells. Plasmid DNA entered the cells through the bacterial cell membrane during a 30 seconds heat-shock at 42 °C, before the cells were cooled on ice for one minute to keep the plasmids inside the bacterial cells. The rapid temperature eliminates many of the cells; however, surviving cells harboring the plasmid are recovered with the addition of 250 µl pre-warmed medium. After addition of the pre-warmed medium, the cells were incubated for one hour at 37 °C with shaking, before 50-150 µl of each transformation reaction were spread on pre-warmed LB agar plates supplied with ampicillin (100 µg/ml). The plates were incubated overnight at 37 °C, and pre-cultures of clones were made by inoculation of 5 ml medium containing 50 µg/ml ampicillin with transformants as described in section 3.1.3. Glycerol stocks of transformed cells were made according to section 3.2. Verification of transformed products was performed using PCR with Red Taq DNA Polymerase Master Mix (VWR) as described in section 3.4.2.

3.8.2 Transformation of XL1-Blue Supercompetent Cells

Materials

- S.O.C. medium
- QuikChange II XL Site-Directed Mutagenesis kit (Agilent Technologies)
    - XL1-Blue Supercompetent cells

Procedure

XL1-Blue Supercompetent cells were used for plasmid amplification of vector DNA generated by site-directed mutagenesis, as described in section 3.4.3. During incubation with shaking, several different speeds (rpm) were used, ranging from 150 to 180 rpm.

XL1-Blue Supercompetent cells were thawed on ice, and for each transformation reaction, 45 µl of ultracompetent cells were aliquoted to prechilled 13 ml polypropylene tubes (Sarstedt). From the *Dpn* I-treated vector DNA, 2 µl was transferred to the supercompetent cells. After 30 minutes of incubation on ice, the transformation reaction was heat-shocked in a 42 °C water bath for 30 seconds. The cells were then placed on ice for two minutes, before 500 µl pre-warmed SOC medium were added to the tube. After incubation for one hour at 37 °C with shaking, the cells were spread out on pre-warmed LB agar plates supplemented with ampicillin. The plates were incubated overnight at 37 °C, and pre-cultures of clones were made by inoculation of overnight cultures with transformants as described in section 3.1.3. Glycerol stocks of transformed cells were made according to section 3.2.

3.9 Protein expression

For optimal expression of both wild type and mutant LPMOs, One Shot® *E. coli* BL21 Star (DE3) strains harboring the various plasmid constructs were cultivated in different media (LB or TB, as described in section 3.1.1), and different cultivation conditions (temperature and length of cultivation) were tested. Table 3.9 shows an overview of conditions used for expression of all proteins in this study.

Table 3.9. Overview of conditions used for expression of all proteins in this study. *The table includes volume of cultivation, as well as the temperature and shaking speed used during incubation of the proteins. All bacterial cultures were supplemented with 50 µg/ml ampicillin and*

incubated with shaking overnight. Note: Different incubation temperatures were used for optimal expression of MaLPMO10B^{cd} and MaLPMO10B^{cd} W82Y/N85F, despite the initial assumption that the mutant would be optimally expressed at the same temperature (37 °C) as the wild type protein.

Protein	**Volume (ml)**	**Temperature (°C)**	**Shaking speed (RPM)**
*Ma*LPMO10B	600	30	170
*Ma*LPMO10B W82Y/N85F	500	30	140
*Ma*LPMO10B Q141G	500	30	140
*Ma*LPMO10B W82Y	500	30	160
*Ma*LPMO10B N85F	500	30	160
*Ma*LPMO10B F88L	500	30	160
*Ma*LPMO10B F88Y	500	30	160
*Ma*LPMO10B Q141W	500	30	160
*Ma*LPMO10B D140A	300	30	180
*Ma*LPMO10B^{cd}	150	37	170
*Ma*LPMO10B^{cd} W82Y/N85F	300	30	140
*Ma*LPMO10D	300	30	180
*Ma*LPMO10D^{sl}	140	30	170

3.10 Periplasmic extracts of *E. coli*

Genes encoding *Ma*LPMO10 variants were cloned into the expression vector pRSET-B containing an N-terminal signal peptide (*Ma*LPMO10B, native signal peptide; *Ma*LPMO10D, cbp21 signal peptide). The signal peptide direct translocation of the LPMOs to the periplasmic space before it is cleaved. Lysis of the bacterial cells, using a cold osmotic shock method (described below), releases proteins from the periplasmic space, and thus allowing the extraction of recombinant LPMOs.

Materials

- Spheroplast buffer
    - 68.4 g sucrose
    - 40 ml 1 M Tris, pH 8.0
    - 400 µl 0.5 M EDTA, pH 8.0

      The ingredients were mixed and dissolved in dH_2O to a final volume of 400 ml and kept at 4 °C.
- Sterile dH_2O
- 20 mM $MgCl_2$

Procedure

In this study, a great variety of cultivation volumes were used, depending on how much protein was needed for the research to be done. For this reason, only relative volumes of ingredients are included in the procedure. All centrifugations were performed at 4 °C, using a Beckman Coulter Avanti® J-25 centrifuge, with either a JA-10 or a JA-14 rotor, depending on the volume of cultivation used. All volumes noted below are relative to the initial volume of the bacterial culture.

Overnight cultures of One Shot® *E. coli* BL21 Star (DE3) (see section 3.1.3) were transferred to centrifugation bottles and cells were harvested by centrifugation for 12 minutes at 6,000 rpm. After discarding the supernatant, the pellet was resuspended in 0.1 volumes of ice cold spheroplast buffer, causing disruption of the outer membrane of bacterial cells. After 5 minutes of incubation on ice, the cell suspension was centrifuged for 12 minutes at 8,000 rpm. Following centrifugation, the supernatant was discarded and the cell pellet was incubated at room temperature for 30 minutes. The pellet was then resuspended in $\frac{1}{12}$x volumes of ice cold water and $\frac{1}{240}$x volumes of 20 mM $MgCl_2$, causing the outer membrane of the cells to burst, and release the proteins from the periplasmic space. Separation of cells and periplasmic extracts was done by centrifugation at 8,000 rpm for 10 minutes. The final supernatant, containing the periplasmic extract, was filtered through a 0.45 µm sterile filter (VWR) and stored at 4 °C until further use.

3.11 Protein purification

For protein characterization, it is important to work with highly pure protein solutions. In this study, different purification techniques were tested and optimized for the respective proteins. All proteins were purified by the use of ion exchange chromatography, and most of the proteins were further purified by size-exclusion chromatography. Hydrophobic interaction chromatography was initially tested in attempts to purify *Ma*LPMO10B, although the enzyme did not interact with any of the tested column materials and the conditions used (20 mM Tris, pH 8.0).

3.11.1 Ion exchange chromatography

Purification of proteins by ion exchange chromatography (IEC) is based on separation of proteins according to their affinity to the ion exchanger resin. IEC can be subdivided into anion exchange chromatography and cation exchange chromatography; the former binding negatively charged molecules to a positively charged exchange resin, and the latter involves binding of opposite charges. The chromatographical resin typically contain cellulose or agarose beads modified with positively or negatively charged functional groups. Following loading of a column containing an IEC resin with protein sample, a low-salt binding buffer is applied in order to wash proteins with low binding affinity off the column. Binding affinity is dependent on a protein's isoelectric point (pI) and the pH value of the buffer used. A protein's pI is the pH value at which the protein carries no net charge. At pH values above the isoelectric point, the protein will carry a negative net charge, and thus will bind to an anion exchanger. Likewise, a pH value lower than a protein's pI, leads to a positive net charge and binding to a cation exchanger. Separation of the proteins bound to an ion exchange resin is obtained by increasing the salt concentration in the elution buffer, either stepwise or as a linear gradient. This results in protein release from the column, where proteins with weak and strong ionic interactions with the IEC resin elute first and last, respectively.

For purification of LPMOs in this study, several different buffers were used (table 3.10). Common to all was pH values greater than the theoretical pI of the various LPMOs (calculated using the ProtParam (ExPASy; http://web.expasy.org/protparam) tool; table 3.10), allowing binding to an anion resin.

Materials

- HiTrap DEAE Sepharose FF, 5 ml column (GE Healthcare)
- Binding buffers (see Table 3.8 below). Sterile filtered (0.45 µm)
- Elution buffers (Binding buffer containing 1 M NaCl). Sterile filtered (0.45 µm)
- ÄKTA systems:
    - ÄKTA pure (GE Healthcare)
    - ÄKTApurifier (GE Healthcare)
    - ÄKTAprime plus (GE Healthcare)

***Table 3.10**. Overview of optimal binding buffers used for purification of the various LPMOs in this study. The table also lists the predicted isoelectric point (pI) of the proteins.*

Protein	**Isoelectric point (pI)**	**Binding buffer**
*Ma*LPMO10B	7.10	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B W82Y/N85F	7.09	50 mM Tris-HCl pH 9.6
*Ma*LPMO10B Q141G	7.10	50 mM Tris-HCl pH 9.6
*Ma*LPMO10B W82Y	7.09	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B N85F	7.10	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B F88L	7.10	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B F88Y	7.09	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B Q141W	7.10	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B D140A	7.78	25 mM CAPS pH 10.5

*Ma*LPMO10B^{cd}	6.19	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10B^{cd} W82Y/N85F	6.19	25 mM Bis-Tris propane pH 9.5
*Ma*LPMO10D	6.63	25 mM CAPS pH 10.5
*Ma*LPMO10D^{sl}	6.63	25 mM Bis-Tris propane pH 9.5

Procedure

The HiTrap DEAE FF ion exchange column was connected to an ÄKTA purifier chromatographic system (GE Healthcare), and washed with elution buffer to remove any contaminants in the column and subsequently equilibrate with the binding buffer. The pH of the sample (i.e. periplasmic extract; section 3.10) was adjusted to the pH of the buffer by the addition of concentrated buffer (1 M concentrated buffer was added until concentration of the sample was at the same concentration of the binding buffer used) to avoid diluting the sample more than necessary. The sample was loaded onto the column by the use of a sample pump at a flow rate of 1 ml/min. Next, the column was washed with elution buffer (1 ml/min flow rate) to wash out unbound proteins until the UV baseline was stabilized. The protein of interest was eluted by applying a linear NaCl gradient from 0 % to 50 % elution buffer (i.e. 0-0.5 M NaCl) over 100 minutes, using a flow rate of 3 ml/min. Eluted proteins were detected by online monitoring of the absorption at 280 nm and collected using a fraction collector (3 ml per fraction). The protein content of each fraction was analyzed using SDS-PAGE according to section 3.14. Fractions containing the protein of interest, identified by molecular mass, were pooled and concentrated using Amicon® Ultra-15 Centrifugal Filters with a molecular mass cut-off (MWCO) of 10 kDa (Millipore). For selected proteins, a second purification step was performed using size-exclusion chromatography (see section 3.11.2).

The aforementioned purification steps were subject to several adjustments during this study, primarily to optimize purity of the protein of interest. Also, in some cases the LPMO was eluted

during washing of the column with binding buffer due to low affinity of the LPMO to the ion exchange column. In these cases, the eluted LPMO was collected in the flow-through using a fraction collector. Conveniently, most *E. coli* proteins in the samples of these special cases bound to the column leading to an almost pure LPMO containing flow-through wash fraction. In these cases, LPMO containing fractions were pooled and concentrated before further purification by size-exclusion chromatography.

3.11.2 Size-exclusion chromatography

Size-exclusion chromatography (SEC), also known as gel filtration, separates macromolecules in solution based on their size. The SEC column is composed of a gel matrix (e.g. agarose) that consists of spherical beads that all contains pores of different sizes. Small proteins diffuse into the pores, whereas larger proteins pass by the pores and move through the gel at a higher rate than the smaller proteins.

Materials

- HiLoad 16/60 Superdex 75 PG, 120 ml column (GE Healthcare; Old model)
- HiLoad 16/600 Superdex 75 PG, 120 ml column (GE Healthcare; New model)
- Running buffer: 50 mM Tris-HCl, pH 7.5; 200 mM NaCl. Sterile filtered (0.45 µm)
- ÄKTA systems:
    - ÄKTApurifier (GE Healthcare)
    - ÄKTAprime plus (GE Healthcare)

Procedure

A size exclusion column, a HiLoad 16/60 (in the case of *Ma*LPMO10B, *Ma*LPMO10B^{cd}, *Ma*LPMO10B D140A, *Ma*LPMO10D and *Ma*LPMO10D^{sl}) or a HiLoad 16/600 (in the case of *Ma*LPMO10B^{cd} W82Y/N85F and *Ma*LPMO10B mutants W82Y/N85F, W82Y, N85F, F88L, F88Y, Q141G and Q141Y) was connected to an ÄKTA purifier chromatographic system (GE Healthcare). A concentrated protein solution (≥ 1 ml) obtained from ion exchange chromatography (see section 3.11.1) was applied through a 2 ml sample loop and loaded onto the column with a flow rate of 0.8 ml/min, followed by one column volume (120 ml) of running buffer. Protein elution was monitored by recording absorbance at 280 nm and the eluate was

collected in fractions of 1 ml. Fractions putatively containing the protein of interest were analyzed by SDS-PAGE according to section 3.14. Fractions containing LPMO verified by SDS-PAGE were pooled and concentrated to approximately 1 ml using Amicon® Ultra-15 Centrifugal Filters with a MWCO of 10 kDa (Millipore).

While doing this study, the HiLoad 16/60 Superdex size exclusion column was replaced by the newer HiLoad 16/600 Superdex column. The transition to a new column did not affect purification of the truncated LPMOs; however, it was no longer possible to purify full-length LPMOs. As a result of this replacement, all mutated full-length LPMO versions (except *Ma*LPMO10B D140A) were only purified using ion exchange chromatography (see section 3.11.1).

3.12 Protein concentration measurement

Two different methods for measurement of protein concentration were used: Quick Start™ Bradford Protein Assay, and direct photometric measurement of protein (A_{280}), using the theoretical molecular extinction coefficient to determine protein concentration. Measuring A_{280} was the preferred method for concentration measurement of highly pure proteins, whereas Bradford Protein Assay was used for proteins that were not completely purified (by IEC only).

3.12.1 Quick Start™ Bradford Protein Assay

Quick Start™ Bradford Protein Assay is a method for determining the concentration of proteins in a sample, which is based on the binding of Coomassie Brilliant Blue G-250 dye to proteins (Bradford, 1976). When the dye binds to protein it is converted from the cationic form (A_{max} = 470 nm) to the stable unprotonated blue form (A_{max} = 595 nm), which can be detected using a spectrophotometer. Protein concentration is calculated by using a standard curve made using a protein standard with pre-determined concentration [Bovine serum albumin (BSA)].

Materials

- Protein Assay Dye Reagent Concentrate (Bio-Rad)
- Polystyrene cuvettes, 1.5 ml (Brand)

- Sample buffer: 25 mM Bis-Tris propane, pH 9.5
- Protein of interest

Procedure

Sample dilution was performed before each protein concentration measurement. The spectrophotometer, BioPhotometer (Eppendorf), was pre-calibrated with bovine serum albumin (BSA) as standard with concentrations ranging from 1.25 to 10 µg/ml, using 20 mM Tris-HCl pH 8 as sample solution. The standard curve had a standard error of 6.3 %. The degree of dilution of each protein sample was dependent on the amount of protein in each sample. For some measurements, several dilutions had to be made in order to obtain concentration values that fell within the range of the linear standard curve.

Before measurement of the protein samples, a blank sample was prepared. The reference sample contained 800 µl sample buffer mixed with 200 µl Protein Assay Dye Reagent. After 5 minutes of incubation, the sample was used as a blank to set the A_{595} to zero on the spectrophotometer.

Concentrated protein from each protein sample was diluted in sample buffer to a total of 800 µl and mixed with 200 µl Protein Assay Dye Reagent. All work with protein samples was performed in triplicates. After 5 minutes of incubation, the absorbance at 595 nm was measured by using the Bradford micro program, which is pre-programmed in the BioPhotometer. The sample concentration was calculated as the mean value of the three replicates.

3.12.2 Direct photometric measurement of protein (A_{280})

Measurement of concentration in protein samples by directly measuring absorbance is a fast and convenient method, requiring no additional reagents or incubations. As opposed to the Bradford method, no standard samples are needed to perform the measurements, and if measurements are taken place by using undiluted protein samples, no protein is consumed. Using a spectrophotometer, absorbance of ultraviolet light at a wavelength of 280 nm (A_{280}) in a protein sample is measured. Tryptophan and tyrosine residues are the primary reason for the absorbance peak at 280 nm.

Calculation of concentration values is done according to Beer's law: $A = \varepsilon \times b \times c$. Beer's law states that absorbance (A) is proportional to the concentration (c) of a solution that can absorb electromagnetic radiation. In the equation, b is the path length of the ultraviolet light, and is in this study always 1 cm. The molar absorptivity coefficient (or extinction coefficient), ε, is wavelength-dependent and indicates how much light a given protein absorbs at a certain wavelength. Values of molar extinction coefficient for all proteins in this study can be found in Table 4.1. The values were obtained using ProtParam (ExPASy), a tool which computes various physical and chemical parameters, including the theoretical molar extinction coefficient, for a given amino acid sequence. The tool computes two separate predicted values, in which the value obtained from "assuming all pairs of Cys residues form cysteines" where used.

By using Beer's law, the following equation can be used to determine a protein's molar concentration (c), where A is the absorbance values obtained from direct photometric measurement:

$$c = \frac{A}{\varepsilon \times b}$$

Multiplication of the determined molar concentration value with the protein's molecular weight gives the mass concentration (e.g. g/l) of the protein sample.

Materials

- Uvette® disposable cuvettes (Eppendorf)
- Sample buffer: 10 mM Bis-Tris, pH 6.0
- Protein of interest

Procedure

The A_{280} program of the BioPhotometer (Eppendorf) spectrophotometer was used, and sample buffer was used as a blank. Protein samples were diluted with sample buffer to obtain absorbance values below 1. The resulting absorbance values were then used to calculate the mass concentration of the samples by using Beer's law, as previously described.

3.13 Copper saturation of purified proteins

LPMOs are copper-dependent enzymes, and contain a mononuclear type 2 copper site. Purified LPMOs (see section 3.11) were copper saturated using Cu(II)SO_4 with a 1:3 molar ratio (LPMO:Cu(II)), ensuring that each catalytic module in the sample contains a copper(II) ion. Non-bound copper was removed from the solution by passing the protein through a desalting column (see below)

Materials

- 50 mM Cu(II)SO_4
- PD MidiTrap G-25 (GE Healthcare) kit:
    - PD MidiTrap™ columns containing 3.5 ml of Sephadex™ G-25 medium
- Equilibration buffer: 10 mM Bis-Tris, pH 6.0
- 20 % ethanol (0.45 µm sterile filtered)
- Purified LPMO, concentration range 1.2-30 mg/ml

Procedure

The concentration of purified protein was measured according to section 3.12, and for each mole of protein, three moles of Cu(II)SO_4 was added to the sample. Excess Cu^{2+} was removed by separation through a PD MidiTrap™ column according to the Gravity Protocol supplemented with the PD MidiTrap G-25 kit.

The column was prepared by removing the top and the bottom caps of the column, allowing the storage solution to drain out of the column. Column equilibration was performed by filling the column with 5 ml equilibration buffer, allowing the buffer to enter the packed bed completely. Following discarding of the flow-through, the equilibration process was repeated twice. A maximum of 1 ml of Cu^{2+} saturated protein sample independent on protein concentration was added to the column. However, if the protein sample volume was less than 1 ml, equilibration buffer was added to the column after the protein sample had entered the packed bed completely to adjust the volume up to 1 ml. After discarding the flow-through, a 1.5 ml Eppendorf tube was placed under the column and the protein was eluted by adding 1 ml equilibration buffer to the column. The eluate was collected and protein concentration was measured before the sample was

stored at 4 °C. Only the first milliliter of eluted protein was collected and used for further work, as any additional elute may contain free copper ions that may interfere with enzyme assays.

3.14 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used technique for separation of proteins according to their mass. The use of precast gels and automated gel imaging systems allows safe and rapid visualization and analysis of protein samples in stain-free polyacrylamide gels.

Separation of proteins according to size was performed using 1-D polyacrylamide gel electrophoresis. Subsequent to gel electrophoresis, the proteins were solubilized by adding detergents and reducing agents to the samples. The addition of the anionic detergent, lithium dodecyl sulfate (LDS; see below), and the reducing agent, dithiothreitol (DTT), denatures and reduces protein disulfide bonds and provides each protein with a uniform negative charge. During gel electrophoresis, the negatively charged proteins migrate through the gel, towards the anode. Due to the equal charge-to-mass ratio of all proteins in the sample, separation of the proteins is only based on the protein's molecular weight. The pore structure of acrylamide gels, serving as size-selective filters, allows smaller proteins to migrate through the gel faster than larger proteins when an electric current is applied to the gel. A molecular marker containing proteins with known masses is used to estimate the mass of the sample proteins.

Materials

- Sample Buffer 2x
    - 1 ml NuPage® LDS Sample Buffer 4x (Thermo Fisher)
    - 0.4 ml NuPage® Sample Reducing Agent 10x (Thermo Fisher)

      Both ingredients mixed with dH_2O to a final volume of 2 ml
- Tris/Glycine/SDS Electrophoresis Buffer 10x (Bio-Rad)

  Diluted 10 times in dH_2O before use
- Any kD™ Mini-PROTEAN® TGX Stan-Free™ Precast Gels (Bio-Rad)
- BenchMark™ Protein Ladder (Thermo Fisher)
- Protein sample of interest (concentration range: 0.4-30 mg/ml)

Procedure

To 15 µl protein sample, 15 µl 2x Sample Buffer was added and the sample was boiled for 10 minutes. A Mini-PROTEAN® TGX Stan-Free™ Precast Gel cassette was installed in a Mini-PROTEAN® Tetra Electrode Assembly (Bio-Rad) and the internal chamber was filled with 1x Tris/Glycine/SDS electrophoresis (TGS) buffer. The electrode assembly was placed in a buffer tank (Bio-Rad) and the surrounding chamber was filled with 1x TGS buffer. In the first well, 10 µl protein ladder was applied, followed by loading of 15 µl of each protein sample in the remaining wells. A lid with power cables was mounted on the buffer tank and connected to a PowerPac 3000 (Bio-Rad) power supply. After running the gel for 18 minutes at 280 V, the gel cassette was removed from the electrode assembly and released from the cassette. 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 imaging method uses the intrinsic fluorescence of the aromatic amino acids in proteins to detect the protein bands. It should be noted that for proteins not containing aromatic amino acids, this imaging method will not work. However, all LPMOs contain multiple aromatic amino acids (more than average compared to most other soluble proteins).

3.15 Analysis of enzyme activity

Analysis of enzyme activity of wild type and mutant LPMOs were performed by product analysis by mass spectrometry (MS) and high-pressure liquid chromatography (HPLC). MALDI-TOF MS provided a qualitative analysis of soluble oxidized products formed after incubation of reactions containing LPMOs and various substrates. Two different HPLC methods were used, depending on the substrate used in the enzyme reactions. High-performance anion exchange chromatography (HPAEC) was used when analyzing products formed from cellulose, whereas Ultra performance liquid chromatography [UPLC; i.e. Hydrophobic Interaction Liquid Chromatography (HILIC)] was used to analyze products derived from chitin.

All reactions set up to analyze enzyme activity contained 1 µM LPMO, 50 mM Bis-Tris pH 6.0 (pH 6.39 at room temperature) and 2 mM ascorbic acid. The ascorbic acid was prepared by dissolving L-Ascorbic acid (Sigma Aldrich) in dH_2O to a concentration of 100 mM. The mixture was then sterile filtered into PCR tubes that were stored at -20 °C until use. Thawed ascorbic acid

solutions were not re-used. Substrate concentrations in the reactions were 10 mg/ml Avicel® PH-101, shrimp shell α-chitin or squid pen β-chitin, or 1 mg/ml phosphoric acid swollen cellulose (PASC). All reactions were carried out in 2 ml microcentrifuge tubes and incubated in a ThermoMixer® C (Eppendorf) at 40 °C and with shaking at 1,000 rpm. The incubation times of the reactions varied, depending on the purpose of the experiment. Enzyme reactions used in time-course experiments, were pre-incubated for a period of approximately 30 minutes before the reactions were started, in order to allow binding of the enzyme to the substrate (this is a relatively slow binding equilibrium). Starting of the reactions was achieved by adding ascorbic acid, an electron donor which is required for catalysis by LPMOs.

Enzyme reactions were stopped by separating soluble products from the insoluble substrates (all LPMOs analyzed in this study are only active towards insoluble polysaccharides). Separation was performed by filtration of the samples using MultiScreen® Filter Plates with Durapore® Membrane (Millipore) and a MultiScreen® Vacuum Manifold (Millipore) connected to a Membrane Dry Vacuum Pump/Compressor VP80 (VWR). Filtrates were transferred to new microcentrifuge tubes and stored at -20 °C or 4 °C until further use.

Products arising from LPMO activity vary highly in degree of polymerization (DP), making activity analysis challenging. Therefore, prior to quantitative analysis using HPLC (i.e. HPAEC and UPLC), soluble fractions containing such oligomeric products were further degraded, reducing the complexity of the sample by generating shorter oligosaccharides with a DP of only 2 and 3. Degradation was performed by one or more enzymes, depending on the substrate used and the type of products to be generated. Degradation of soluble oxidized products originating from cellulose substrates was performed using cellulases (1 µM overnight, unless stated otherwise). The endoglucanase Cel5A (Made in-house) from the bacterium *Thermobifida fusca* and the cellobiohydrolase Cel7A (Made in-house) from the fungus *Trichoderma reesei* were added to the samples and incubated overnight at 37 °C to generate mainly DP2-oxidized products. Products formed by chitin oxidation were further degraded by incubating the samples with the chitobiase (CHB; made in-house) from *Serratia marcescens*, allowing quantification of chitobionic acid (GlcNAcGlcNAc1A).

In-house made DP2-oxidized standards were produced by oxidizing a known concentration of cellobiose (Glc_2) or chitobiase [$GlcNAc_2$; Loose et al. (2014)] using cellobiose dehydrogenase

(CDH; 400 µM *Mt*CDH in 20 mM NaOAc pH 5.5) from *Myriococcum thermophilum* and chitooligosaccharide oxidase (ChitO) from *Fusarium gramineearum*, respectively.

3.15.1 Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS)

MALDI-TOF mass spectrometry is usually used as a non-quantitative analysis method used to determine and identify the molecular weight of biomolecules and large organic molecules. Ionization of fragile molecules, such as large peptides, proteins, oligonucleotides and oligosaccharides, requires the use of a soft ionization technique that allows molecules to be ionized without fragmentation. The use of a MALDI matrix, providing both desorption and ionization, ensures that energy is absorbed by the matrix and transferred indirectly to the sample, thereby reducing sample decomposition. Ionization by MALDI is achieved in two steps. First, the sample is dissolved in a matrix, consisting of small organic molecules in solution, on a metal plate. The sample-matrix mixture is then dried, before the metal plate is transferred to the inside of the mass spectrometer. Under vacuum conditions, multiple energy pulses are passed to the sample-matrix mixture from a laser, resulting in ionization of sample molecules by being either protonated or deprotonated. Analysis of sample molecules occur by acceleration of charged molecules from the ion source into a mass analyzer. Time-of-flight (TOF) mass analyzers determines an ion's mass-to-charge ratio by measuring the time it takes for the ion to travel a known distance and subsequently reaching a detector. A difference in applied potential creates a strong electric field which accelerates the ions towards the field-free flight tube where separation of ions occurs before reaching the detector. Ions are separated according to their velocities; smaller ions travelling faster than larger ions.

In this study, MALDI-TOF mass spectrometry was used for product analysis of soluble oxidized products formed after incubation of the LPMOs with various substrates.

<u>Materials</u>

- Matrix:
    - 4.5 mg DHB (2,5-Dihydroxybenzoic acid)
    - 150 µl acetonitrile

    - 350 µl dH_2O

      DHB and acetonitrile were mixed and completely dissolved by vortexing, before dH_2O was added. The matrix solution was stored at 4 °C.
- MTP 384 ground steel BC, MALDI target plate (Bruker)
- MTP Target Frame III (Bruker)
- MALDI-TOF mass spectrometer, Ultraflex (Bruker)
- Sample from LPMO reaction (Concentration range: 0.4-21.9 mg/ml)
- 6.25 mM NaCl solution
- Hairdryer

Procedure

Samples to be analyzed were prepared by separating soluble and insoluble products in the sample solution. If not filtered (see section 3.15) the samples were centrifuged for 5 minutes at 13,000 rpm, creating a supernatant consisting of soluble products and a pellet of insoluble polysaccharide substrate.

Prior to analysis, the samples were diluted in either dH_2O or in NaCl. Dilution (5x) of samples increased the quality of product crystallization on the matrix. Sodium saturation (samples adjusted to 5 mM NaCl) was performed in order to reduce or eliminate potassium adducts from the mass spectra, making mass interpretation easier.

Spotting of the sample on the MALDI target plate was achieved by mixing 2 µl matrix solution with 1 µl of sample solution on a clean spot on the target plate surface. The matrix-sample spot was dried completely by a warm stream of air (achieved by using a hairdryer) before the MALDI target plate was placed on a target frame and inserted into the MALDI-TOF spectrometer. The software FlexControl version 3.3 (Bruker) was used for controlling the system, and FlexAnalysis version 3.3 (Bruker) as well as mMass version 5.5.0 (Niedermeyer THJ and Strohalm M) was used for analysis and processing of data. Most samples were analyzed using laser beam intensities between 40 % and 60 % of maximum capacity, shooting each spot approximately 400 times.

3.15.2 High-Performance Liquid Chromatography (HPLC)

High-performance liquid chromatography is an analytical technique used for separation, identification and quantification of different compounds in a mixture. An HPLC instrument typically includes a (auto)sampler, a column, pumps, and one or more detectors. The sampler contains all samples to be analyzed, and from there they are injected into the mobile phase of the system. The mobile phase carries the sample mixture into the column, which contains a stationary phase that retains the components of the sample, thereby separating their relative position in the mobile phase. The mobile phase and the sample constituents are driven by the pumps through the column, eventually reaching the detector. The type of detector varies depending on the nature of the sample (e.g. UV detector for UV absorbing samples, refractive index for non-UV absorbing samples etc.). HPLC instruments usually contain a UV-detector which monitors the quantity of sample compounds emerging from the column. The detector generates signals proportional to the quantity of the sample compounds. The time it takes for an injected sample and a particular compound reaches the detector is called the retention time. Different compounds have different retention times, resulting in separate peaks along the time axis of a chromatogram.

HPLC was used for both qualitative and quantitative analysis of soluble oxidized products. Qualitative analysis yielded chromatograms showing product profiles (i.e. the distribution of oxidized oligosaccharides produced by a specific LPMO) for each LPMO acting on a specific substrate. Quantitative analysis was performed by measuring peak areas (integrals) to determine the concentration of a specific oligosaccharide in a sample. Standards were used to allow annotation and quantification of the various products analyzed.

3.15.2.1 Analysis of oxidized oligomeric products from cellulose by HPAEC-PAD

Soluble oxidized products released from cellulose substrates by the various LPMOs were analyzed using high-performance anion exchange chromatography (HPAEC) coupled with pulsed electrochemical detection (PAD). Carbohydrates are weak acids, with acid dissociation constants (pK_a) in the range of 12 to 14, that can be transferred into oxyanions by using HPAEC with high pH, and subsequently separated using anion exchange columns.

Materials

- Column: CarboPac PA1 carbohydrate column (Dionex)
- Sample from LPMO reaction (Concentration range: 0.4-21.9 mg/ml)
- Eluents:
    - A: 0.1 M NaOH
    - B: 1 M NaOAc in O.1 M NaOH
    - C: Milli-Q water
- HPAEC systems:
    - Dionex ICS-5000^{+} DC (Thermo Scientific)
    - Dionex ICS-3000 (Thermo Scientific)

Procedure

Analysis of cello-oligosaccharides in soluble fractions generated by the various LPMOs (see section 3.15) was performed by HPAEC using a Dionex BioLC connected to a CarboPac PA1 column. Sample components were separated using 0.1 M NaOH (Eluent A) at a flow rate of 0.25 ml/min and a column temperature of 30 °C. Separation and the subsequent elution of native and oxidized products were performed using a stepwise gradient with 0.1 M NaOH and 1 M NaOAc (Eluent B). Qualitative analyses (longer products; > 2-10 DP) were carried out using the following gradient: 0-10 % B over 10 minutes, 10-30 % B over 25 minutes, 30-100 % B over 5 minutes, 100-0 % B over 1 minute, and 0 % B over 9 minutes. Quantitative analyses (shorter products) were carried out using a steeper gradient, as follows: 0-10 % B over 10 minutes, 10-18 % B over 10 minutes, 18-30 % B over 9 minutes, 30-100 % B over 1 minute, 100-0 % B over 0.1 minute, and 0% B over 13.9 minutes. Eluted cello-oligosaccharides were detected by PAD and chromatograms recorded using Chromeleon 7.0 (Dionex) software.

3.15.2.2 Analysis of oxidized oligomeric products from chitin by UPLC

The activity of Cu(II)-saturated LPMOs on different chitin substrates were analyzed by examining oxidized and native chito-oligosaccharides released from α-chitin and β-chitin. Analysis was performed using an ultra-performance liquid chromatography (UPLC) system, capable of work at much higher pressures and utilizing columns made of much smaller particle sizes than conventional HPLC systems, allowing faster run times and better resolution.

Materials

- Columns: Hydrophilic interaction chromatography (HILIC) column (Acquity UPLC BEH Amide, 50 mm and 150 mm)
- Sample from LPMO reaction (Concentration range: 0.4-21.9 mg/ml)
- Eluents:
    - A: 15 mM Tris-HCl, pH 8
    - B: 100 % Acetonitrile
- UPLC system:
    - 1290 Infinity Binary LC System (Agilent Technologies)

Procedure

Analyses of released chitobionic acid (GlcNAcGlcNAc1A) from chitin by the various LPMOs were performed by UPLC using hydrophilic interaction chromatography (HILIC) column. A short (50 mm) column was used for quantitative analyses, whereas a longer (150 mm) column was used for qualitative analyses. The amount of sample injected to the column was 7 µl and the UPLC was operated with a flow rate of 0.4 ml/min. Using ACN (Eluent A) and 15 mM Tris-HCl pH 8.0 (Eluent B), the gradient was carried out as follows: 78 % A and 22 % B held for 4 minutes, and then a 1 minute gradient to 62 % A and 38 % B. The column was reconditioned by performing a 1 minute gradient back to the initial conditions of 78 % A and 22 % B followed by running at these conditions for 1 minute. Eluted products were detected by measuring absorbance at either 195 nm (qualitative analyses) or 205 nm (quantitative analyses of chitobionic acid). For quantitative analysis, a standard sample of A2ox (Loose et al., 2014) were used. The recorded chromatograms were analyzed using Chromeleon 7.0 software.

3.16 Binding assays

Substrate binding affinities of purified LPMOs were examined by incubating the LPMOs with various carbohydrate substrates. Three separate binding assays were conducted: 1) binding of full-length and truncated *Ma*LPMO10B to Avicel and β-chitin; 2) binding of *Ma*LPMO10D and *Ma*LPMO10D^{sl} to Avicel; and 3) pH-dependent binding of *Ma*LPMO10B to Avicel. All assays were conducted using the same experimental design. The binding was analyzed by measuring

(A_{280}) free (unbound) protein in the samples, collected at certain time intervals, and thus the enzymes binding affinity to specific substrates could be determined. To calculate the amount of unbound protein in each sample, four different groups of samples were set up:

- Buffer sample: A 1.5 ml Eppendorf tube, filled with 800 µl buffer, was used to blank the spectrophotometer before use.
- Control samples: Mixtures of buffer and enzyme; these samples were stored at room temperature during the course of the experiment. Absorbance was measured at the start and at the end of the experiment to observe any changes in protein concentration.
- Background samples: Consisting of buffer and substrate. Background samples were incubated with shaking at 1,000 rpm and 22 °C (ThermoMixer® C, Eppendorf), followed by filtration of samples, collected at the start and at the end of the experiment, using MultiScreen® Filter Plates with Durapore® Membrane (Millipore) and a MultiScreen® Vacuum Manifold (Millipore) connected to a Membrane Dry Vacuum Pump/Compressor VP80 (VWR). Filtrates were transferred to new 1.5 ml Eppendorf tubes and stored at 4 °C, before A280 was measured at the end of the experiment. The purpose of these samples was to measure the absorbance of possible UV absorbing components in the substrate, and thus be able to subtract background coming from the substrate from unbound protein in the enzyme-substrate samples.
- Enzyme-substrate samples: Consisting of substrate, buffer and enzyme. Samples were incubated with shaking at 1,000 rpm and 22 °C, followed by filtration. Filtrates were transferred to new 1.5 ml Eppendorf tubes and stored at 4 °C until the end of the experiment. Samples were collected at certain time intervals, ranging from 2.5 minutes to 280 minutes.

The amount of free protein in a sample was calculated by the following formula:

$$Free\ protein, \% = \frac{A\ (Binding\ sample) - A\ (Background\ average)}{A\ (Control\ samples\ average)} \times 100$$

where A is absorbance values.

For all binding assays, the following concentrations were used:

- Substrate: 10 mg/ml

- Buffer: 50 mM
- LPMO: 0.08 µg/µl

3.16.1 Binding of full-length and truncated *Ma*LPMO10B to Avicel and β-chitin

Materials

- Enzymes:
  - *Ma*LPMO10B (1.77 mg/ml)
  - *Ma*LPMO10B^{cd} (21,91 mg/ml)
- Substrates:
  - Avicel® PH-101 (50 mg/ml stock suspension)
  - Squid pen β-chitin (20 mg/ml stock suspension)
- 1 M Bis-Tris, pH 6.0
- Deionized H_2O
- Uvette® disposable cuvettes (Eppendorf)

Procedure

Absorbance (A_{280}) of all samples were measured using a BioPhotometer (Eppendorf). For each enzyme, the same cuvette was used, measuring absorbance in the order of increasing concentrations.

A buffer sample was prepared, filling a 1.5 ml microcentrifuge tube with 50 mM Bis-Tris pH 6.0, and stored at room temperature. Control samples, consisting of buffer and each respective LPMO, were mixed to a final volume of 300 µl. A280 was measured, and the samples were stored at 4 °C until the end of the experiment. Background samples were mixed to a final volume of 300 µl, consisting of buffer and each respective substrate. After mixing of the samples, 100 µl was collected from each sample and the remaining samples were incubated with shaking until the end of the experiment. The collected samples were filtrated, and the filtrates were stored at 4 °C until the end of the experiment. Binding samples, consisting of buffer, substrate and enzyme were mixed to a final volume of 800 µl. The samples were placed in a thermomixer and samples were

collected after 5, 15, 30, 60, 120 and 240 minutes. The collected samples were filtered, and the filtrates were stored at 4 °C until the end of the experiment.

After 240 minutes, incubation of all samples was completed. The buffer sample was used to zero the spectrophotometer, and absorbance of control samples was measured. Background samples and the remaining enzyme-substrate samples were collected and filtered. Absorbance was measured for all the samples.

3.16.2 Binding of *Ma*LPMO10D and *Ma*LPMO10D^{sl} to Avicel

Materials

- Enzymes:
  - *Ma*LPMO10D (0.39 mg/ml)
  - *Ma*LPMO10D^{sl} (1,69 mg/ml)
- Substrate:
  - Avicel® PH-101 (50 mg/ml stock suspension)
- 1 M Bis-Tris, pH 6.0
- Deionized H_2O
- Uvette® disposable cuvettes (Eppendorf)

Procedure

A_{280} for all samples were measured using a BioPhotometer (Eppendorf). For each enzyme, the same cuvette was used, measuring absorbance in the order of increasing concentrations.

The experiment was conducted according to section 3.16 and 3.16.1 and enzyme-substrate samples were collected at the following time intervals: 1.5, 3, 4.5, 15, 30 and 60 minutes.

3.16.3 pH-dependent binding of *Ma*LPMO10B to Avicel

The assay was conducted to examine the pH-dependency of *Ma*LPMO10B binding to Avicel when mixed with buffers of a wide range of pH values. Four different buffers were used, with pH values ranging from 3.5 to 10.5. Three of the buffers included overlapping pH values to examine

how buffers of the same pH value influence binding affinities. CAPS, pH 10.5, was included to observe the effect of a high pH on binding.

Materials

- Enzymes:
  - *Ma*LPMO10B (1.77 mg/ml)
- Substrate:
  - Avicel® PH-101 (50 mg/ml stock suspension)
- Buffers:
  - 200 mM Citrate, pH 3.5, 4.5, 5.5 and 6.5
  - 200 mM Phosphate, pH 6.5 and 7.5
  - 200 mM Bis-Tris propane, pH 7.5, 8.5 and 9.5
  - 200 mM CAPS, pH 10.5

    All buffers were sterile filtered (0.45 µm)
- Deionized H_2O
- Uvette® disposable cuvettes (Eppendorf)

Procedure

Absorbance of the unbound protein was measured using a BioPhotometer (Eppendorf) at 280 nm. For each buffer with a specific pH value, the same cuvette was used, measuring absorbance in the order of increasing concentrations. The BioPhotometer was blanked before use of each respective buffer.

The experiment was conducted according to section 3.16 and 3.16.1 and enzyme-substrate samples were collected at the following time intervals: 2.5, 5, 15, 30, 60, 120 and 180 minutes.

3.17 Synergy experiment

Synergy has previously been demonstrated for a strict C1 oxidizing and a C1/C4 oxidizing LPMO acting on cellulose (Forsberg et al., 2014b). Analysis of the C4 oxidizing activity of *Ma*LPMO10B mutant W82Y/N85F had demonstrated that the C4 oxidizing activity was nearly abolished (section 4.4.6), and a synergy experiment was therefore conducted to investigate

possible synergistic effects of mixing *Ma*LPMO10B W82Y/N85F (C1 oxidizing and weak C4 oxidizing) with wild type *Ma*LPMO10B (C1/C4 oxidizing). LPMO10C from the bacterium *Streptomyces coelicolor* was included to investigate possible synergistic effects when incubating this LPMO with wild type or mutant (i.e. W82Y/N85F) *Ma*LPMO10B.

Materials

- LPMOs:
    - *Ma*LPMO10B (1.77 mg/ml)
    - *Ma*LPMO10B W82Y/N85F (4.77 mg/ml)
    - *Sc*LPMO10C (3.82 mg/ml)
- Substrate:
    - PASC (10 mg/ml stock suspension)
- 1 M Bis-Tris, pH 6.0
- Ascorbic acid (100 mM)
- Deionized H_2O
- Cel5A from *Thermobifida fusca* (58 µM)

Procedure

Reaction mixtures containing 1 mg/ml PASC, 50 mM Bis-Tris pH 6.0 and 1 µM LPMO in a final volume of 300 µl were prepared in 2 ml microcentrifuge tubes. In samples where two different LPMOs were included, 0.5 µM of each enzyme were combined according to table 3.11:

Table 3.11. Overview showing concentration of enzyme added to the various samples.

Enzymes	**Concentration of enzymes added**
*Ma*LPMO10B	1 µM
*Ma*LPMO10B W82Y/N85F	1 µM
*Sc*LPMO10C	1 µM
*Ma*LPMO10B + *Ma*LPMO10B W82Y/N85F	0.5 µM + 0.5 µM
*Ma*LPMO10B + *Sc*LPMO10C	0.5 µM + 0.5 µM

*Ma*LPMO10B W82Y/N85F + *Sc*LPMO10C	0.5 µM + 0.5 µM

Following incubation with shaking at 1,000 rpm and 40 °C for 20 minutes (ThermoMixer® C, Eppendorf), the reactions were started by adding ascorbic acid to a final concentration of 2 mM in each sample. From each sample, 60 µl of the reaction mixture was collected 1, 2, 3 and 6 hours after the addition of ascorbic acid, followed by filtration of samples using MultiScreen® Filter Plates with Durapore® Membrane (Millipore), and a MultiScreen® Vacuum Manifold (Millipore) connected to a Membrane Dry Vacuum Pump/Compressor VP80 (VWR). Filtrates were transferred to new 1.5 ml Eppendorf tubes and stored at -20 °C.

Before performing quantitative analysis using HPAEC-PAD, all solubilized products were further degraded by the endoglucanase *Tf*Cel5A in order to simplify quantification (see section 3.15). From all filtered samples, 35 µl were transferred into new 1.5 ml Eppendorf tubes. The samples were then mixed with 35 µl of a solution containing 1 µM *Tf*Cel5A in 100 mM Bis-Tris pH 6.0 and dH_2O, resulting in a final concentration of 0.5 µM *Tf*Cel5A and 50 mM Bis-Tris pH 6.0 in the samples (and a 2-fold dilution of the LPMO products). The samples were incubated overnight at 37 °C, followed by quantitative analysis by HPAEC-PAD.

3.18 Protein crystallization

Determining the structure of proteins require growth of protein crystals and subsequent analysis by X-ray crystallography. Protein crystals are composed of an ordered three-dimensional array of identical protein molecules, and by radiating a crystal with a beam of X-rays, numerous scattered beams are produced. The scattered beams hit a detector, creating a diffraction pattern of spots. Gradually rotating the crystal will allow the recording of intensities and angles of each scattered beam, creating a three-dimensional map of scattered spots. The position of every atom in the crystallized protein molecule is then calculated using a computer.

In this study, two LPMOs were successfully crystallized with subsequent data collection and their structures determined. The method of hanging drop vapor diffusion was used to grow crystals of the catalytic domain of both wild type and mutant *Ma*LPMO10B (*Ma*LPMO10B^{cd} and *Ma*LPMO10B^{cd} W82Y/N85F; residues 37-230). In this method, a droplet containing a mixture of

protein sample and reagent (crystallization condition) is placed on an inverted siliconized cover slide, which is then placed on top of a liquid reservoir containing the same reagent but in higher concentration. Over time, water vapor leaves the hanging drop and ends up in the liquid reservoir, and at the same time the reagent concentration in the sample drop increases due to loss of water. Equilibrium is reached when the reagent concentration is the same in the drop as in the reservoir. Crystal growth in the drop will occur if the correct crystallization conditions have been used. Finding the correct crystallization condition can be difficult and time-demanding; however, using a sparse matrix design, numerous different crystallization conditions can easily be screened. In this study, the JCSG-*plus*™ MD1-37 kit, containing 96 different crystallization conditions, was used for initial screening experiments; however, this was only done with *Ma*LPMO10B^{cd}. Due to the similar sequences of the wild type and the mutant, subsequent optimization experiments were successfully performed using the same initial condition for both proteins.

3.18.1 Crystallization of *Ma*LPMO10B^{cd}

Materials

- JCSG-*plus*™ MD1-37 (Molecular Dimensions)
- VDXm Plate with sealant, 24 well (Hampton Research)
- VDX Plate with sealant, 48 well (Hampton Research)
- Siliconized Cover Slides, 18 mm (Hampton Research)
- Siliconized Cover Slides, 12 mm (Hampton Research)
- 1 M potassium phosphate monobasic
- 40 % w/v PEG 8000
- 85 % v/v glycerol
- *Ma*LPMO10B^{cd} (21.9 mg/ml)

Procedure

The JCSG-*plus*™ MD1-37 kit was used for initial screening experiments of *Ma*LPMO10B^{cd}. The protein was stored in 10 mM Bis-Tris pH 6.0 sample solution, and protein concentration was 21.9 mg/ml. Two 48 well VDX plates with sealant were used, and into each well, 200 µl of each respective reservoir solution was transferred. Onto 12 mm siliconized cover slides, 1 µl protein

sample was mixed with 1 µl of each respective reservoir solution. Using a tweezer, the slides were placed on top of each corresponding well, covering the wells completely, with the drops hanging on the underside of each slide. Both plates, containing all 96 conditions mixed with protein, were incubated at room temperature, and examined on a regular basis using a microscope.

*Ma*LPMO10B^{cd} successfully crystallized in five different conditions; however, all crystals were in the shape of needles (1D growth). Further optimization experiments were carried out using the crystallization condition which yielded the largest and most defined crystals. Optimization of 0.04 M potassium phosphate monobasic, 16 % w/v PEG 8000 and 20 % v/v glycerol were performed using 24 well VDXm plates with sealant and 18 mm siliconized cover slides. Stock solutions of each reagent were mixed to a volume of 1 ml, with different concentrations, into each well. Cover slides with protein sample and crystallization solution were prepared as previously described, and incubated at room temperature. Several optimized conditions yielded crystals, in the shape of large single crystals (3D growth). Crystals of optimized conditions were transferred to cryo vials and flash frozen in liquid nitrogen. The vials were stored in a nitrogen dewar during shipment to the European Synchrotron Radiation Facility (Grenoble, France) for analysis.

3.18.2 Crystallization of *Ma*LPMO10B^{cd} W82Y/N85F

Materials

- VDXm Plate with sealant, 24 well (Hampton Research)
- Siliconized Cover Slides, 18 mm (Hampton Research)
- 1 M potassium phosphate monobasic
- 40 % w/v PEG 8000
- 85 % v/v glycerol
- *Ma*LPMO10B^{cd} W82Y/N85F (18.06 mg/ml)

Procedure

Due to the success of producing crystals of *Ma*LPMO10B^{cd}, no initial screening experiments were performed with *Ma*LPMO10B^{cd} W82Y/N85F. Instead, optimization experiments were

carried out, using the same optimization setup as with *Ma*LPMO10B^{cd}. Crystals of *Ma*LPMO10B^{cd} W82Y/N85F were obtained as described in section 3.19.1, using concentrated protein in 10 mM Bis-Tris pH 6.0 storage solution, and a protein concentration of 18.06 mg/ml. Crystals of optimized conditions were transferred to cryo vials and flash frozen in liquid nitrogen. The vials were stored in a nitrogen dewar during shipment to a synchrotron for analysis.

3.18.3 Data collection and structure determination

Data collection and structure determination were performed by Dr. Bjørn Dalhus, Oslo University Hospital

Data sets for *Ma*LPMO10B^{cd} and *Ma*LPMO10B^{cd} W82Y/N85F crystals were collected at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, using the MASSIF ID30A beamline. A data set of *Ma*LPMO10B^{cd} was collected at 1.08 Å. The data was integrated, scaled and analyzed using XDS, Aimless and CCP4i. The structure was solved by molecular replacement using Phaser using a model of LPMO10B from *Streptomyces coelicolor* (PDB 4OY6). The model was rebuilt into the sequence of *Ma*LPMO10B^{cd} by using PHENIX and final refinement and adjustments of the model were performed using RefMac5 and Coot.

3.19 Bioinformatics

Table 3.12. Overview of various bioinformatics tools used during this study.

Tools	**Purpose**	**Web address**
Expasy ProtParam	Computation of physical and chemical parameters of LPMOs	http://web.expasy.org/protparam
Pfam	Annotation of domain structures of LPMOs	http://pfam.xfam.org

CAZy	Annotation of carbohydrate-active enzymes	http://www.cazy.org
T-Coffee Server Expresso	To make multiple sequence alignments (MSA) of various LPMOs	http://tcoffee.crg.cat/apps/tcoffee/do:expresso

4 Results

4.1 Bioinformatics

4.1.1 Physicochemical properties and domain structure

Physical and chemical properties of all proteins described in this study were calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam, (Gasteiger et al., 2005)). The physicochemical properties of *Ma*LPMO10B (UniProt entry D9SZQ3; GenBank entry ADL45185) and *Ma*LPMO10D (UniProt entry D9T1F0; GenBank entry ADL47477) variants are shown in table 4.1.

Table 4.1. ***Physicochemical properties of MaLPMO10B and MaLPMO10D variants.*** *The physical and chemical properties of all LPMOs expressed in this study were calculated using the ExPASy ProtParam tool. Calculations do not include signal peptides; only mature proteins and their truncated and/or mutated forms. The table shows the number of base pairs (bp) and amino acids (aa), molecular weight (mw), isoelectric point (pI), and extinction coefficient (ε) of each protein. The number of base pairs include the entire codon optimized gene (MaLPMO10D does not include a signal sequence; see section 4.1.3), whereas the number of amino acids only include the mature protein.*

Protein	**bp**	**aa**	**Mw (Da)**	**pI**	**ε**
*Ma*LPMO10B	1,101	330	34,906.2	7.10	93,765
*Ma*LPMO10B W82Y/N85F			34,919.3	7.09	89,755
*Ma*LPMO10B Q141G			34,835.2	7.10	93,765
*Ma*LPMO10B W82Y			34,883.2	7.09	89,755
*Ma*LPMO10B N85F			34,939.3	7.10	93,765
*Ma*LPMO10B F88L			34,872.2	7.10	93,765
*Ma*LPMO10B F88Y			34,922.2	7.09	95,255
*Ma*LPMO10B Q141W			34,964.3	7.10	99,265
*Ma*LPMO10B D140A			34,862.2	7.78	93,765
*Ma*LPMO10B^{cd}	690	194	21,559.8	6.19	67,170
*Ma*LPMO10B^{cd} W82Y/N85F			21,569.8	6.19	63,160

*Ma*LPMO10D	963	320	34,017.4	6.63	100,755
*Ma*LPMO10D^{sl}	933	310	33,030.3	6.63	100,755

Analysis of the domain structures of *Ma*LPMO10B and *Ma*LPMO10D was performed using the Pfam database (http://pfam.xfam.org). According to this database, both proteins consists of a signal peptide, an N-terminal auxiliary activity family 10 (AA10) LPMO domain and a family 2 carbohydrate binding domain (CBM2; Figure 4.1). The AA10 domain is connected to the CBM2 domain through a low complexity linker region containing several PPPTT repeats. The amino acid sequences of *Ma*LPMO10B and *Ma*LPMO10D are shown in Appendix C.

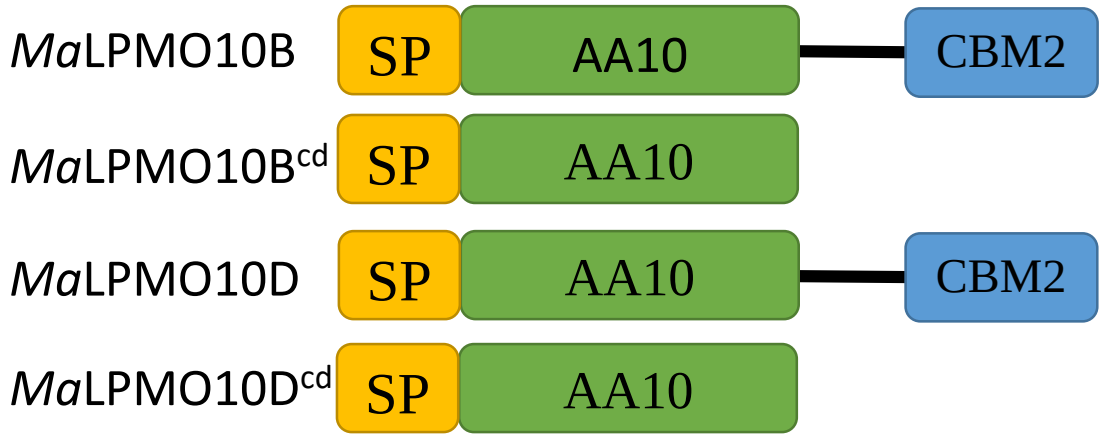


Figure 4.1. Domain structure of MaLPMO10B and MaLPMO10D variants. *All four proteins contain a signal peptide (SP; yellow color) and a N-terminal AA10 LPMO domain (AA10; green color). MaLPMO10B and MaLPMO10D contain a CBM2 domain (blue color) connected to the AA10 LPMO domain through a linker region (solid black line). MaLPMO10B^{cd} and MaLPMO10D^{cd} do not contain a CBM (are thus referred to as truncated proteins), and therefore consists only of a AA10 LPMO domain and a signal peptide.*

4.1.2 Multiple sequence alignment (MSA)

A multiple sequence alignment (MSA) was made to compare the amino acid sequence of four C1/C4-oxidizing and three C1-oxidizing LPMOs (Figure 4.2). The complete amino acid sequence of each LPMO was used for the MSA. All LPMOs contain an N-terminal signal peptide joined to the catalytic domain. All but two of the LPMOs contain a CBM (i.e. CBM2) domain and an associated linker region. LPMO10B from the bacterium *Thermobifida fusca* contains an

additional domain (Fibronectin type III domain) located between the catalytic domain and the CBM domain. Residues of *Ma*LPMO10B that were targeted for site-directed mutagenesis (highlighted red and blue in the MSA) are shown in the structure of *Ma*LPMO10B in Figure 4.3.

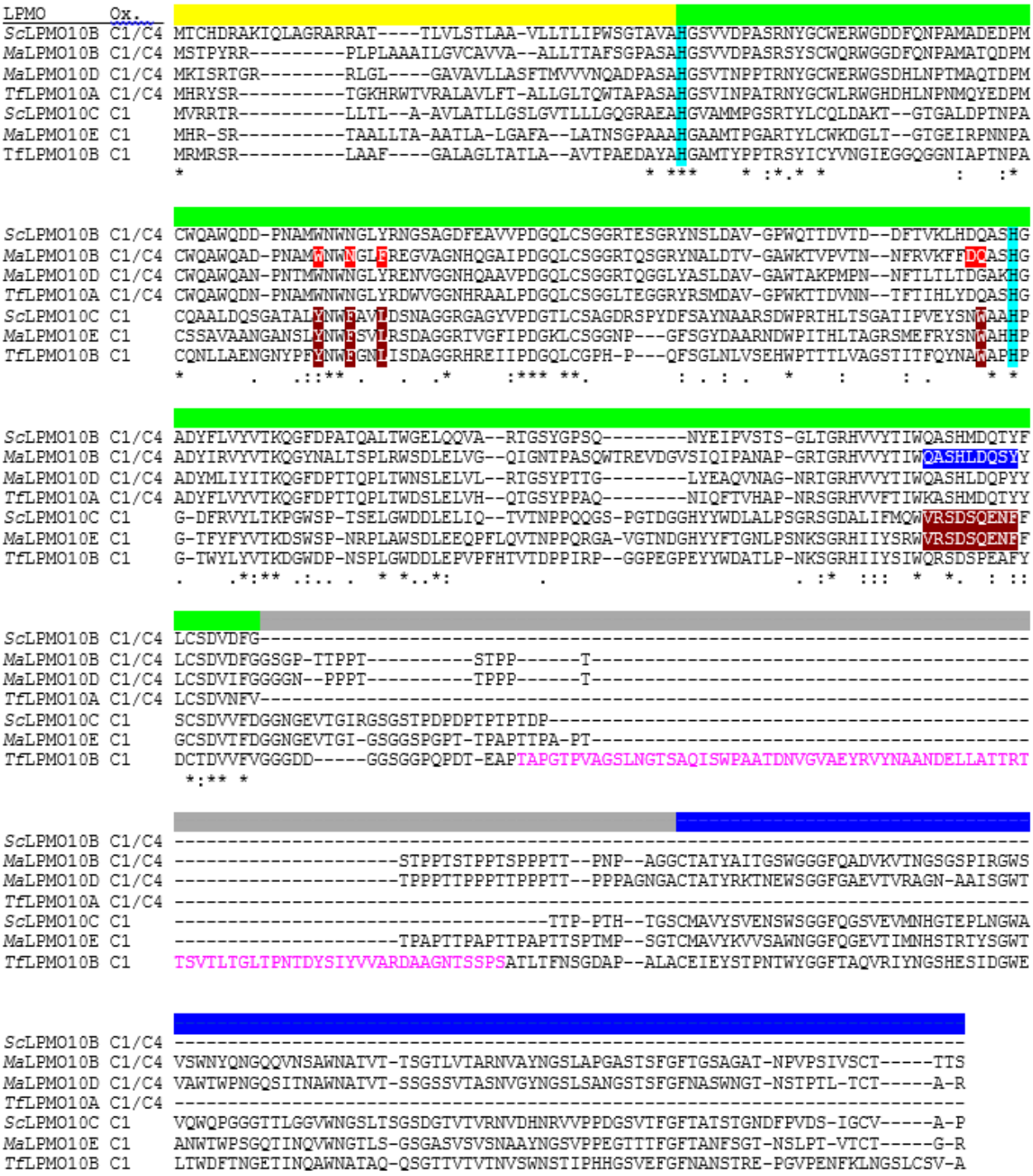


Figure 4.2. Multiple sequence alignment (MSA) of full-length C1/C4- and C1-oxidizing LPMOs. *Domain structures are indicated by colored lines above the sequences. All the LPMOs contain a signal peptide (yellow) joined to a catalytic domain (green). The catalytic domain is connected to a CBM domain (blue) through a linker region (grey). ScLPMO10B and TfLPMO10A do not contain a CBM domain. TfLPMO10B contains a fn3 domain (colored pink in the sequence) connected to catalytic and CBM domains by linker regions. Histidines coordinating the copper ion are highlighted in turquoise. Residues of MaLPMO10B targeted for site-directed mutagenesis are highlighted in red (single or double point*

mutations) or blue (mutation of a short sequence), and residues highlighted in brown where the target for each of the corresponding mutations. Conserved residues are marked by symbols below the sequence, where fully conserved residues are marked with an asterisk (*) and residues with strongly or weakly similar properties are marked with a colon (:) or a period (.), respectively.

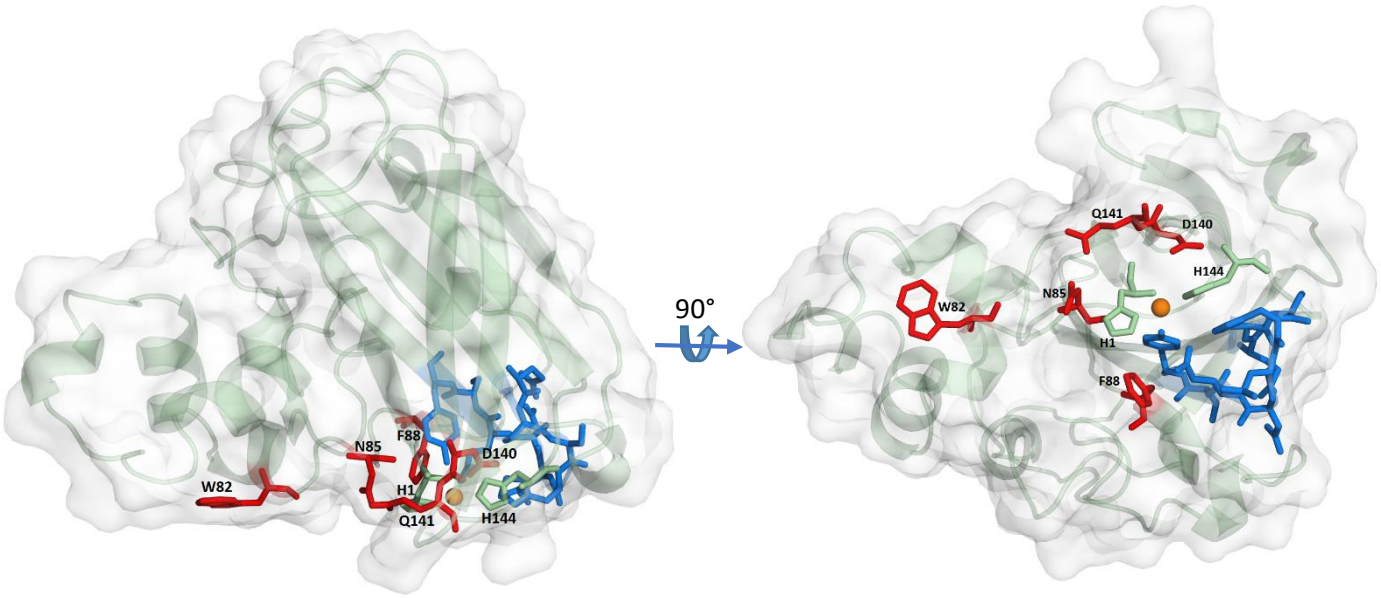


Figure 4.3. Residues of MaLPMO10B targeted for site-directed mutagenesis. Targeted residues and copper active site histidines are represented as sticks. The copper ion is represented as a sphere and is colored orange. Residues involved in single and double-mutations are colored red. Residues comprising the mutated loop are colored blue. Figure made using PyMOL (Delano & Lam, 2005).

4.1.3 Codon optimization

The genes *Malpmo*10B and *Malpmo*10D were both expressed recombinantly in *E. coli* BL21 Star™ (DE3) cells. Since both originates from the Gram-positive bacterium *Micromonospora aurantiaca*, the genes were codon optimized for optimal *E. coli* expression before being cloned into the pRSET-B expression vector. Multiple codons often encode the same amino acid, and for each amino acid, a tRNA with a corresponding anti-codon exists. The abundancy of each specific tRNA varies between different species, and to optimize protein expression, the more abundant tRNAs have to be utilized during translation. By introducing single nucleotide mutations to a DNA sequence, codon usage can be altered without affecting the translated protein sequence.

Wild type sequences of *MaLPMO10B* and *MaLPMO10D* were codon optimized for expression in *E. coli* by OptimumGene™ (GenScript®). Guanine-Cytosine (GC)-content of the optimized genes was reduced from 70.93 % to 60.16 % for *MaLPMO10B,* and from 69.15 % to 60.05 % for *MaLPMO10D.* The optimized sequences of both genes are included in Appendix B.

4.2 Cloning, mutagenesis and transformation

4.2.1 Cloning of *Malpmo10B* and *Malpmo10D*

The expression vector used for expression of *Ma*LPMO10B and *Ma*LPMO10D in *E. coli* was a pRSET-B (see Appendix A for plasmid map of pRSET-B) variant that contained the *cels2-n2* gene and a sequence in the upstream region encoding a signal peptide derived from the *Serratia marcescens cbp21* gene (Forsberg et al., 2014a). In order to facilitate cloning of the *Ma*LPMO10B and *Ma*LPMO10D genes, the *cels2-n2* gene including and excluding the signal peptide sequence was excised using the following restriction endonucleases:

- Including *cbp21* signal sequence: *Bsm*I and *Hind*III
- Removing *cbp21* signal sequence: *Nde*I and *Hind*III

Figure 4.4 shows the linearized pRSET-B expression vectors after treatment with restriction endonucleases, i.e. either containing or lacking the *cbp21* signal sequence.

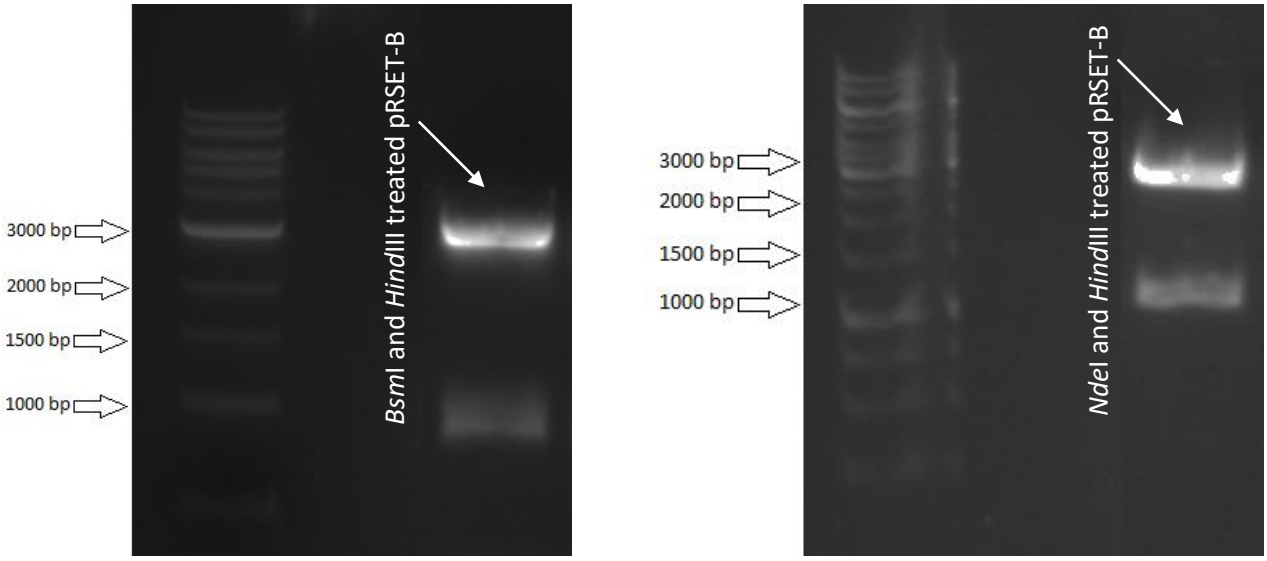


Figure 4.4. Agarose (1 %) gel electrophoresis of double restriction digested pRSET-B expression vector. *The pRSET-B_cels2-n2 expression vector was treated with HindIII and BsmI (left) and HindIII and NdeI (right), resulting in linearized vectors (upper bands) both with (left) and without (right) the cbp21 signal sequence. The left lane shows a 1 kb DNA ladder (NEB). The sizes of selected marker fragments are labeled. The predicted sizes of the PCR fragments were 2,835 bp for HindIII and BsmI treated pRSET-B; and 2,754 for HindIII and NdeI treated pRSET-B.*

Codon optimized genes (see section 4.1.3) encoding *Ma*LPMO10B and *Ma*LPMO10D were amplified by PCR (see section 3.4.1) using primers listed in section 2.5. Two different reverse primers (see section 2.5) were used for cloning the full-length LPMO and the catalytic LPMO domain (designated as superscript cd for catalytic domain). The following LPMO genes were cloned: *Malpmo*10B, *Malpmo*10D and *Malpmo*10D^{cd}, yielding PCR products of expected size (Figure 4.5).

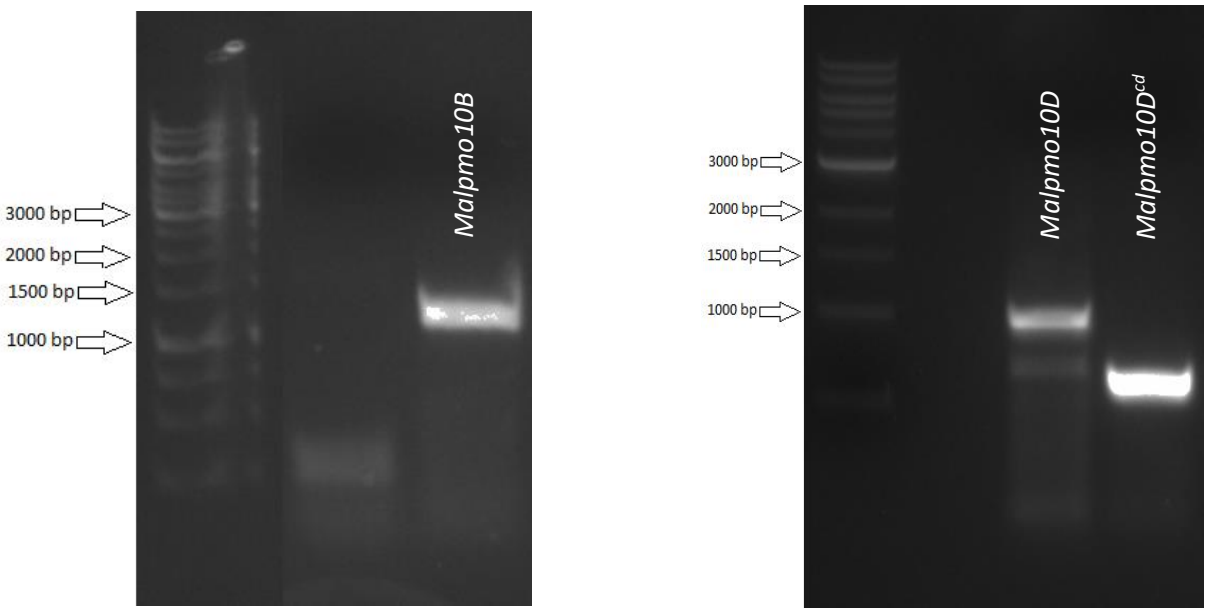


Figure 4.5. Agarose (1 %) gel electrophoresis of PCR amplified Malpmo10B and Malpmo10D genes. *The agarose gel images show DNA fragments obtained by PCR using Phusion® High-Fidelity DNA Polymerase. The left lane shows a 1 kb DNA ladder (NEB). The sizes of selected marker fragments are labeled. The predicted sizes of the PCR fragments were 1,132 bp for Malpmo10B, 996 bp for Malpmo10D, and 594 bp for Malpmo10D^{cd}.*

Following extraction of DNA fragments from agarose gels (see section 3.4.5), all three amplified gene fragments were cloned into linearized pRSET-B using the In-Fusion HD Cloning kit as described in section 3.5. The following constructs were obtained:

- pRSET-B/*Malpmo10B*, with native signal sequence
- pRSET-B/*Malpmo10D*, with *cbp21* signal sequence
- pRSET-B/*Malpmo10D*^{cd}, with *cbp21* signal sequence

The results indicated successful cloning of *Malpmo*10B, while sequencing of *Malpmo*10D and *Malpmo*10D^{cd} were incomplete. Sequencing results of *Malpmo*10D revealed 60 missing nucleotides in the linker region; however, the rest of the protein, including the LPMO and the CBM domains, remained intact, without a change in the reading frame. No further work was done using this plasmid. A second attempt to clone *Malpmo*10D resulted in 30 missing nucleotides in the linker region, again with the rest of the protein intact. The latter plasmid was further used in this study, and was named pRSET-B/*Malpmo10D*^{sl} (sl for shortened linker). A third attempt to clone *Malpmo*10D was successful. An analysis of the linker region of all three

*Ma*LPMO10D genes are included in Appendix C. All attempts to clone *Malpmo*10D^{cd} were unsuccessful, and all further work involving *Malpmo*10D^{cd} was discontinued.

In addition to the successfully cloned plasmids described above, the following pre-made plasmid constructs were used in this study:

- pRSET-B/*Malpmo10B* (*cbp21* signal sequence)
- pRSET-B/*Malpmo10B*^{cd} (native signal sequence)

All plasmids were subsequently transformed into chemically competent One Shot® BL21 (DE3) cells (see section 3.8.1) and cultivated for protein expression.

4.2.2 Site-directed mutagenesis

The pRSET-B/*Malpmo10B* (native signal sequence) plasmid (see section 4.2.1) was used as a template for site-directed mutagenesis of most amino acids mutated in this study. The pRSET-B/*Malpmo10B*^{cd} (native signal sequence) plasmid was used as a template for site-directed mutagenesis of W82Y/N85F in truncated MaLPMO10B. Mutants were generated using the QuikChange II XL Site-Directed mutagenesis kit as described in section 3.4.3. Mutated expression vectors were verified by DNA sequencing, before being transformed into chemically competent One Shot® BL21 (DE3) cells (see section 3.8.1). Sequencing results were analyzed using GENtle (Mantle 2003) and the results are listed in table 4.2.

Table 4.2. Overview of results from DNA sequencing of mutated expression vectors. *The table shows which mutations were successfully introduced into the pRSET-B/Malpmo10B (native signal sequence) plasmid. Section 2.5 contains an overview of all primers used for site-directed mutagenesis in this study.*

Mutations	**Successful (+) or unsuccessful (-)**
W82Y	+
N85F	+
W82Y/N85F	+
F88L	+
F88Y	+
D140A	+
Q141G	+

Q141W	+
213QASHILDQSY → 213VRSDSQENF ("Loop")	-
W82Y/N85F/Q141G	-
*Ma*LPMO10B^{cd} W82Y/N85F	+

All successfully mutated expression vectors were subsequently transformed into chemically competent One Shot® BL21 (DE3) cells (see section 3.8.1) and cultivated for protein expression.

4.3 Protein expression and purification

4.3.1 Protein expression

For protein expression, fresh colonies or pieces of glycerol stocks were inoculated into LB or TB medium supplemented with 50 µg/ml ampicillin. Based on previous studies (Forsberg et al., 2014a and 2014b), incubation of bacterial cultures was tested at both 30 °C and 37 °C. Expression of both full-length and truncated *Ma*LPMO10B, as well as *Ma*LPMO10D^{sl}, were successful (Figure 4.6). The results indicate that *Ma*LPMO10B containing a native signal peptide was best expressed at 30 °C. The results also indicate that both native and CBP21 signal peptides leads the expressed protein to the periplasmic space of the One Shot® BL21 (DE3) cells. All four periplasmic fractions (i.e. native and CBP21 signal peptide, expressed at 30 and 37 °C) of *Ma*LPMO10B were combined prior to protein purification. Optimal expression of *Ma*LPMO10B^{cd} and *Ma*LPMO10D^{sl} was achieved at 37 °C and 30 °C, respectively (Figure 4.6, Panel B). All mutants, of both full-length and truncated *Ma*LPMO10B, were successfully expressed.

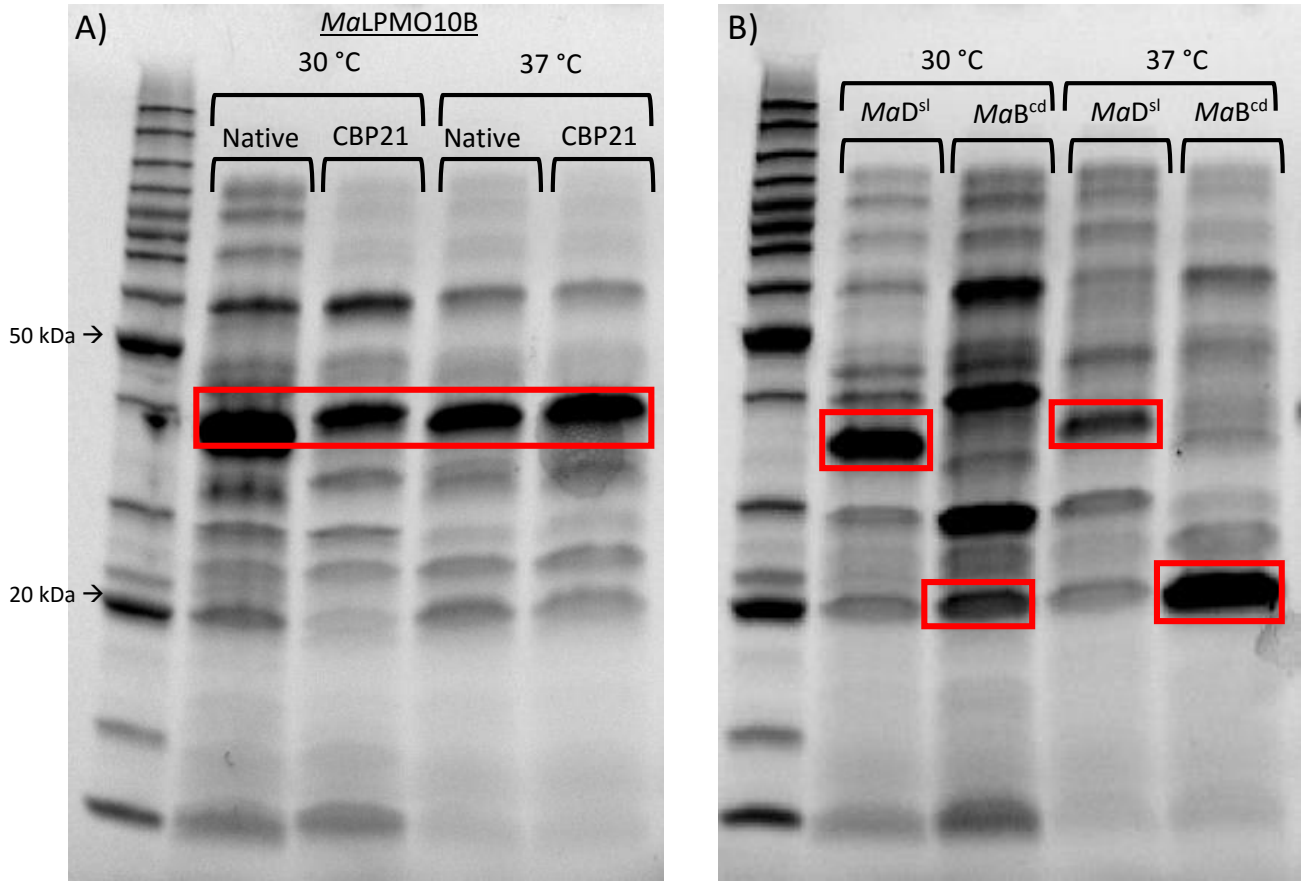


Figure 4.6. Expression of MaLPMO10B, MaLPMO10B^{cd} and MaLPMO10D^{sl}. *The figure shows two SDS-PAGE gels containing a protein marker (Thermo Fisher) and periplasmic extracts from MaLPMO10 producing strains. Two of the marker proteins on the left gel are labeled by their molecular weight. Proteins of interest are indicated by a red square. Panel A) shows periplasmic extracts containing MaLPMO10B incubated at both 30 °C and 37 °C. Panel B) shows periplasmic extracts of MaLPMO10B^{cd} and MaLPMO10D^{sl} incubated at 30 °C and 37 °C. Abbreviations: MaD^{sl}, MaLPMO10D^{sl}; MaB^{cd}, MaLPMO10B^{cd}. Predicted sizes of proteins: MaLPMO10B, 34.9 kDa; MaLPMO10B^{cd}, 21.6 kDa; MaLPMO10D^{sl}, 33.0 kDa.*

4.3.2 Protein purification

Purification of all proteins in this study was performed by ion exchange chromatography as described in section 3.11.1. Not all LPMOs bound fully to the ion exchange column (Table 4.3), some bound with weak affinity to the column and eluted during washing of the column with binding buffer. Analysis of eluted LPMOs using SDS-PAGE, revealed that LPMO eluted during washing of the column appeared purer than LPMO eluted during application of a linear salt

gradient (see figure 4.6). Most LPMOs did not bound at all, and eluted in its entirety during the washing procedure; however, as described above, these LPMO fractions were still fairly pure. Selected proteins were further purified by size-exclusion chromatography (Table 4.3). Figure 4.7 shows SDS-PAGE analyses of LPMOs purified by both ion exchange chromatography and size-exclusion chromatography. An SDS-PAGE analysis, including all purified LPMOs in this study, is shown in figure 4.8.

Table 4.3. Overview of proteins purified by ion exchange chromatography (IEC) and size-exclusion chromatography (SEC). *The table shows which LPMOs bound to the IEC column and which proteins were further purified by SEC. LPMOs that bound weakly to the IEC column eluted both as flow-through during application of binding buffer and early during application of a linear salt gradient (elution buffer). Purification of the listed LPMOs by SEC were performed using the HiLoad 16/60 Superdex 75 pg, 120 ml column, as described in section 3.11.2.*

Protein variant	**Bound to IEC column**	**Size-exclusion chromatography**
*Ma*LPMO10B	Weakly	Yes
*Ma*LPMO10B W82Y/N85F	Weakly	No
*Ma*LPMO10B Q141G	Yes	No
*Ma*LPMO10B W82Y	No	No
*Ma*LPMO10B N85F	No	No
*Ma*LPMO10B F88L	No	No
*Ma*LPMO10B F88Y	No	No
*Ma*LPMO10B Q141W	No	No
*Ma*LPMO10B D140A	No	Yes
*Ma*LPMO10B^{cd}	Yes	Yes
*Ma*LPMO10B^{cd} W82Y/N85F	Yes	Yes
*Ma*LPMO10D	No	Yes
*Ma*LPMO10D^{sl}	Weakly	Yes

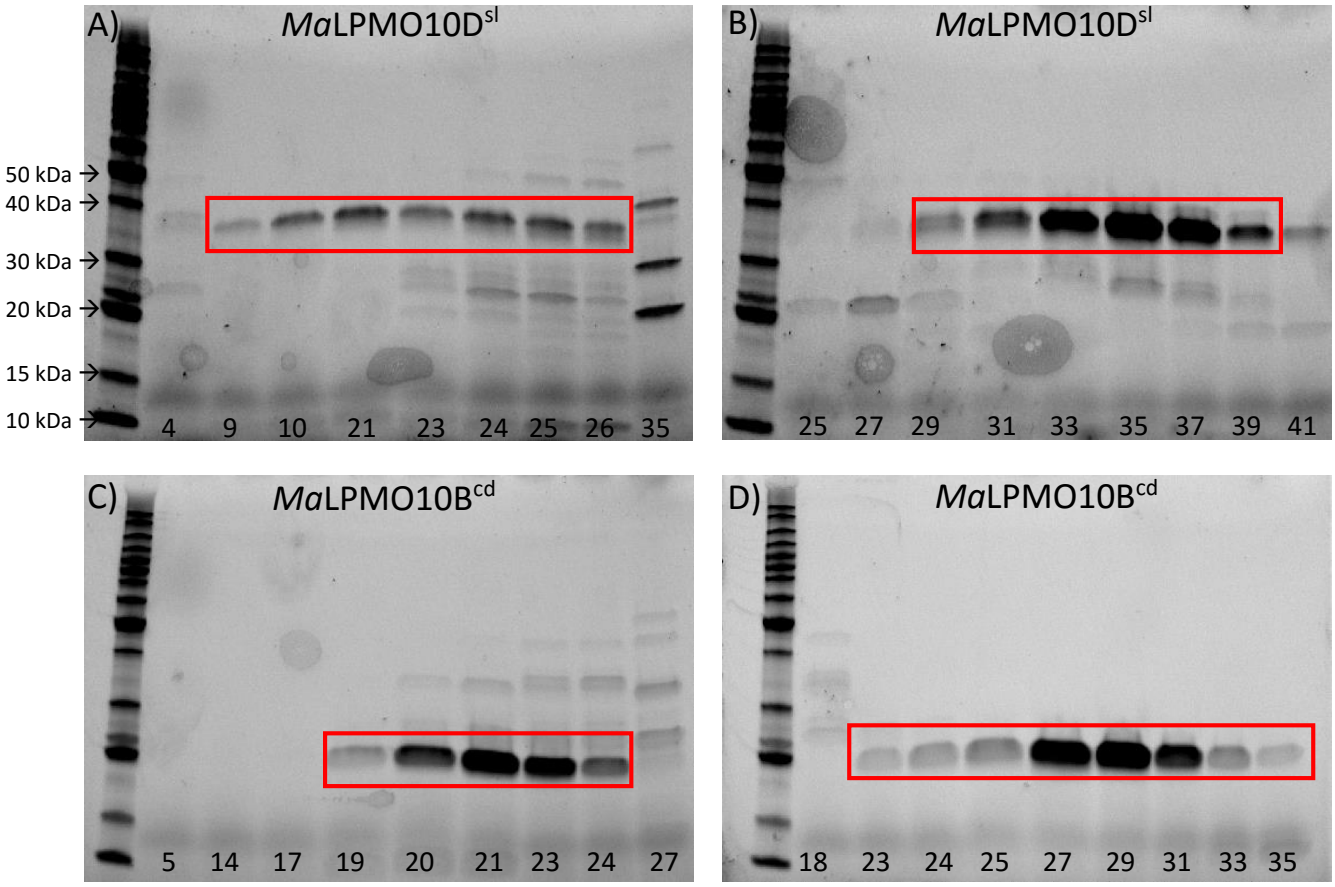


Figure 4.7. SDS-PAGE analysis of proteins purified by ion exchange chromatography (IEC) and size-exclusion chromatography (SEC). *The left lane of all four gels contains a protein ladder (Thermo Fisher) with masses labeled on the first gel for reference. All sample lanes are labeled with the corresponding fraction number collected during purification of the proteins (see Appendix D for chromatograms of the respective proteins). Proteins of interest are indicated by a red square. Panel A) shows MaLPMO10D$_{sl}$ after purification by IEC. The protein was weakly bound to the column, and eluted both as flow-through during application of binding buffer (fractions 9 and 10) and during application of a linear salt gradient at a concentration of approximately 45 mM NaCl (fractions 21 to 26). Fractions 8 to 26 were combined and purified by SEC as shown in panel B). The LPMO was not completely purified by SEC, as some smaller proteins are visible on the gel. Panel C) shows MaLPMO10B^{cd} after purification by IEC. The LPMO bound completely to the column, and was eluted by application of a linear salt gradient. Fractions 18 to 25 were combined for further purification by SEC, shown in panel D). Purification of MaLPMO10B^{cd} by SEC resulted in a highly purified LPMO. Predicted sizes of proteins: MaLPMO10B^{cd}, 21.6 kDa; MaLPMO10D^{sl}, 33.0 kDa.*

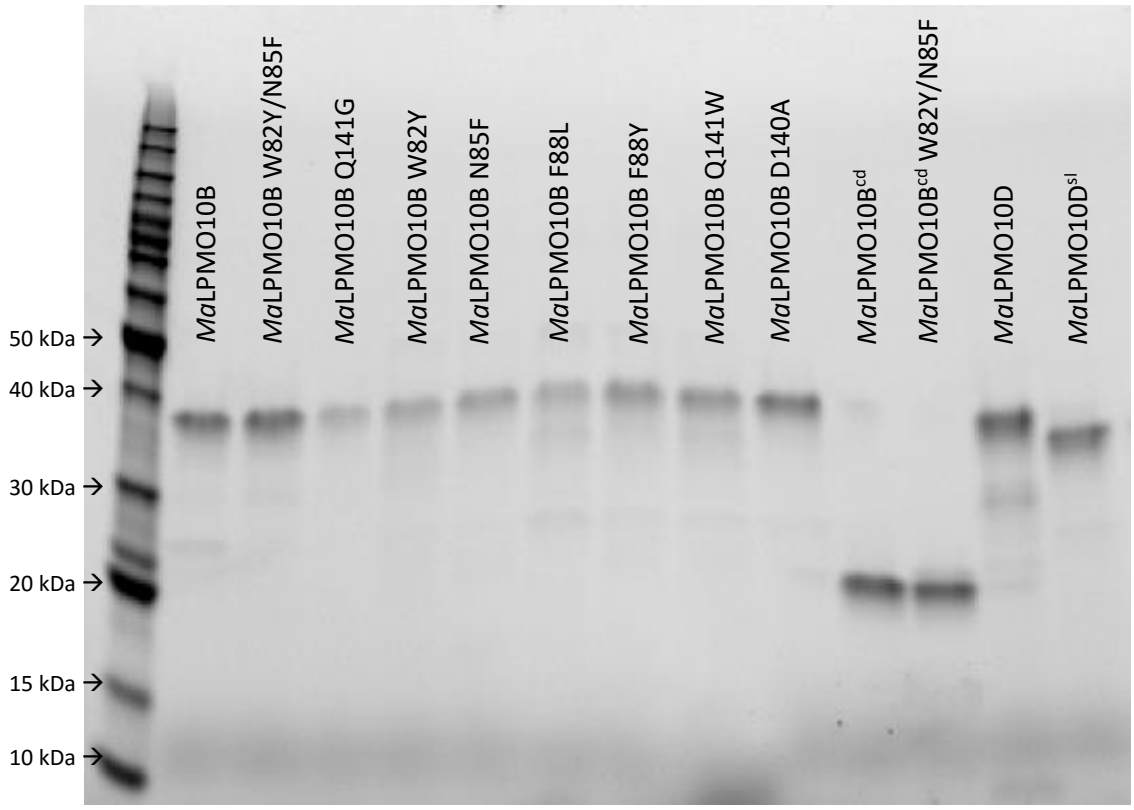


Figure 4.8. SDS-PAGE analysis of all purified LPMOs in this study. *The left lane contains a protein ladder with selected bands labeled by molecular weight. The other lanes contain all 14 LPMOs purified in this study. See table 4.3 for information about purification methods involving each respective protein. Predicted sizes of LPMOs: all MaLPMO10B variants, approximately 34.9 kDa; both MaLPMO10B^{cd} variants, 21.6 kDa; MaLPMO10D, 34.0 kDa; MaLPMO10D^{sl}, 33.0 kDa.*

4.4 Analysis of LPMO activity

4.4.1 Initial investigation of activity of wild type LPMOs towards various substrates

In order to determine the substrate preferences of *Ma*LPMO10B, *Ma*LPMO10Bcd and *Ma*LPMO10D^{sl} (since the experiment was conducted before successful cloning of full-length *Ma*LPMO10D, *Ma*LPMO10D^{sl} was used in its place), the enzymatic activity of the enzymes was determined qualitatively using shrimp shell α-chitin, squid pen β-chitin and microcrystalline cellulose (Avicel® PH-101). Analysis of reaction products showed that all three LPMOs were active on Avicel (Figure 4.9) and β-chitin (Figure 4.10). Analysis of the chromatograms indicated

that products oxidized at the C1 carbon (aldonic acids), C4 carbon (geminal diols), as well as double-oxidized (oxidized at both C1 and C4 carbons) products were formed from incubation with cellulose, whereas only aldonic acids were detected as products when the enzymes were incubated with β-chitin. The product profiles of oxidized products formed from β-chitin indicate that the various LPMOs have a tendency to form even-numbered oxidized products as previously shown (Vaaje-Kolstad et al., 2010). Product formation of truncated *Ma*LPMO10B was significantly lower than that of full-length *Ma*LPMO10B. The chromatograms also indicated that none of the enzymes were active on α-chitin, which was further verified by MALDI-TOF MS.

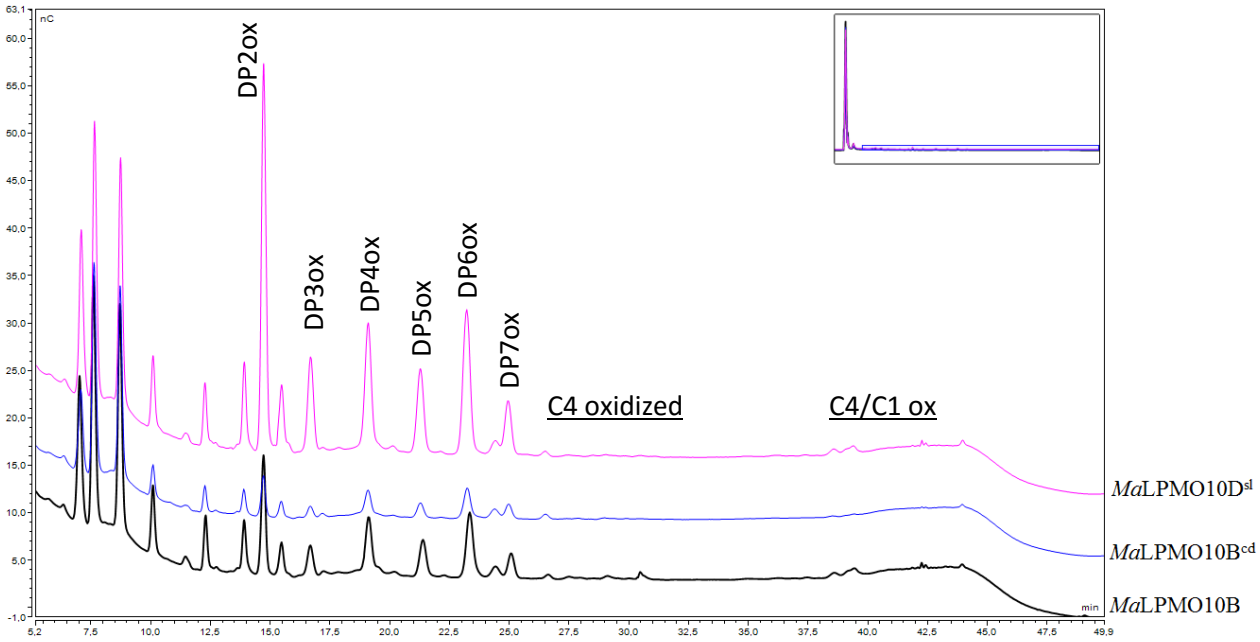


Figure 4.9. Product profiles of full-length and truncated MaLPMO10B and MaLPMO10D^{sl}. *Purified LPMOs at a concentration of 1 µM were mixed with 10 mg/ml substrate, 50 mM Bis-Tris buffer (pH 6.0) and 2 mM ascorbic acid. Samples were incubated overnight at 40 °C and 1,000 rpm, and analyzed by HPAEC-PAD. All reactions were carried out in triplicates (n=3). DP2ox-DP7ox represents oxidized cello-oligomers released form the microcrystalline cellulose by the various LPMOs.*

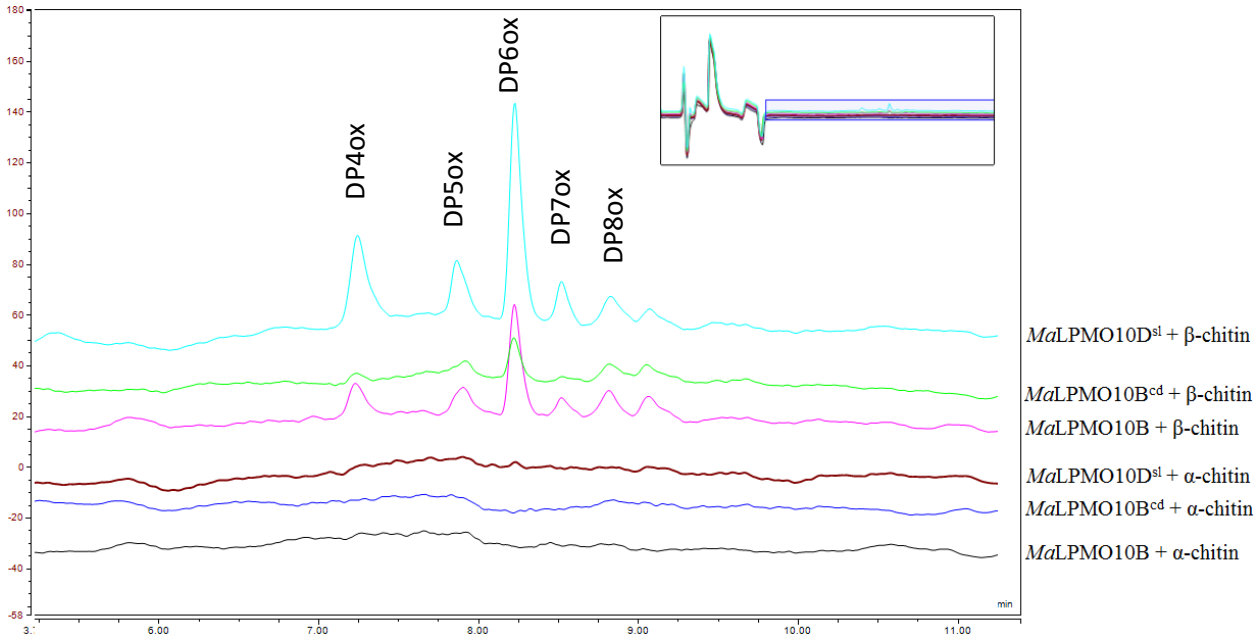


Figure 4.10. Product profiles of full-length and truncated MaLPMO10B and MaLPMO10D^{sl}. Purified LPMOs at a concentration of 1 µM were mixed with 10 mg/*ml substrate, 50 mM Bis-Tris buffer (pH 6.0) and 2 mM ascorbic acid. Samples were incubated overnight at 40 °C and 1,000 rpm, and analyzed by UPLC. All reactions were carried out in triplicates (n=3). DP4ox-DP8ox represents oxidized chito-oligomers released form microcrystalline cellulose by the various LPMOs.*

4.4.2 Time course of released oxidized products from Avicel by *Ma*LPMO10B and *Ma*LPMO10B^{cd}

A quantitative analysis was performed to investigate product formation by full-length and truncated (lacking the CBM2 domain) *Ma*LPMO10B over time when acting on microcrystalline cellulose (Avicel® PH-101). The analysis (Figure 4.11) showed that formation of cellobionic acid (GlcGlc1A) by full-length *Ma*LPMO10B were nearly 4-fold that of truncated *Ma*LPMO10B. Activity levels were similar the first two hours of incubation, before activity of the truncated LPMO ceased.

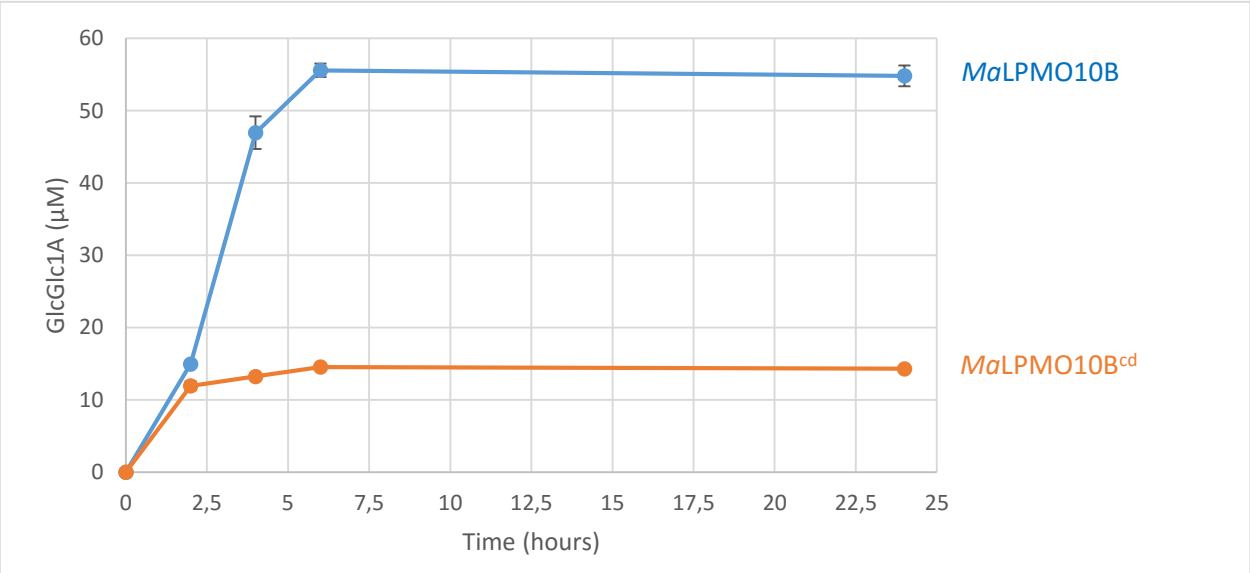


Figure 4.11. Time course of oxidized products released from Avicel by full-length and truncated MaLPMO10B. *The graph shows oxidation of 10 mg/ml Avicel by 1 µM MaLPMO10B (blue line and data points) and MaLPMO10B^{cd} (orange line and data points) in 50 mM Bis-Tris (pH 6.0) and 2 mM ascorbic acid incubated at 40 °C. Due to the product complexity, all samples were further degraded by TfCel5A (1 µM end concentration) to allow simplified quantification (see section 3.15 for further explanation). Data are mean values ± standard deviations (n=3); standard deviations indicated by error bars at each point in the graph.*

4.4.3 Time course of released oxidized products from Avicel by *Ma*LPMO10D and *Ma*LPMO10D^{sl}

In order to demonstrate the effect of a shortened linker region in *Ma*LPMO10D has on product formation, a quantitative analysis was performed on microcrystalline cellulose (Avicel® PH-101). Analysis of the formation of cellobionic acid (GlcGlc1A; Figure 4.12) over the course of 24 hours showed that *Ma*LPMO10D with a complete linker region oxidized the substrate at a higher rate the first hours of incubation; however, at the end of the experiment, both *Ma*LPMO10D variants had generated similar amounts of cellobionic acid.

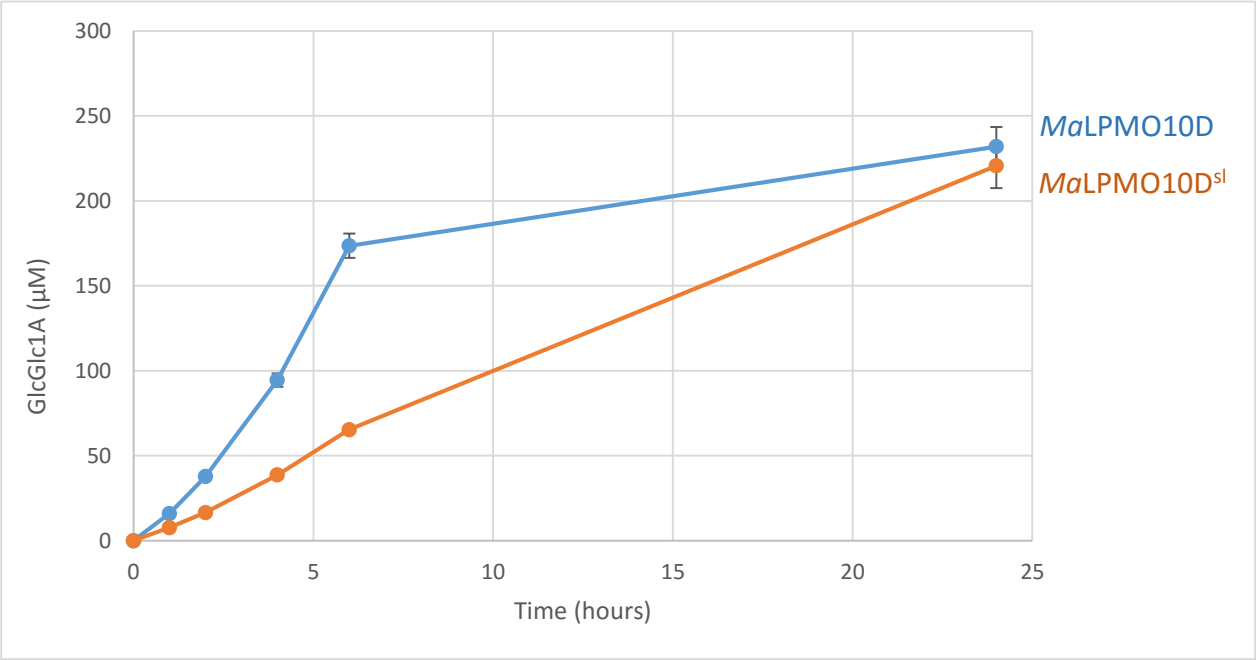


Figure 4.12. Time course of oxidized products released from Avicel by MaLPMO10D and MaLPMO10D^{sl}. *The graph shows oxidation of 10 mg/ml Avicel by 1 µM MaLPMO10D (blue line and data points) and MaLPMO10D^{sl} (orange line and data points) in 50 mM Bis-Tris (pH 6.0) and 2 mM ascorbic acid incubated at 40 °C. Due to the product complexity, all samples were further degraded by TfCel5A (1 µM end concentration) to allow simplified quantification (see section 3.15 for further explanation). Data are mean values ± standard deviations (n=3); standard deviations indicated by error bars at each point in the graph.*

Some of the reaction mixtures collected after 24 hours (prior to addition of cellulase) were used to perform a qualitative analysis of released oxidized cello-oligosaccharides. The product profiles of *Ma*LPMO10D and *Ma*LPMO10D^{sl} (Figure 4.13) demonstrates that both enzymes oxidize highly crystalline cellulose in a similar manner, generating the various cello-oligosaccharides in equal amounts.

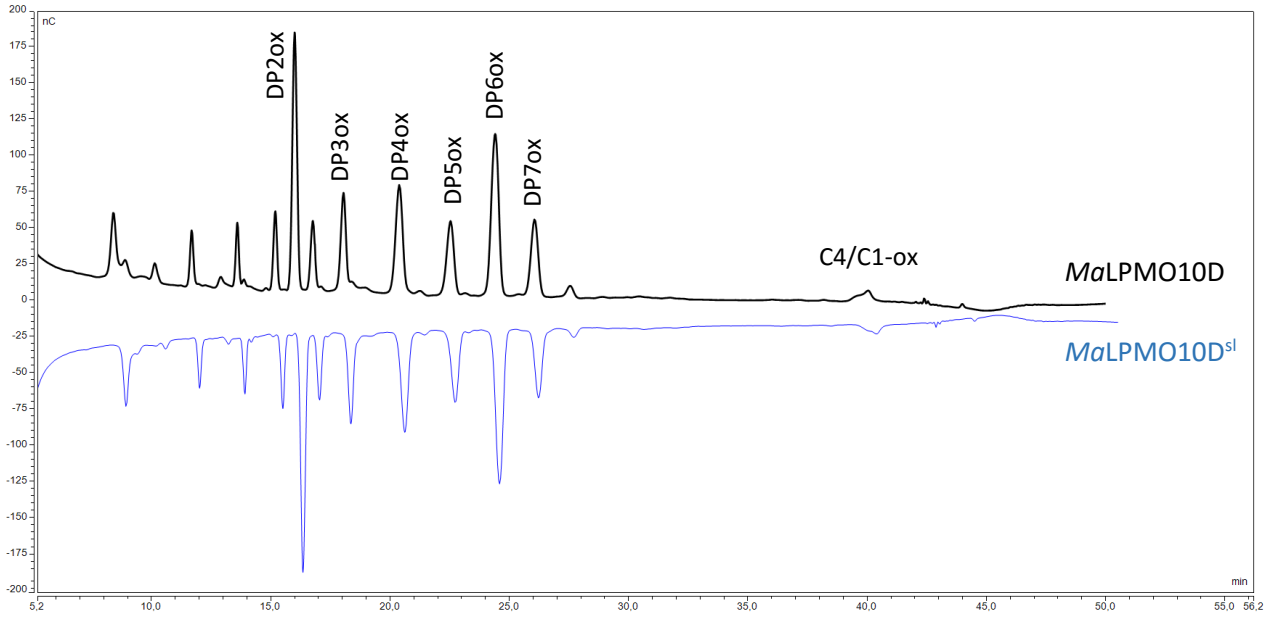


Figure 4.13. Chromatographic comparison of product profiles of MaLPMO10D and MaLPMO10D^{sl}. *The chromatogram shows the product profiles of both enzymes after 24 hours of incubation with Avicel. Reaction mixtures contain 1 µM LPMO, 10 mg/ml Avicel, 50 mM Bis-Tris (pH 6.0) buffer and 2 mM ascorbic acid. Soluble fractions were analyzed by HPAEC-PAD. DP2ox-DP7ox represents oxidized cello-oligomers released from microcrystalline cellulose by the MaLPMO10D variants.*

4.4.4 Activity of wild type and mutant *Ma*LPMO10B on β-chitin

In order to determine the activity of all *Ma*LPMO10B mutants on β-chitin, a qualitative analysis was performed to determine the enzymatic activity of the enzymes, using wild type *Ma*LPMO10B as reference (Figure 4.14). Analysis of reaction products showed that all LPMOs were active on β-chitin, and that only aldonic acids were detected as products. The chromatograms reveal that all *Ma*LPMO10B variants have a tendency to form even-numbered oxidized products.

The product analysis indicated that two *Ma*LPMO10B mutants, Q141G and D140A, generated more oxidized chito-oligosaccharides than the wild type. For that reason, a quantitative analysis was set up to compare the release of chitobionic acid (GlcNAcGlcNAc1A) from β-chitin by wild type *Ma*LPMO10B and three mutants; D140A, Q141G and Q141W. To allow quantification of chitobionic acid (GlcNAcGlcNAc1A), all oxidized products were further degraded by a GH20 β-*N*-acetylhexosaminidase (chitobiase CHB from *Serratia marcescens*). CHB was added to a final

concentration of 1.6 µM and incubated at 37 °C for 2 hours prior to HILIC analysis. Figure 4.15 shows the quantification of chitobionic acid in the various samples. The results demonstrate that wild type *Ma*LPMO10B is more active than the various *Ma*LPMO10B mutants on β-chitin.

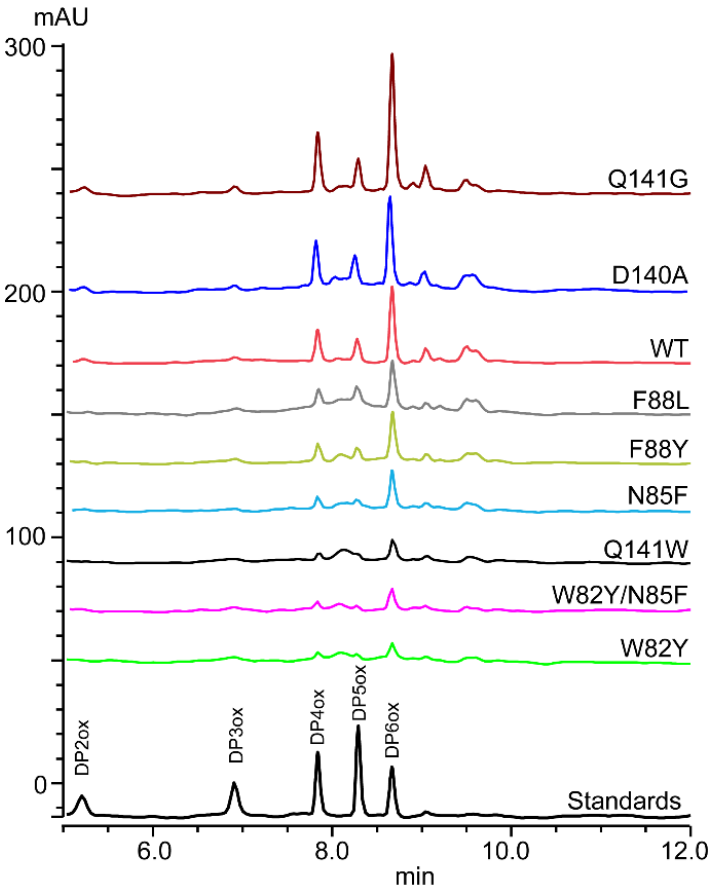


Figure 4.14. Qualitative analysis of activity of wild type and mutant MaLPMO10B towards β-chitin. *The chromatograms show released oxidized chito-oligosaccharides from β-chitin by wild type and mutant MaLPMO10B. Cu(II)-saturated LPMO (1 µM) were mixed with 10 mg/ml β-chitin, 50 mM Bis-Tris (pH 6.0) buffer and 2 mM ascorbic acid. Reactions were incubated overnight at 40 °C and with shaking at 1,000 rpm before soluble fractions were analyzed by HILIC. Standards of Chit-O oxidized chito-oligosaccharides (DP2-DP6) was used as a reference to allow peak annotation. DP2ox-DP6ox represents oxidized chito-oligomers released from β-chitin by the various LPMOs.*

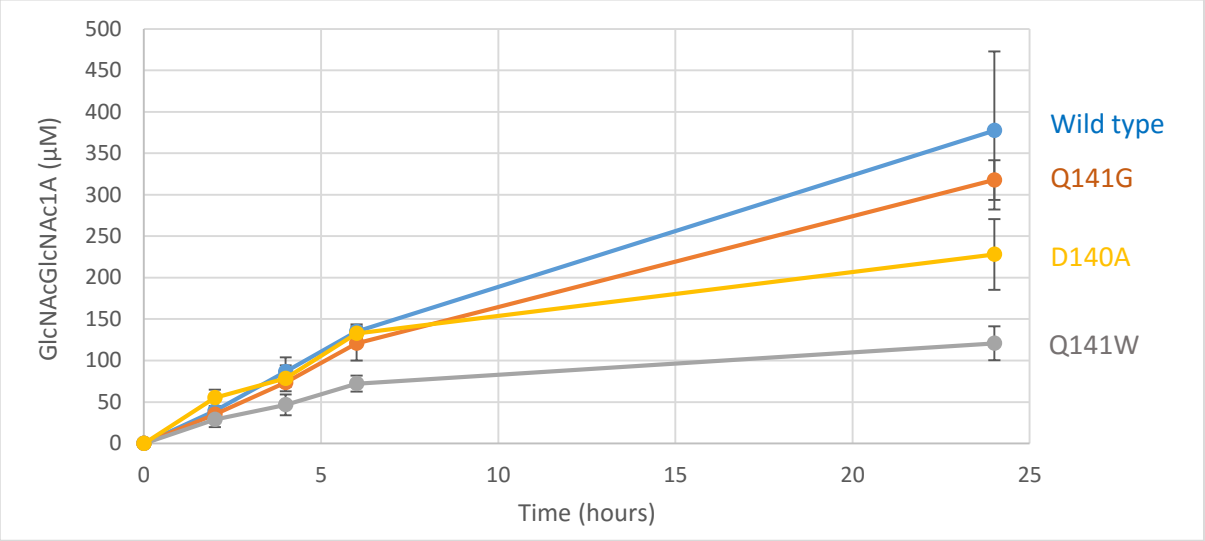


Figure 4.15. Quantitative analysis of released chitobionic acid from β-chitin by wild type and three mutants of MaLPMO10B. *Reactions of 1 µM Cu(II)-saturated LPMO degrading 10 mg/ml β-chitin in 50 mM Bis-Tris (pH 6.0) were pre-incubated for 30 minutes at 40 °C and with shaking at 1,000 rpm prior to addition of 2 mM ascorbic acid. Soluble fractions were collected after 2, 4, 6 and 24 hours of incubation subsequent to CHB (1.6 µM end concentration) degradation. Known concentrations of GlcNAcGlcNAc1A were used to generate a standard curve, thus allowing calculation of concentration of chitobionic acid based on peak areas. Data are mean values of three independent reactions and standard deviations are indicated by error bars at each point in the graph.*

4.4.5 Activity of wild type and mutant *Ma*LMO10B on Avicel and PASC

Qualitative analyses were conducted to compare the product profiles of *Ma*LPMO10B wild type and mutants when compared to a strict C1-oxidizing LPMO (LPMO10C from *Streptomyces coelicolor*; Forsberg et al., 2014b). Wild type *Ma*LPMO10B generates C1-, C4-, and double (C4/C1)-oxidized cello-oligosaccharides, and any change in activity of mutants will be visualized on a HPAEC-generated chromatogram.

Analysis of released oxidized products from Avicel demonstrated that all *Ma*LPMO10B variants were active on microcrystalline cellulose (Figure 4.16), generating oxidized oligomers in the range of DP2ox to DP8ox. The product profile of wild type *Ma*LPMO10B includes the addition of C4- and C4/C1-oxidized products, in contrast to the product profile of *Sc*LPMO10C, which contains exclusively C1-oxidized products. No clear C4-oxidized products appeared to have been

generated by the *Ma*LPMO10B mutants; however, the area of C4-oxidized products formed by the wild type *Ma*LPMO10B is very low when using Avicel as substrate. Although, the presence of C4/C1-oxidized products (Figure 4.16B) demonstrates that most mutants had retained the ability to oxidize the C4 carbon. The enlarged region where C4/C1-oxidized products elutes in the chromatogram indicates that mutants W82Y, N85F, F88L and W82Y/N85F have little or no C4-oxidizing activity.

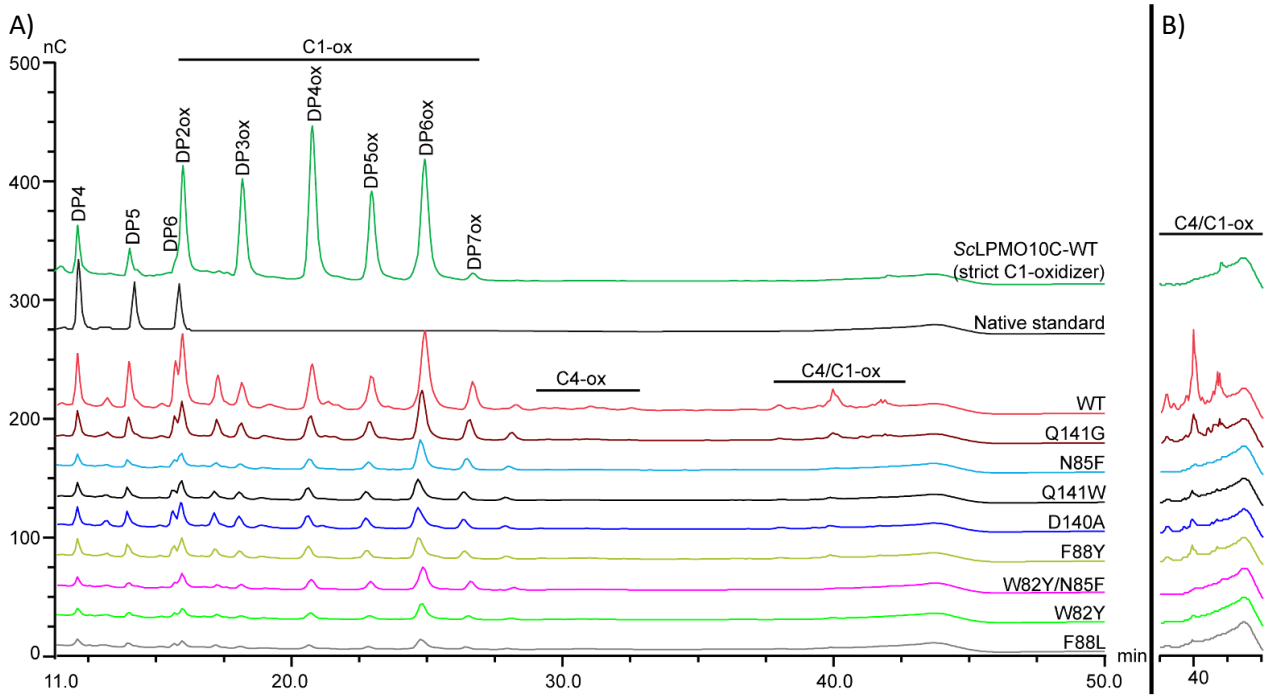


Figure 4.16. Qualitative analysis of oxidized products generated by wild type and mutant MaLPMO10B when acting on Avicel. *Each reaction contained 1 µM LPMO, 10 mg/ml Avicel, 50 mM Bis-Tris (pH 6.0) buffer and 2 mM ascorbic acid. Reactions were incubated overnight at 40 °C and 1,000 rpm followed by analysis by HPAEC-PAD. A reaction mixture containing ScLPMO10C, a strict C1-oxidizing LPMO, was used as for comparison due to the absence of C4- and C4/C1-oxidized products. A native standard, containing cello-oligosaccharides of DP4-DP6 was also included to facilitate peak annotation. The full chromatogram (Panel A) includes labels of the regions containing C1-, C4- and C4/C1-oxidized products, respectively. Panel B) is an enlargement of the region containing C4/C1-oxidized products. DP4-DP6 represents native products, whereas DP2ox-DP7ox represents oxidized cello-oligomers.*

In order to easier detect C4 and C4/C1 oxidized products generated by the various *Ma*LPMO10B mutants, another qualitative analysis was conducted using phosphoric acid swollen cellulose (PASC) as a substrate (Figure 4.17). PASC is made from Avicel® PH-101, and phosphoric acid pretreatment results in substrate swelling and thus more amorphous Avicel, and thereby increased accessibility of the cellulose (Hall et al., 2010). Degradation studies using PASC as a substrate

detect increased amounts of oxidized products compared to using Avicel as a substrate. This is the case when comparing the chromatograms in figure 4.16 to the chromatograms in figure 4.17; i.e. larger amounts of all products are observed when using PASC as a substrate instead of Avicel. Among the *Ma*LPMO10B mutants, all but W82Y/N85F produced C4-oxidized products, and by looking at C4/C1-double oxidized products, only small amounts were detected for W82Y/N85F. Small amounts of C4- and C4/C1-oxidized products were detected for the W82Y and N85F mutants.

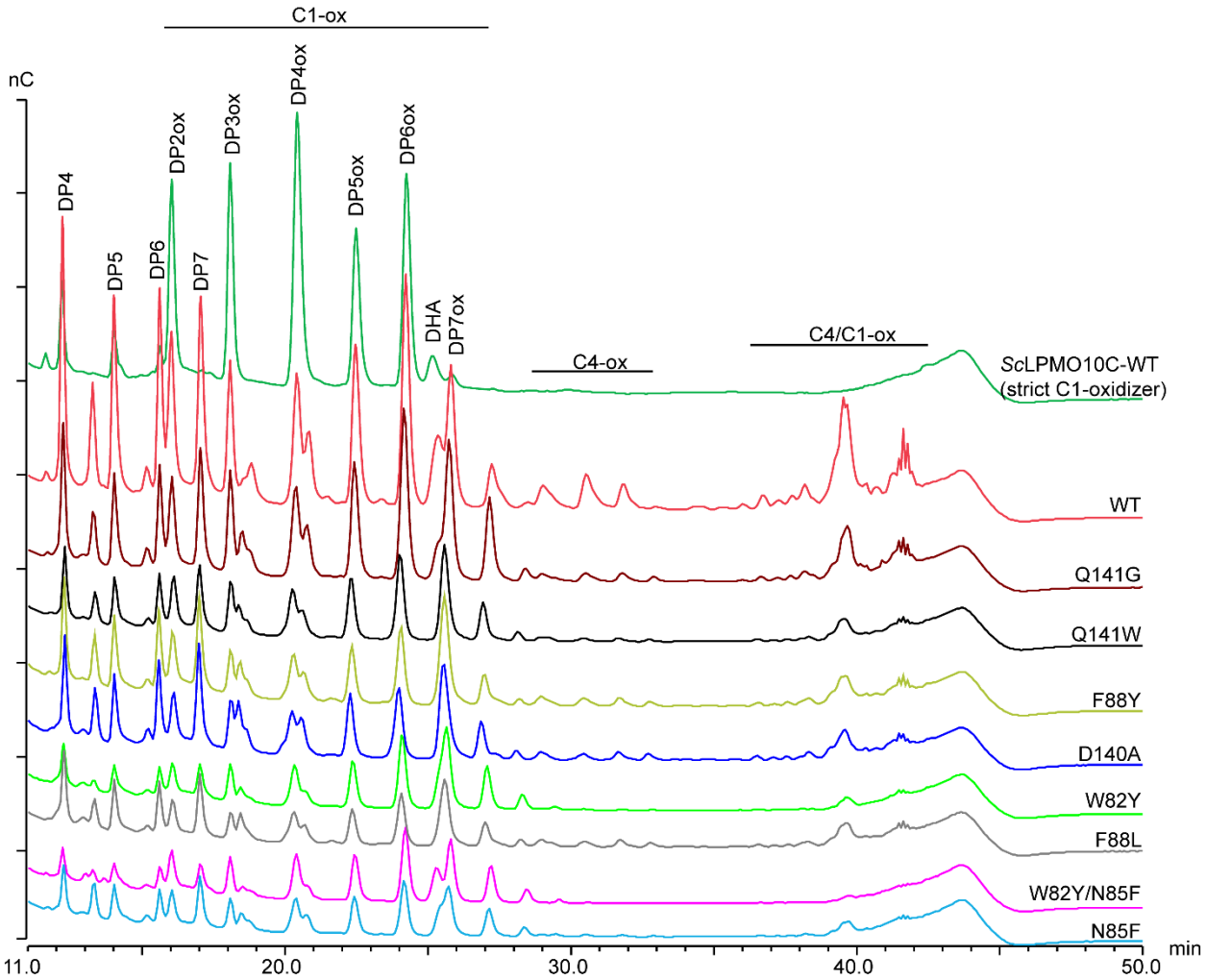


Figure 4.17. Qualitative analysis of oxidized products generated by wild type and mutant MaLPMO10B on PASC. *In each reaction, 1 µM LPMO was mixed with 1 mg/ml Avicel, 50 mM Bis-Tris (pH 6.0) buffer and 2 mM ascorbic acid. Reactions were incubated overnight at 40 °C and 1,000 rpm followed by analysis by HPAEC-PAD. A reaction mixture containing ScLPMO10C, a strict C1-oxidizing LPMO, was used as a standard due to the absence of C4- and C4/C1-oxidized products. The chromatogram includes labels of the regions containing C1-, C4- and C4/C1-oxidized products, respectively. DP4-DP7 represents native products, whereas DP2ox-DP7ox represents oxidized oligomers. DHA, dehydroascorbate.*

The overnight incubated reactions containing PASC were also analyzed using MALDI-TOF MS. The analysis demonstrated that no C4 or C4/C1 oxidizing products had been generated by *Ma*LPMO10B W82Y/N85F (Figure 4.18; larger version is included in Appendix E). Comparison of the spectra of wild type and mutant (i.e. W82Y/N85F) *Ma*LPMO10B and *Sc*LPMO10C demonstrate a change in the C4-oxidizing activity of the W82Y/N85F mutant, which generate the same products as the C1-oxidizing *Sc*LPMO10C. No C4/C1-oxidized products are visible in the W82Y/N85F specter, in contrast to the wild type enzyme.

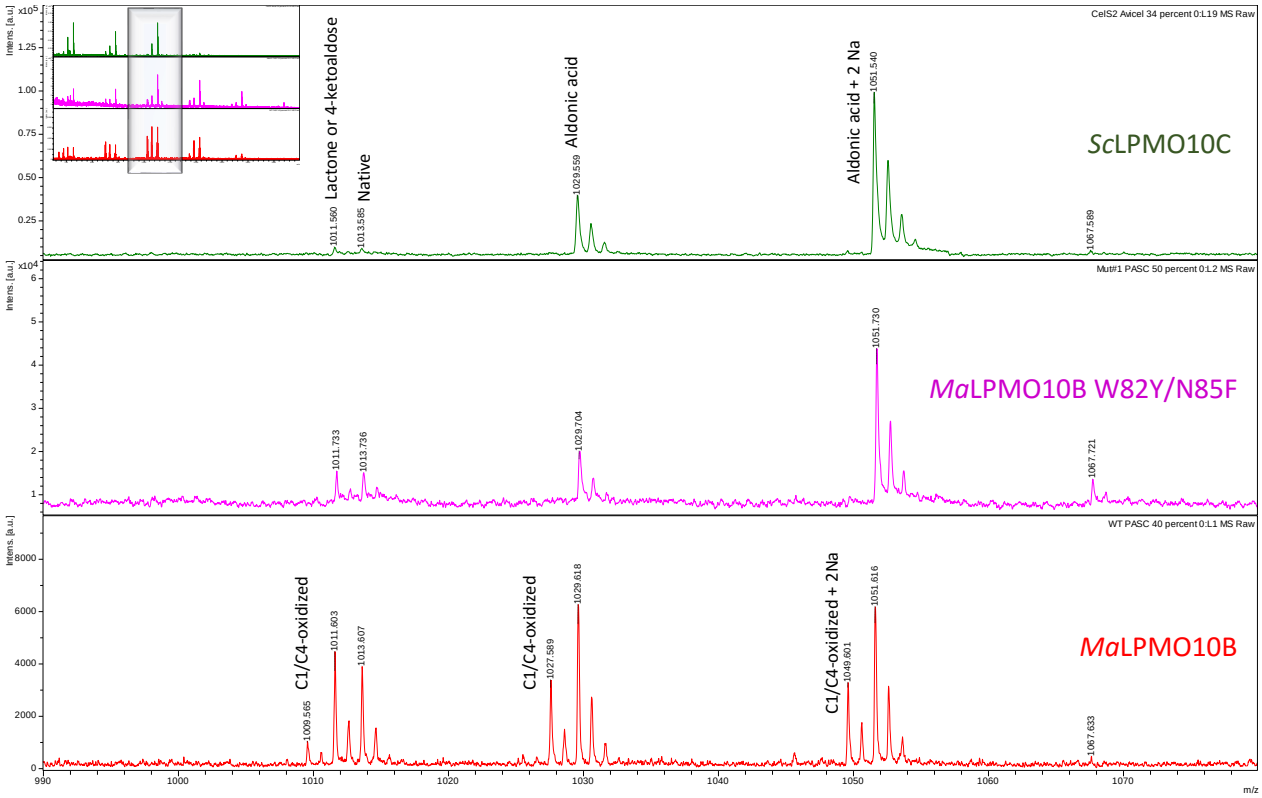


Figure 4.18. MALDI-TOF MS analysis of released oxidized cello-oligosaccharides from PASC by ScLPMO10C, MaLPMO10B and MaLPMO10B W82Y/N85F*. Overnight reactions were mixed with DHB in 1:2 ratios and air-dried prior to analysis by MALDI-TOF MS. The figure shows an enlargement of the DP6 ion cluster (highlighted in the top left corner) of each LPMO. Note the absence of C4/C1-oxidized products (labeled) in the MaLPMO10B W82Y/N85F specter. A larger version of the figure is included in Appendix F.*

4.4.6 Detailed examination of the C4 oxidizing activity of *Ma*LPMO10B W82Y/N85F

The apparent loss of ability of the *Ma*LPMO10B W82Y/N85F mutant to oxidize the C4 carbon of cellulose was quite remarkable and required a deeper analysis of the activity of this enzyme for verification of the initial results obtained. Analysis of C4/C1 oxidized products generated by *Ma*LPMO10B W82Y/N85F, was therefore performed by further degradation of the oxidized products using additional cellulases and a cellobiose dehydrogenase (*Mt*CDH; regenerates oxidized species allowing easier detection of C4/C1 oxidized products). Control reactions including *Sc*LPMO10C and *Ma*LPMO10B were used for comparison of chromatograms

generated by HPAEC-PAD. The chromatograms (Figure 4.19) demonstrated that no C4 or C4/C1-double oxidized products had been generated from Avicel by *Ma*LPMO10B W82Y/N85F, as the product profile was similar to that of the strict C1 oxidizing *Sc*LPMO10C.

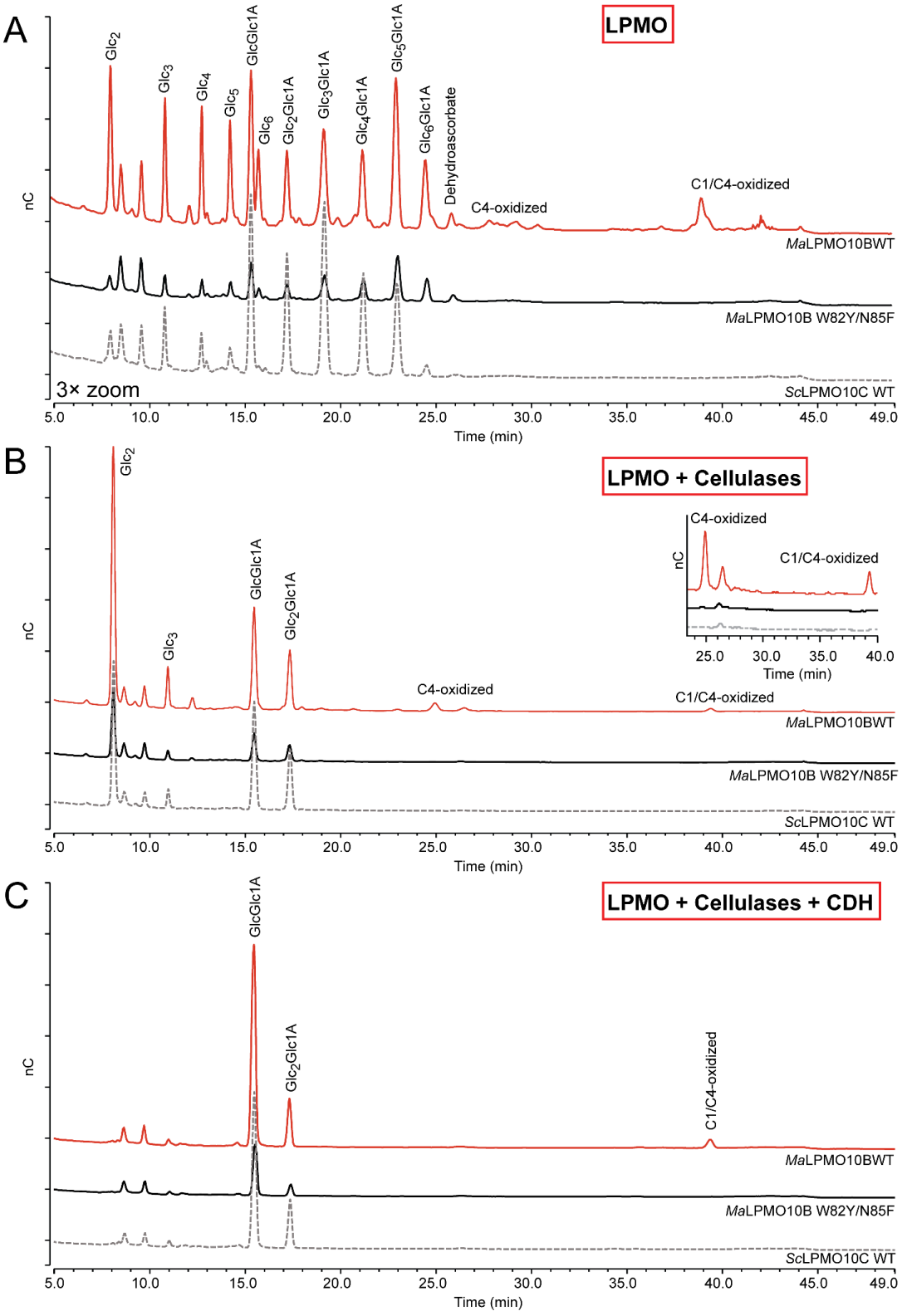


Figure 4.19. Analysis of C4/C1-oxidized products generated by MaLPMO10B W82Y/N85F acting on Avicel. *Reactions of ScLPMO10C (strict C1 oxidizer) or wild type MaLPMO10B were used for comparison. Each reaction contained 1 µM LPMO, 10 mg/ml Avicel, 50 mM Bis-Tris (pH 6.0) buffer and*

2 mM ascorbic acid (Panel A). Reactions were incubated overnight at 40 °C and 1,000 rpm, before spilt into two. To one part, 1 µM TfCel5A and 1 µM TrCel7A was added (Panel B). After incubation overnight (37 °C), samples were again split into two; one part was mixed with 1 µM MtCDH followed by incubation overnight at 37 °C (Panel C). Samples were analyzed by HPAEC-PAD.

Further examination of the C4 oxidizing activity of *Ma*LPMO10B W82Y/N85F was done using PASC as a substrate. A chromatographic analysis was performed, including cellulases to hydrolyze oxidized products into shorter fragments (DP2ox and DP3ox), thereby simplifying analysis. All *Ma*LPMO10B mutants, as well as wild type *Ma*LPMO10B and *Sc*LPMO10C, were included, to examine the distribution of DP2-, DP3- and C4-oxidized products in the reactions. No standard samples of known concentration were included, meaning all data are based on mean peak area (n=3), thus only giving a relative comparison of activity. Figure 4.20 shows the distribution of C1- and C4-oxidized (main peak of unknown DP, as shown in Figure 4.21) products in the various samples. Chromatograms of W82Y/N85F and wild type *Ma*LPMO10B and *Sc*LPMO10C are shown in figure 4.21. Both figures demonstrate the presence of small amounts of C4-oxidized products generated by W82Y/N85F. As the figures illustrate, the C1- and C4-oxidizing activity of W82Y/N85F is lower than that of the wild type enzyme; however, the relative distribution of oxidized products indicate that the overall activity level of the mutant is not lowered in equal amounts; the C4-oxidizing activity is highly reduced compared to the C1-oxidizing activity.

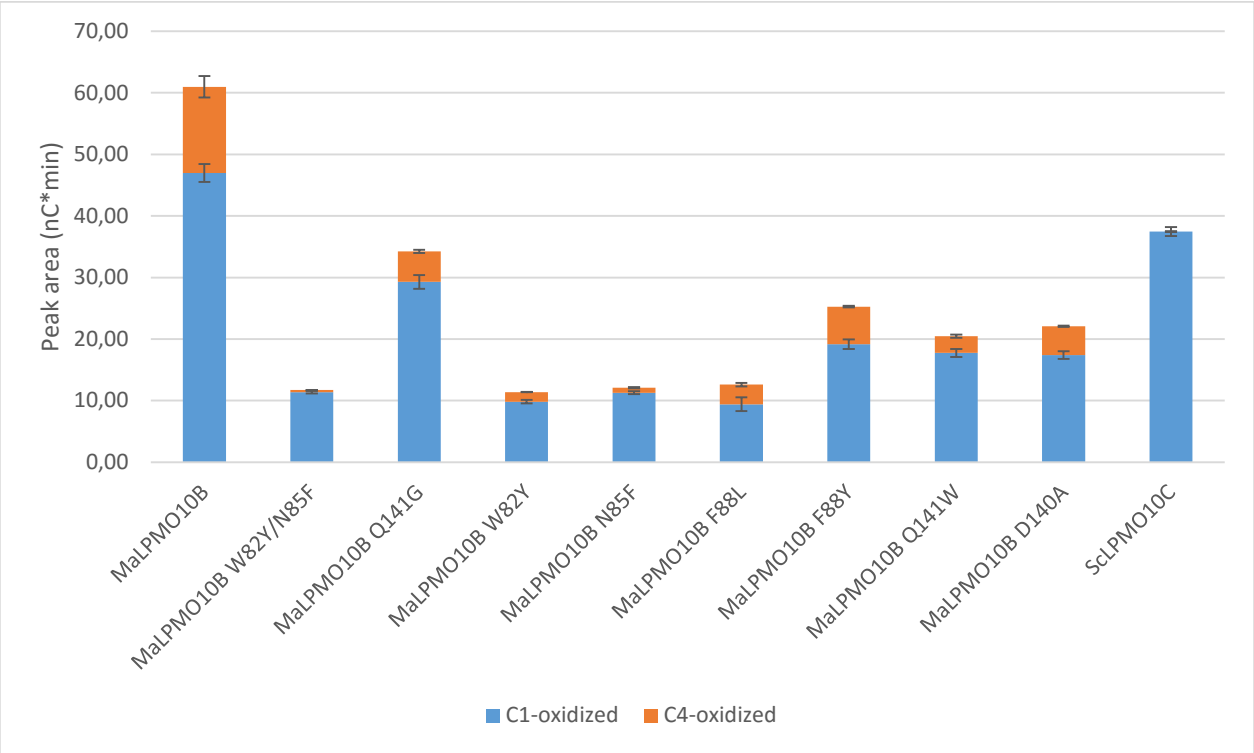


Figure 4.20. Distribution of oxidized products among wild type and mutant MaLPMO10B and ScLPMO10C. *Cu(II)-saturated LPMO, at a concentration of 1 µM, was mixed with 1 mg/ml PASC in 50 µM Bis-Tris (pH 6.0) buffer in the presence of 2 mM ascorbic acid. Reactions were incubated overnight at 40 °C and 1,000 rpm, followed by addition of 1 µM TfCel5A and 1 µM TrCel7A to the soluble fractions. After incubation overnight at 37 °C, samples were analyzed using HPAEC-PAD. Data are based on mean peak area of three independent reactions. Error bars indicate standard deviations.*

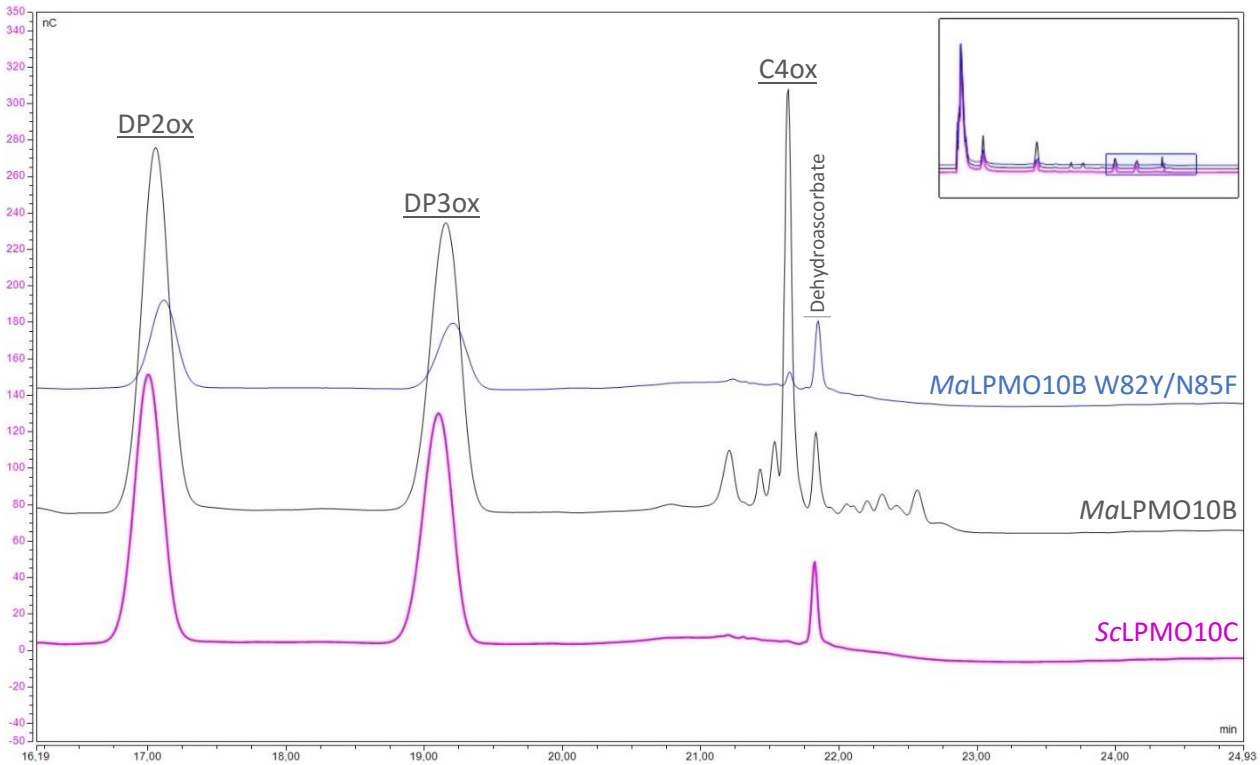


Figure 4.21. Chromatographic analysis of oxidized products released from PASC by MaLPMO10B W82Y/N85F and wild type MaLPMO10B and ScLPMO10C. *The figure shows an enlargement of the DP2-, DP3- and C4-oxidizing regions of the chromatograms. Notice the presence of C4-oxidizing products generated by W82Y/N85F. The chromatograms demonstrate the relative amounts of oxidized products generated by the various LPMOs, and indicates that the C4-oxidizing activity of W82Y/N85F is highly reduced compared to that of the wild type enzyme.*

An additional approach used to determine the C4-oxidizing activity of *Ma*LPMO10B W82Y/N85F was accomplished by concentrating the reaction mixture sample prior to analysis by HPAEC-PAD. Non-concentrated reactions containing oxidized products generated by *Ma*LPMO10B and *Sc*LPMO10C were used for comparison. The soluble fraction of *Ma*LPMO10B W82Y/N85F was vacuum dried (Concentrator plus, Eppendorf), resulting in a 3x concentration of the sample. The chromatogram (Figure 4.22) of W82Y/N85F reveals several small peaks in the C4-oxidized products region. The chromatograms also illustrate the big difference in C1- and C4-oxidized products generated by *Ma*LPMO10B and *Ma*LPMO10B W82Y/N85F; however, conclusions cannot be drawn in regard to quantities, since the qualitative analysis only allow relative comparison of the product profiles.

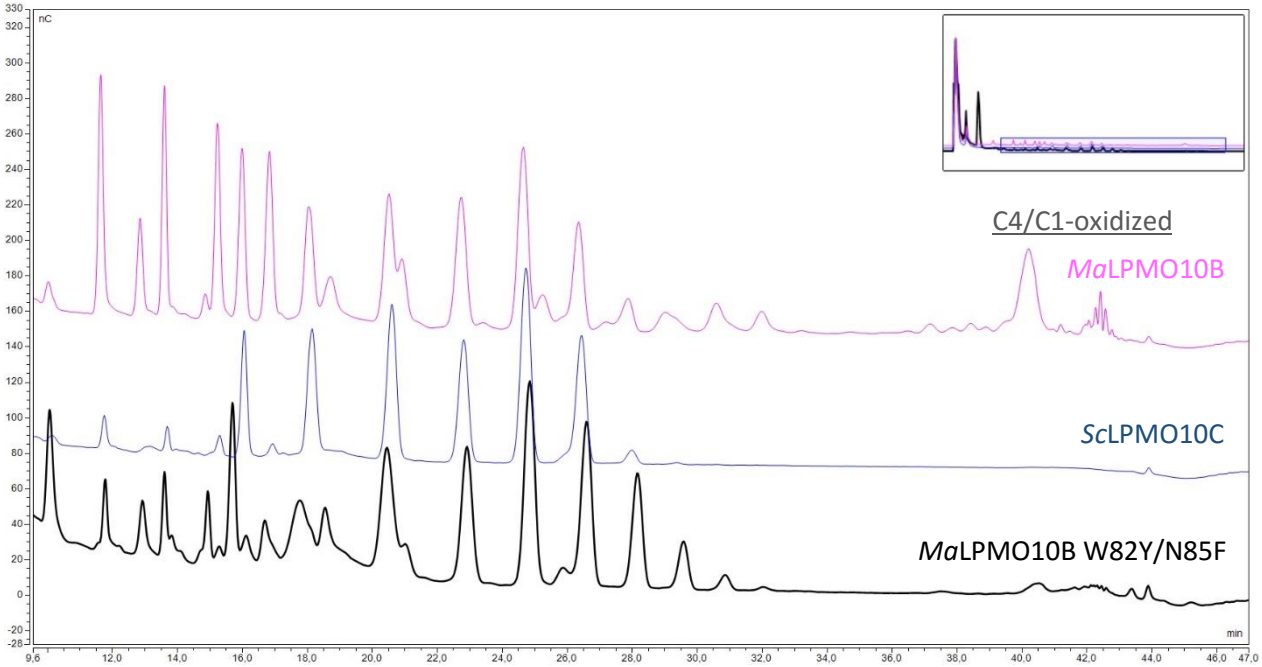


Figure 4.22. Comparison of concentrated oxidized products of MaLPMO10B W82Y/N85F to non-concentrated oxidized products of MaLPMO10B and ScLPMO10C using HPAEC-PAD. *The soluble fraction of MaLPMO10B W82Y/N85F was vacuum dried and rehydrated using dH_2O, resulting in a 3x concentration of the sample. Soluble fractions of degradation reactions of MaLPMO10B and ScLPMO10C, including the concentrated W82Y/N85F sample, were analyzed using HPAEC-PAD. The region of interest (i.e. C4/C1-double oxidized products) is labeled. Note the presence of small peaks of C4/C1-double oxidized products in the chromatogram of MaLPMO10B W82Y/N85F.*

4.4.7 Synergy experiment

Since LPMOs with different activities may target different regions of the substrate an experiment was set up to investigate possible synergistic effects obtained by combining such activities (the C1/C4 oxidizing LPMO (the C1/C1 oxidizing *Ma*LPMO10B, C1 oxidizing *Sc*LPMO10C and the *Ma*LPMO10B W82Y/N85F mutant). The data (Figure 4.23) indicate no clear synergistic effects at the end point of analysis (24 hours). On the other hand, the combination of *Ma*LPMO10B and *Ma*LPMO10B W82Y/N85F seemed to show a synergy in the degradation of PASC in the early phase of the reaction. The synergistic effect is apparent until the third hour of incubation, where the concentration of released cellobionic acid is similar. After one hour of incubation, both the

wild type and the mutant *Ma*LPMO10B have generated equal amounts of cellobionic acid, but the C1-oxidizing activity of the mutant ceases during the next hour of incubation, where no more cellobionic acid are generated by the mutant. This is also the case for the reaction containing only *Ma*LPMO10B W82Y/N85F; after one hour of incubation the activity ceases.

The graph in figure 4.24 presents the amount of released C4-oxidized products as peak area. Although being non-quantitative, the graph still illustrates relative amounts of oxidized products among the different conditions. A synergistic effect is observed after the first two hours of incubation with a combination of both wild type and mutant *Ma*LPMO10B; however, the C4-oxidizing activity ceases within the next hour of incubation. Incubation of *Ma*LPMO10B W82Y/N85F alone, or in combination with *Sc*LPMO10C, generated no C4-oxidized products. Interestingly, the reaction containing a combination of *Ma*LPMO10B and *Sc*LPMO10C were inactive during the first three hours of incubation, before the C4-oxidizing activity was observed.

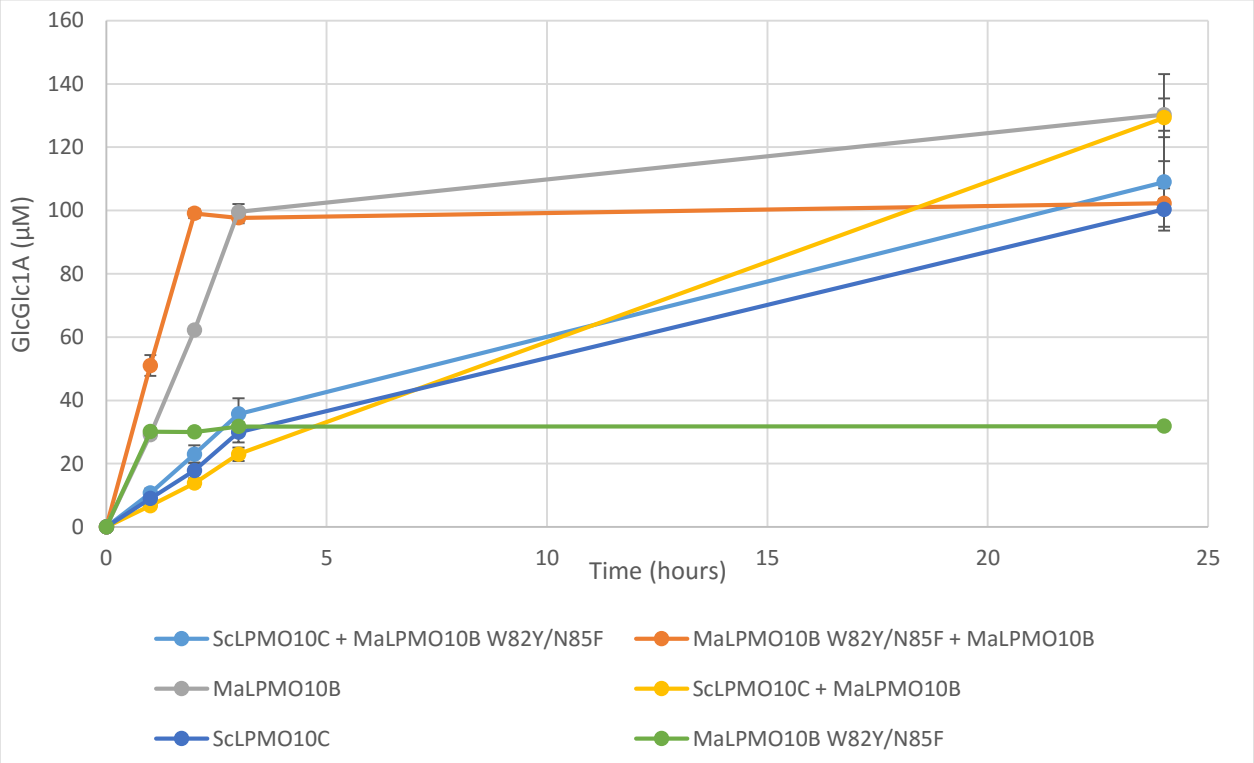


Figure 4.23. Analysis of possible synergism of combinations of MaLPMO10B, MaLPMO10B W82Y/N85F and ScLPMO10C. *Time course of released cellobionic acid (GlcGlc1A) following incubation of 1 µM LPMO (see table 3.9) and 1 mg/ml PASC in 50 mM Bis-Tris (pH 6.0). Reactions were*

pre-incubated for 20 minutes before 2 mM ascorbic acid was added. Samples were collected after 1, 2, 3 and 24 hours of incubation at 40 °C and 1,000 rpm. Soluble fractions were mixed with 1 µM TfCel5A and incubated overnight at 37 °C prior to analysis by HPAEC-PAD. Standard reactions of known concentration were used to generate a standard curve, allowing quantification of cellobionic acid (GlcGlc1A) in each reaction. Data are mean values ± standard deviations (n=3); standard deviations are indicated by error bars at each point in the graph.

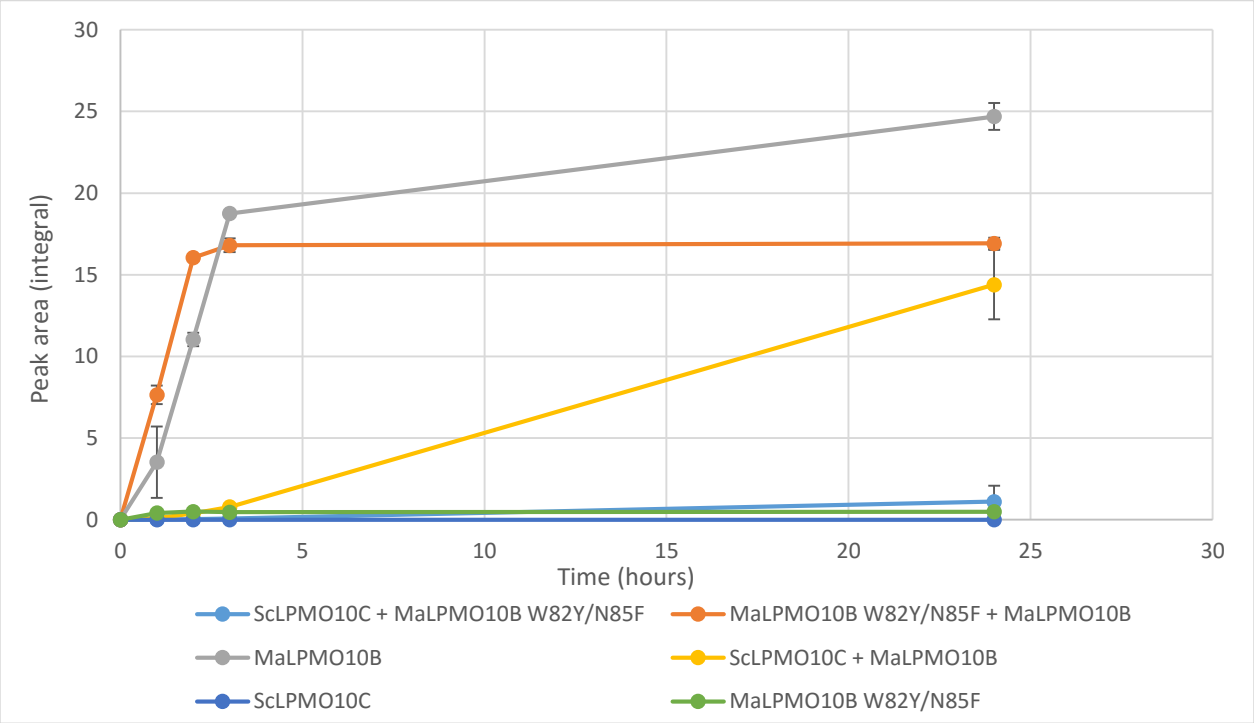


Figure 4.24. Analysis of possible synergism of combinations of MaLPMO10B, MaLPMO10B W82Y/N85F and ScLPMO10C. *Time course of released C4-oxidized products presented as mean peak area (non-quantitative) of three independent reactions. Data are mean values ± standard deviations (n=3); standard deviations are indicated by error bars at each point in the graph.*

4.4.8 Influence of CBM on *Ma*LPMO10B^{cd} W82Y/N85F activity

The presence and absence of the *Ma*LPMO10B CBM may influence the ability of the W82Y/N85F mutant to produce C4 oxidized products. In order to investigate this, the activity of truncated *Ma*LPMO10B W82Y/N85F was compared to that of the full-length *Ma*LPMO10B W82Y/N85F (Figure 4.25). Moreover, due to protein crystallization of *Ma*LPMO10B^{cd}

W82Y/N85F, it was of interest to verify that the truncated mutant was active. The results indicate that the activity of the truncated mutant was reduced compared to that of the full-length mutant.

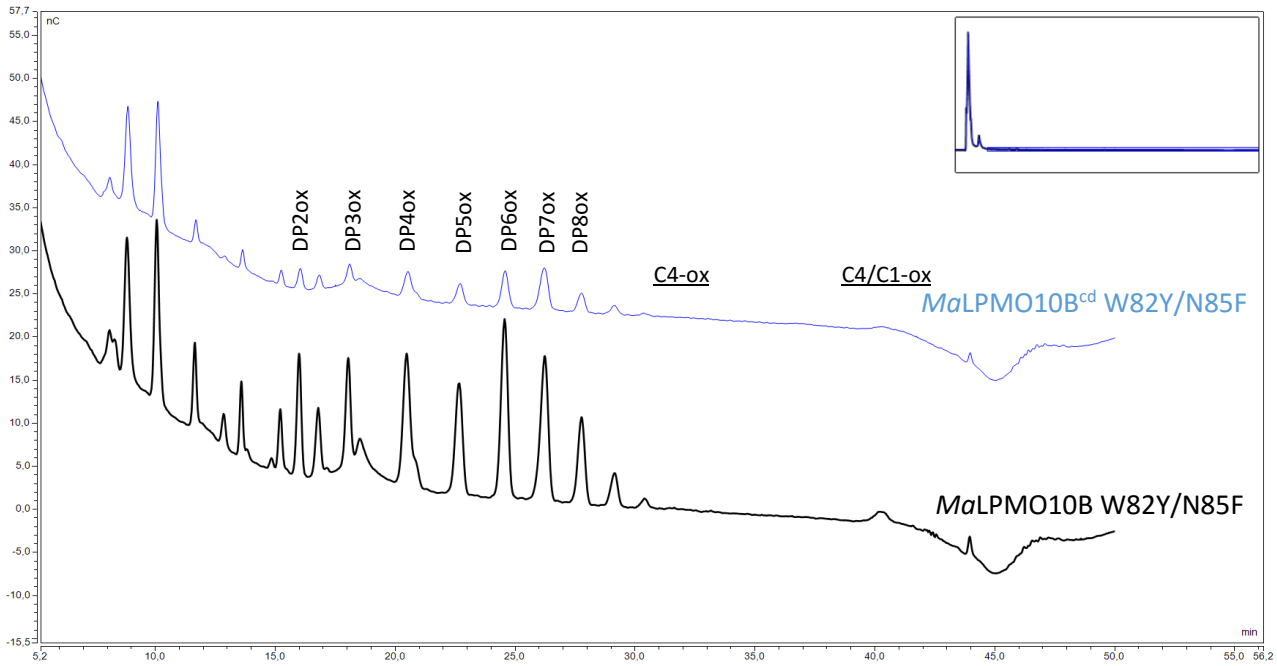


Figure 4.25. Qualitative analysis of the activity of MaLPMO10B^{cd} W82Y/N85F using HPAEC-PAD. *In each reaction, 1 µM Cu(II)-saturated LPMO was mixed with 10 mg/ml Avicel® PH-101 in 50 mM Bis-Tris (pH 6.0) in the presence of 2 mM ascorbic acid. The reactions were incubated overnight at 40 °C and 1,000 rpm before the soluble fractions were analyzed using HPAEC-PAD. Comparison of the product profiles in the figure demonstrates that the activity of truncated MaLPMO10B W82Y/N85F is reduced.*

4.5 Binding assays

4.5.1 Binding of full-length and truncated *Ma*LPMO10B to Avicel and β-chitin

In order to demonstrate the binding affinity of *Ma*LPMO10B towards Avicel and β-chitin, an experiment was set up to measure the amount of unbound protein in soluble fractions of reaction mixtures.

Figure 4.26 shows the binding affinity of full-length and truncated *Ma*LPMO10B to Avicel and β-chitin, presented as percentage of free protein in the soluble fractions as measured by A_{280} over time. The results demonstrate that full-length *Ma*LPMO10B bound equally strong to Avicel and β-chitin, although binding occurred much more quickly to Avicel. Within five minutes of incubation, approximately 80 % of the protein was bound to Avicel, whereas binding of the same

amount of protein to β-chitin took approximately four hours. Truncated *Ma*LPMO10B bound weakly to Avicel and β-chitin, with no more than 20 % protein bound to the substrates at the end of the incubation period. The rate of binding was similar to the full-length protein.

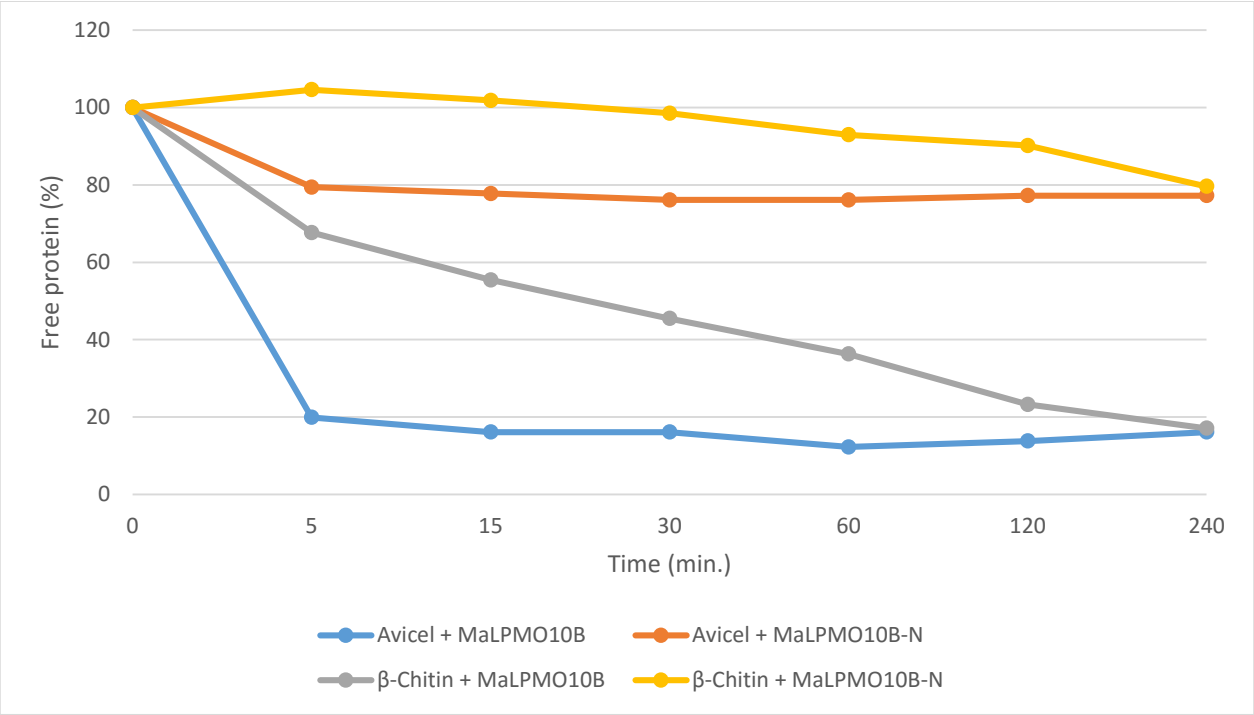


Figure 4.26. Binding of full-length and truncated MaLPMO10B to Avicel and β-chitin. *Reactions of 0.08 mg/ml LPMO and 10 mg/ml substrate in 50 mM Bis-Tris (pH 6.0) buffer was incubated at 22 °C and 1,000 rpm followed by filtration of samples collected after 5, 15, 30, 60, 120 and 240 minutes. The amount of unbound protein in the soluble fractions was measured by A_{280} and presented as the percentage of free protein in the samples over time. Truncated MaLPMO10B is designated as MaLPMO10B-N (for N-terminal domain). Data are based on one reaction (n=1).*

4.5.2 Binding of *Ma*LPMO10D and *Ma*LPMO10D^{sl} to Avicel

In order to investigate the effect of the shortened linker region of *Ma*LPMO10D^{sl} on binding affinity towards Avicel, an experiment was set up to measure the amount of unbound protein in soluble fractions of reaction mixtures. The data (Figure 4.27) demonstrate that the two proteins bound equally strong to Avicel, although *Ma*LPMO10D^{sl} bound more quickly to the substrate.

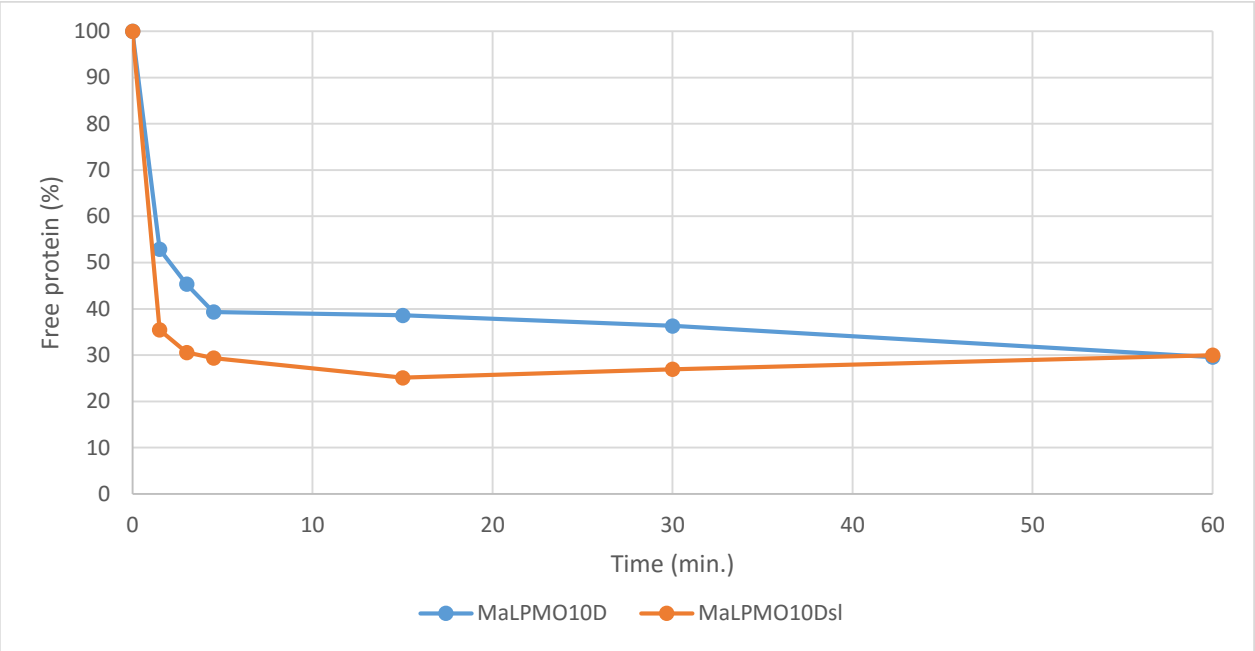


Figure 4.27. Binding of MaLPMO10D and MaLPMO10D^{sl} to Avicel. *Reactions of 0.08 mg/ml LPMO and 10 mg/ml Avicel in 50 mM Bis-Tris (pH 6.0) buffer were incubated at 22 °C and 1,000 rpm before samples were collected and filtrated at various time points (1.5, 3, 4.5, 15, 30 and 60 minutes). The amount of unbound protein in the soluble fractions was measured by A_{280} and presented as the percentage of free protein in the samples over time. Data are based on one reaction (n=1).*

4.5.3 pH-dependent binding of *Ma*LPMO10B to Avicel

In order to investigate the pH-dependency of *Ma*LPMO10B binding to Avicel when mixed with buffers of a wide range of pH values, an experiment was set up including four different buffers with pH values ranging from 3.5 to 10.5. The data (Figure 4.28) demonstrate that *Ma*LPMO10B bound to the microcrystalline cellulose with high affinity in all the various buffers and pH values. More than 80 % of the protein bound to the substrate in all conditions.

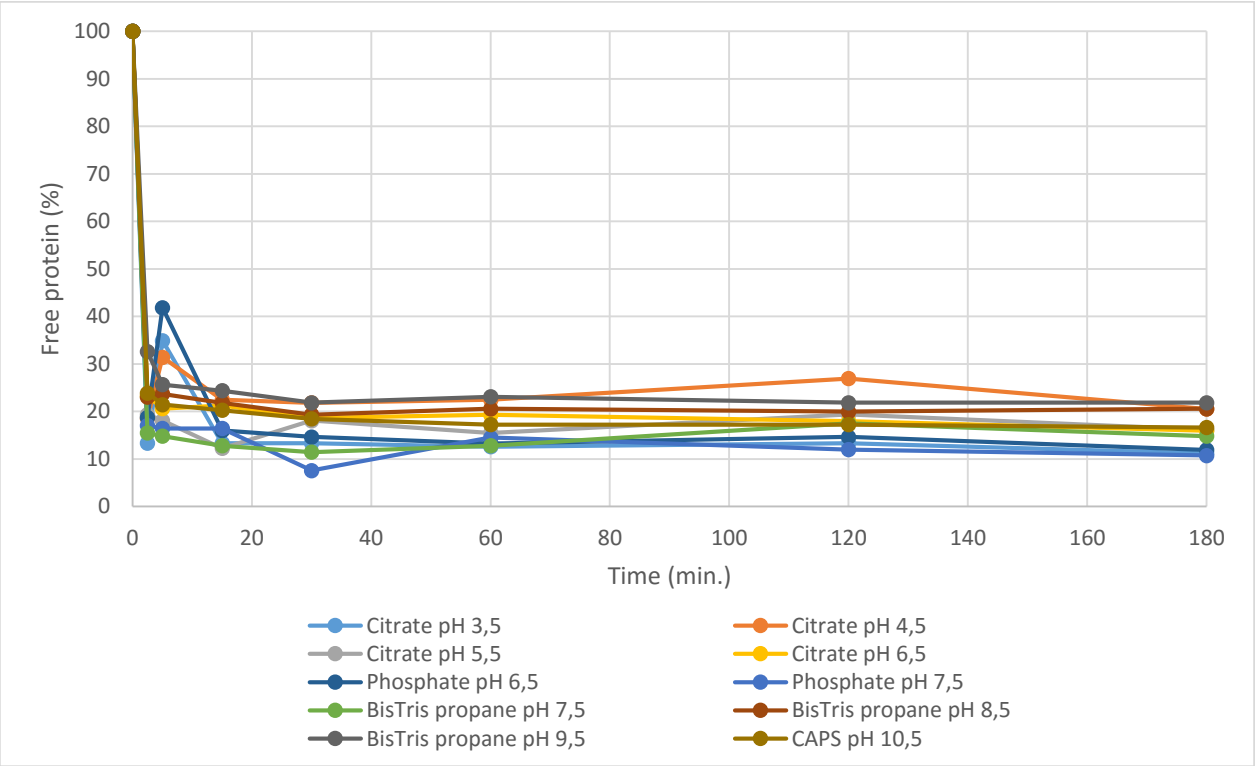


Figure 4.28. pH-dependent binding of MaLPMO10B to Avicel. *Reactions of 0.08 mg/ml LPMO and 10 mg/ml Avicel in 50 mM buffer were incubated at 22 °C and 1,000 rpm before samples were collected and filtrated at various time points (2.5, 5, 15, 30, 60, 120 and 180 minutes). The amount of unbound protein in the soluble fractions was measured by A_{280} and presented as the percentage of free protein in the samples over time. Data are based on one reaction (n=1).*

5 Discussion

The objective of this study was to gain insight into the function and mechanism of cellulose-active AA10 LPMOs and to and to better understand the structural basis of the oxidative mechanism (i.e. regioselectivity; C1 or C1/C4 oxidation) of these proteins. To investigate the mechanisms of oxidative regioselectivity, two putative cellulose-active AA10 proteins, both assumed to have C1/C4 oxidizing activity, were selected from the bacterium *Micromonospora aurantiaca* ATCC 27029 for characterization.

Cloning and expression of *Ma*LPMO10B and *Ma*LPMO10D, both full-length and truncated versions (i.e. lacking the CBM domain), were attempted. *Ma*LPMO10B variants were successfully cloned and expressed, but cloning of *Ma*LPMO10D variants initially failed. Sequencing results of full-length *Ma*LPMO10D indicated that a stretch of nucleotides was missing from the linker region, but interestingly, both the catalytic and the CBM domain remained intact, without a change in the reading frame. The protein, termed *Ma*LPMO10D^{sl}, was successfully expressed. It was suspected that a high GC-content (60.05 %) in the codon optimized gene, or a possible formation of secondary structures in the DNA template caused amplification (by Phusion) of the complete gene to fail. The addition of DMSO and an alternative reaction buffer (Phusion GC buffer) to the PCRs, as well as a raised annealing temperature, led to successful amplification and subsequent cloning and expression of full length *Ma*LPMO10D. Cloning of a truncated version of *Ma*LPMO10D was unsuccessful, for reasons not further investigated.

Purification of truncated LPMOs was successful, and resulted in highly pure proteins; however, purification of full-length proteins was challenging. Due to low affinity of the full-length LPMO to the ion exchange resin, most of the protein was eluted during washing of the column with binding buffer. Fortunately, the LPMO containing flow-through fractions were almost pure, since most *E. coli* proteins bound to the column. One reason for the low binding affinity could be that the predicted isoelectric point (pI) values were wrong (possibly to low), and that the pH value of the binding buffers therefore was too low. The predicted pI of the full-length proteins ranged from 6.63 to 7.78, which means that the proteins should carry a net negative charge and therefore bind to the positively charged resin when a binding buffer of pH 9.5 was used. Further investigation of the protein's physiochemical parameters showed that predicted pI of the CBM

domain alone was as high as 8.82. This is possibly too close to the pH of the binding buffer and a buffer with higher pH was therefore used. However, the increased pH did not result in binding of the LPMO to the ion exchange resin. In retrospect, the pI of the protein should have been investigated physically by isoelectric focusing, but since pure protein was obtained in spite of the protein not binding to the column, this was not pursued in order to save time. It should be noted that purification of a full-length LPMO was also attempted with hydrophobic interaction chromatography, but the protein did not interact with any of the column materials tested.

Initially, purification of full-length LPMOs by size-exclusion chromatography was successful and resulted in pure proteins; however, after replacement of the size-exclusion column, purification of full-length proteins failed. The reason for this is uncertain, since the two columns basically are of the same type (the new model replaced the old model; product data appear to be similar). It is plausible that the CBM bound to strongly to the spherical beads in the gel matrix, and as a result the LPMO did not elute. Indeed, no eluted fractions contained LPMO (i.e. full-length LPMO), and at the same time, the column pressure increased after each successive run, probably as a result of proteins stuck on top of the resin or in the gel matrix. A possible solution would be to use another type of SEC column. As a consequence of the issues regarding SEC purification, some of the proteins (i.e. several mutants) were only purified by IEC, and were thus not completely pure (Figure 4.8).

The use of impure LPMO samples may have had an effect on the outcome of activity assays in this study. On the one hand, the presence of other proteins or biomolecules in a protein sample may have led to incorrect determination of LPMO concentration. This would imply an underestimation of LPMO concentration, which may have had an impact on quantitative activity assays. The presence of contaminating proteins in the various enzyme reactions may have resulted in incorrect quantitative analysis of product formation in the samples if the contaminants had hydrolytic or oxidative activities. However, *E. coli* BL21 expression strains do not contain enzymes with such activities verified by control reactions (Z. Forsberg, personal communication), which means that product profiles (qualitative analyses) of the enzyme reactions would not have been affected in any significant way.

Initial activity assays demonstrated that *Ma*LPMO10B and *Ma*LPMO10D were active on Avicel and β-chitin, but not on α-chitin (Figure 4.9 and 4.10). As expected, both proteins oxidized

crystalline cellulose (i.e. Avicel) at the C1 and the C4 carbon, generating both C4 and C4/C1 oxidized products. Moreover, the importance of the CBM2 domain for binding of *Ma*LPMO10B to Avicel and β-chitin was demonstrated (Figure 4.26), and the results indicated stronger binding affinity towards Avicel than β-chitin. According to previous studies on LPMOs (Forsberg et al., 2014b), loss of binding to the substrate may reduce the activity of the enzyme. Indeed, a quantitative activity assay confirmed the importance of the CBM2 domain for *Ma*LPMO10B activity (Figure 4.11); the full-length protein yielded substantially higher amounts of cellobionic acid than the truncated protein. Interestingly, the activity levels of both proteins were equal the first two hours of incubation, before the activity of the truncated protein ceased. This result suggests that the lack of the CBM causes the catalytic domain to easily dissociate from the substrate, and thus ceasing product formation.

Two variants of *Ma*LPMO10D was expressed and purified in this study: one with a complete linker region (wild type), and one with a shortened linker region (10 missing amino acids). The effects of a shortened linker region in multi-domain LPMOs have to our knowledge not previously been studied. Our results demonstrate that the shortened linker region of *Ma*LPMO10D increased binding affinity of the enzyme towards crystalline cellulose (Figure 4.27). However, as the results are only based on one observation, further research on the binding affinities of the proteins are necessary. Quantitative analysis of product formation (i.e. cellobionic acid) from the two *Ma*LPMO10D variants demonstrated that the shortened linker region clearly decreased the rate of Avicel oxidation, although the amount of released products evened out with time (≤24 hours; Figure 4.12). A possible explanation is that a shortened linker results in the two domains to be in close proximity to one another, which may reduce substrate accessibility of the catalytic domain. Alternatively, the short linker region reduces the flexibility of the linker, which may inhibit the mobility of the catalytic domain on the substrate. CBMs have been shown to reduce the processivity of CAZymes (S. J. Horn et al., 2006; Sorensen et al., 2015), and as the CBM bind to crystalline cellulose much more strongly than the catalytic LPMO domain, the reduced mobility of the catalytic domain may have resulted in a decreased oxidation rate.

Binding of full-length *Ma*LPMO10B to Avicel was tested using several different buffers with a wide range of pH values (pH 3.5-10.5). Since the results are based on only one observation, the results must be carefully interpreted. However, the results suggest that the LPMO bound to the

substrate with high affinity (80-90 % of the protein bound within 2.5 minutes) in all buffers used (Figure 4.28), thus indicating that binding is independent on pH. It is thus possible that the interaction of aromatic amino acids of the protein dominate the interactions with the polysaccharide substrate.

As previously mentioned, the main objective of this thesis was to investigate the function and mechanism of cellulose-active AA10 proteins with the aim to better understand the structural basis for oxidative regioselectivity in these proteins. A phylogenetic analysis divides cellulose-active AA10 proteins into two groups: C1 specific or C1/C4 oxidizing LPMOs (Book et al., 2014). The fact that all AA10 proteins perform C1 oxidation, whereas some proteins also perform C4 oxidation may indicate that at some time in natural evolution some of the enzymes either gained or lost the ability to oxidize the C4 carbon of the crystalline cellulose (B. Bisarro, personal communication, unpublished work). The mechanisms behind oxidative regioselectivity in cellulose-active AA10 proteins can be studied in two ways: either introduce C4 oxidizing ability into a strict C1 oxidizer, or eliminate the C4 oxidizing ability from a C1/C4 oxidizer. In this study, selected residues in a C1/C4 oxidizer were targeted for site-directed mutagenesis, in order to observe any changes in the enzyme's oxidizing activity. A change in the rate of C4 oxidation compared to that of the wild type enzyme may indicate that the mutated residue was important for C4 oxidation.

The prelude to this thesis was the discovery of C4 oxidation of cellulose by the two AA10 proteins *Sc*LPMO10B and *Tf*LPMO10A by Forsberg *et al.* in 2014. Since mutagenesis studies targeting active site or substrate-binding surface residues may lead to reduced activity levels, it was decided to work with an enzyme that contained a CBM domain (*Sc*LPMO10B and *Tf*LPMO10A lacked a CBM) since LPMOs that have a CBM are more active, and detecting the activity of mutants may therefore be easier and more accurate (B. Bisarro, personal communication, unpublished work).

Due to the difficulties with cloning of *Ma*LPMO10D, and the fact that we were unable to obtain a truncated version of this protein, only *Ma*LPMO10B was selected for mutagenesis studies. Furthermore, to determine the three-dimensional structure, it was necessary to work with a protein that could be truncated (crystallization of the full-length protein is not preferable due to the flexible linker region). To identify residues responsible for carrying out the C4 oxidative

activity, site-directed mutagenesis of selected residues especially conserved within C1 or C1/C4 oxidizing AA10 proteins was performed. All residues selected for mutation were chosen on the basis of structure-based multiple sequence alignments (e.g. Figure 4.2), although the specific method for residue selection varied (some residues were selected with the aid of the phylogenetic and mutational analysis performed by B. Bisarro; see below). Since the structure of *Ma*LPMO10B was not determined before the end of this study, the structure of *Sc*LPMO10B [PDB 4OY6; Forsberg et al. (2014b)], a close homolog of *Ma*LPMO10B, was used for structural comparison to C1 oxidizing AA10 proteins.

Judged by structure-based MSAs, the position of phenylalanine 88 in *Ma*LPMO10B is replaced by a leucine in many C1 oxidizers (Figure 4.2). The phenylalanine is positioned close to the substrate-binding surface and in close proximity to the copper active site (approximately 5.8 Å from the copper ion). Interestingly, this position is often a tyrosine in other C1/C4 oxidizers, and the phenylalanine was thus mutated to both a leucine and a tyrosine residue. The mutants were successfully produced, and analysis of activity on both cellulosic substrates (i.e. Avicel and PASC) and β-chitin showed that the activity of *Ma*LPMO10B F88L and F88Y was reduced on all substrates. There were no signs of reduced C4 oxidizing activity compared to C1 oxidized activity, and the product profiles of both mutants were similar to that of the wild type on all substrates. The reduced activity levels suggest that the mutations have caused changes in the conformation at or near the active site, or in the substrate-binding surface of the LPMO. Moreover, the phenylalanine does not appear to play a significant role in the oxidative regioselectivity of *Ma*LPMO10B.

AA10 proteins contain a highly conserved alanine residue (Ala142 in *Ma*LPMO10B) located in a loop close to the active site. The loop hosting the conserved alanine holds different configurations in C1 and C1/C4 oxidizers, and it has been suggested that a highly conserved aspartic acid (position 140 in *Ma*LPMO10B) in C1/C4 oxidizers is responsible for forcing the loop into a different conformation than that of C1 oxidizers. The different loop conformations cause the highly conserved alanine to be differently positioned in C1 and C1/C4 oxidizers, which allows adequate space for a ligand to enter in the solvent-facing axial position in C1/C4 oxidizers. In C1 oxidizers, the alanine is more closely positioned to the active site, which may result in restricted axial access to the copper ion [Figure 5.1; (Forsberg et al., 2014b)]. With an attempt to change

the conformation of the loop, Asp140 in *Ma*LPMO10B was mutated to an alanine. The mutant was successfully produced and analysis of the activity demonstrated that the mutant was active on both cellulose and β-chitin. As with mutants F88L and F88Y, the overall activity level of D140A was reduced on all substrates; however, the activity level of D140A on β-chitin was close to that of the wild type.

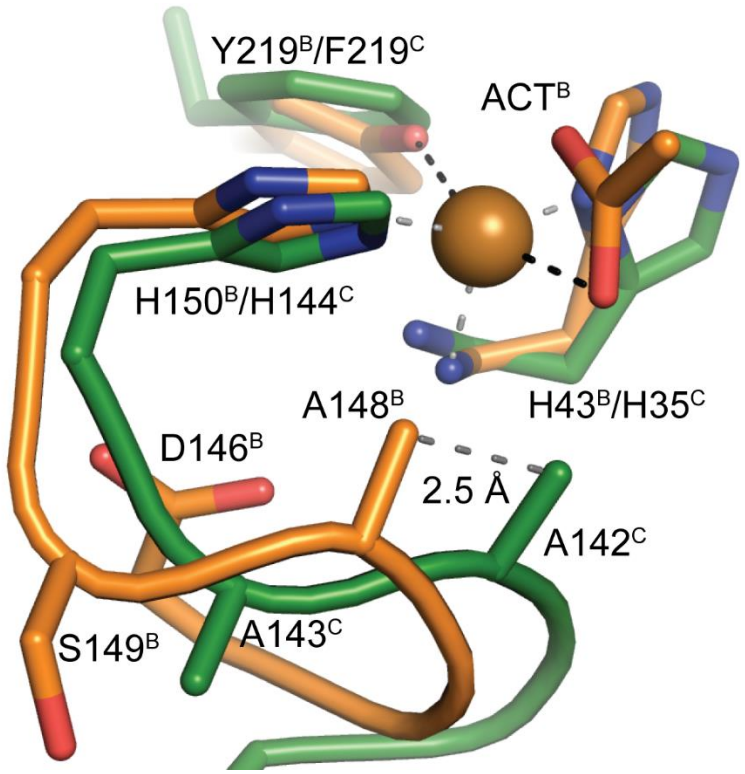


Figure 5.1. Position of the conserved aspartic acid in ScLPMO10B*. The highly conserved aspartic acid (represented as D146 in ScLPMO10B; colored orange) in C1/C4 oxidizing AA10 proteins have been suggested to be responsible for the alternative conformation of the loop which harbors the conserved active site alanine (represented as A148 in ScLPMO10B). Note the alanines in ScLPMO10B and ScLPMO10C that are positioned 2.5 Å from one another. Figure source: Modified from Forsberg et al. (2014b).*

The β-sandwich of LPMOs contains several β-strands connected to one another via short loops. One of these loop is positioned close to the active site of AA10 proteins. The loop, comprising nine residues, has low sequence similarity among C1 and C1/C4 oxidizing AA10 proteins (Figure 4.2), and is positioned more proximal to the active site in C1/C4 oxidizers than in C1 oxidizers (Figure 5.2). As the loop is located at the substrate-binding surface of the proteins, it was

presumed to be relevant for substrate-specificity. Thus, attempts were made to replace all nine residues in the loop of *Ma*LPMO10B with the corresponding residues of *Sc*LPMO10C. However, the attempts were unsuccessful, and the experiment was discontinued.

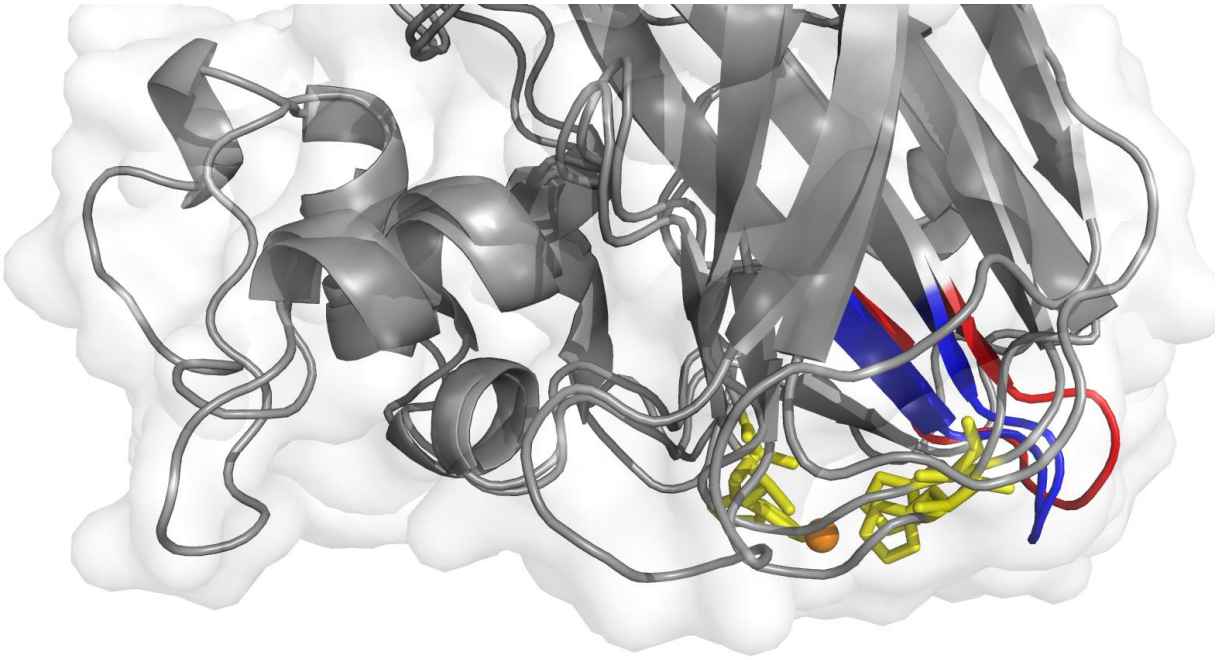


Figure 5.2. Loop conformation in ScLPMO10B and ScLPMO10C. *The copper ion is represented as a sphere (colored orange), whereas the conserved active-site histidines are represented as sticks (colored yellow). The rest of the protein structures are colored grey. Notice the loop orientation of ScLPMO10B (blue), which is positioned more proximally to the active-site copper ion than the loop of ScLPMO10C (red). The loops are located at the substrate-binding surface of the proteins. The structure is that of ScLPMO10B, a close homolog of MaLPMO10B. Figure made using PyMOL (Delano & Lam, 2005).*

As previously mentioned, phylogenetic and correlation mutation analysis [(CMA), performed by B. Bisarro, unpublished work; an outline of the method is included in Appendix F] gave background for an even more advanced analysis of C4-oxidation determining activity than just by analyzing MSAs and protein structure. The method, which involved structure-based multiple sequence alignments combined with CMA of 55 sequences encoding C1 and C1/C4 oxidizing cellulose-active AA10 proteins, led to identification of residues Trp82, Asn85 and Gln141 that were thus targeted for site-directed mutagenesis (Figure 5.3). These residues were identified as

being highly conserved in C1 oxidizing cellulose-active AA10 proteins (as tyrosine, phenylalanine and tryptophan, respectively). The corresponding residues of Trp82 and Asn85 (in *Ma*LPMO10B) are in C1 oxidizers highly conserved as a pair, and therefore it was attempted to study the combined effect of mutating both these positions.

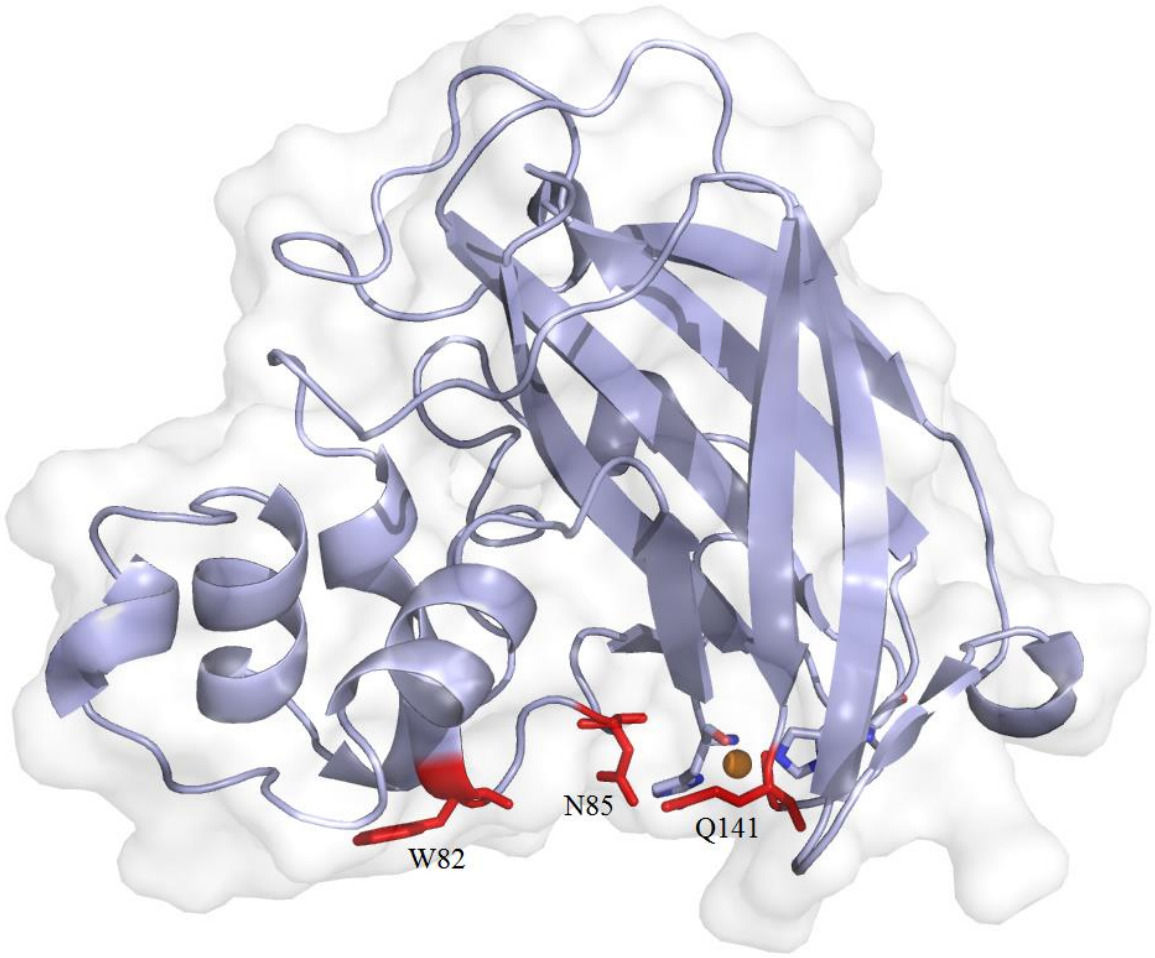


Figure 5.3. Residues targeted for site-directed mutagenesis*. Residues W82, N85 and Q141 are represented as sticks, and colored red. The active site copper ion is represented as a sphere, and colored orange. The structure is that of MaLPMO10B. Figure made using PyMOL (Delano & Lam, 2005).*

Qualitative and quantitative analyses were conducted with all mutants (i.e. W82Y, N85F, Q141W, as well as double mutant W82Y/N85F) acting on Avicel, PASC or β-chitin, and their activity was compared against that of the wild type protein and a strict C1 oxidizing LPMO (i.e. *Sc*LPMO10C). All of the mutants were more or less active on both cellulose (i.e. Avicel and

PASC) and β-chitin; however, one clear finding stood out. The C4 oxidizing activity of the double mutant, W82Y/N85F, was almost completely absent. In initial qualitative experiments, there were indications of total elimination of the C4 oxidizing activity as no C4 or C4/C1 oxidized products were detected using MALDI-TOF MS (Figure 4.18) and HPLC (Figure 4.17. However, further analyses demonstrated that small amounts of C4/C1 oxidized products indeed were generated by the W82Y/N85F mutant (Figures 4.20 and 4.21). Quantification of cellobionic acid (DP2ox) generated by the W82Y/N85F mutant and the wild type LPMO demonstrated that even though the overall activity of the mutant was reduced, the C4 oxidizing activity were clearly impaired relative to the C1 oxidizing activity. Furthermore, by concentrating (3x) the soluble products generated by the W82Y/N85F mutant by vacuum drying, it was demonstrated that the C4 oxidizing activity of the mutant was almost abolished compared to the C1 oxidizing activity (Figure 4.22). It should be noted that vacuum drying might lead to changes in the chemistry of the generated products, which may result in inaccurate identification of oxidized products by HPLC (G. Vaaje-Kolstad, personal communication).

The importance of Trp82 and Asn85 for oxidative regioselectivity in cellulose-active AA10 proteins are further demonstrated by the fact that both W82Y and N85F single-mutants showed reduced C4 oxidizing activity. Both mutants generated far less C4/C1 oxidized products relative to C1 oxidized products. Interestingly, it appears that a combination of the two is important for the C4 oxidizing activity of these proteins, since the C4 oxidizing activity level of the double-mutant is lower than the single-mutants. This raises the question regarding which other positions may be involved in determining the C4/C1 activity of cellulose-active AA10 proteins. One possibility is the glutamine at position 141 in the *Ma*LPMO10B sequence. As with the W82Y and N85F mutants, this residue was identified by structure-based MSA and CMA analysis (B. Bisarro, unpublished work), and activity assays demonstrated that the C4 oxidizing activity of Q141W mutant was reduced. However, as we were unable to obtain a triple-mutant including all three residues, further research is necessary to expose the role of Q141 in oxidative regioselectivity.

As previously mentioned, activity assays demonstrated that all of the mutants were active on β-chitin (Figure 4.14). Interestingly, the mutant that showed least activity (i.e. W82Y) contained a mutated tyrosine in position 82. A tyrosine at the corresponding position in the chitin-active

AA10 protein CBP21 have previously been identified as important chitin binding (Aachmann et al., 2012). One could argue that replacing the tryptophan with a tyrosine would increase the mutants binding affinity towards β-chitin, but this result demonstrates that the substrate-binding specificity of these proteins are highly complex.

With the discovery of C1/C4 oxidizing LPMOs, the study by Forsberg *et al*. in 2014 also demonstrated synergistic action between strict C1 oxidizing and C1/C4 oxidizing LPMOs. Inspired by these findings, we investigated the possibility of synergy between the C1/C4 oxidizing wild type *Ma*LPMO10B and the mainly C1 oxidizing W82Y/N85F mutant (Figure 4.23 and 4.24). Analyzing quantified cellobionic acid, the experiment demonstrated a clear synergistic effect the first two hours of incubation when combining the wild type and the double-mutant. Interestingly, the wild type and the double-mutant generated equal amounts of cellobionic acid during the first hour of incubation, before the activity of the double-mutant ceased. Analysis of C4/C1 oxidized products demonstrated synergism between wild type and mutant (i.e. W82Y/N85F) *Ma*LPMO10B the first two hours of incubation. Surprisingly, combining the strict C1 oxidizing *Sc*LPMO10C with wild type or mutant *Ma*LPMO10B resulted in no C4/C1 oxidized products being formed the first three hours of incubation. As expected, virtually no products were formed from a mixture of *Sc*LPMO10C and *Ma*LPMO10B W82Y/N85F, but the lack of product formation from *Sc*LPMO10C and wild type *Ma*LPMO10B was unexpected. What caused *Ma*LPMO10B to be inactive the first hours of incubation is unknown, and would require further research. The result suggests inhibition effects, rather than synergy effects between the two LPMOs regarding C4 oxidation of the substrate. Interestingly, it appears the C1 activity of the LPMOs is not affected, meaning *Sc*LPMO10C somehow blocks C4 oxidation by *Ma*LPMO10B. The results from the synergy experiment have some limitations. First, dose-response effects were not tested, meaning a linear relationship between LPMO concentration and product formation cannot be claimed. In other words, the effect of 1 µM of a specific LPMO might not be twice the effect of 0.5 µM of the same LPMO. The synergistic effects of the reactions are therefore inconclusive. Second, due to the fact that some of the LPMOs were not completely purified, means that LPMO concentrations may be inaccurate. Last, quantification of C4 oxidized products have not been performed, meaning conclusions regarding the amount of C4 oxidized products formed cannot be drawn.

The structure of the catalytic domain of both wild type (1.08 Å) and mutant (i.e. W82Y/N85F; 2.7 Å; the structure was not determined by the end of this study) were solved by X-ray crystallography (performed by Dr. Bjørn Dalhus, Oslo University Hospital). The structure of wild type *Ma*LPMO10B was compared to the structures to other known C1/C4 oxidizing cellulose-active AA10 proteins, namely *Tf*LPMO10A and *Sc*LPMO10B. The structural alignment demonstrated that *Ma*LPMO10B contains a prolonged loop, connecting two β-strands in the β-sandwich region of the catalytic domain (Figure 5.4). The prolonged loop, comprising eight additional residues, are turning away from the substrate-binding surface. Since all three LPMOs are both cellulose (C1/C4 oxidizing) and chitin (C1 oxidizing) active, the prolonged loop of *Ma*LPMO10B does not appear to be important for substrate specificity or oxidative regioselectivity.

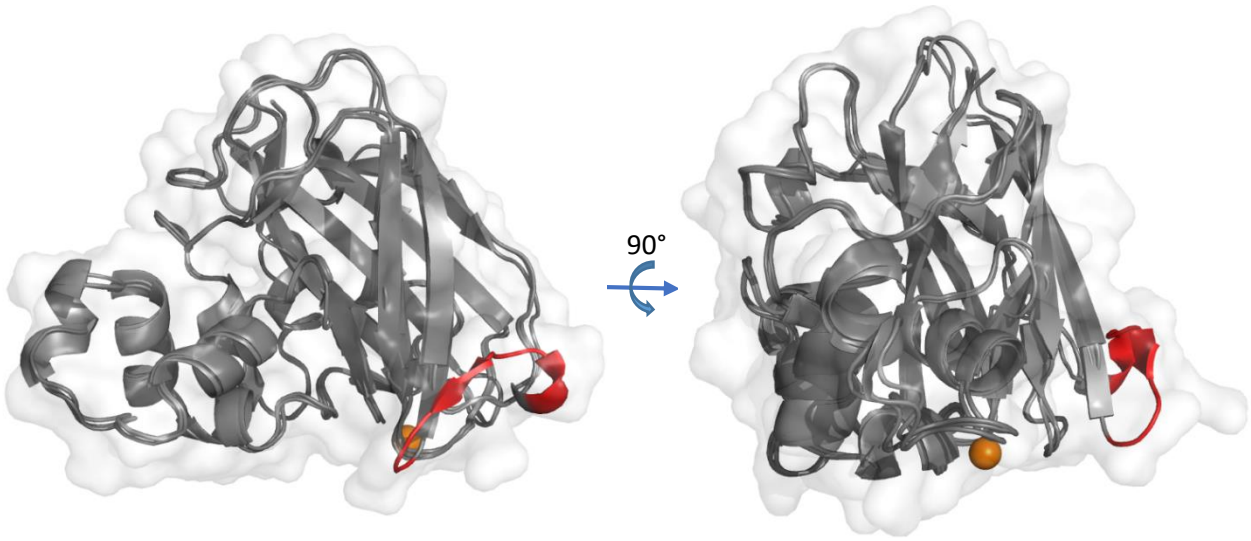


Figure 5.4. Structural alignment of three cellulose-active C1/C4 oxidizing AA10 proteins. *The structures of ScLPMO10B, TfLPMO10A and MaLPMO10B are colored grey. The active site copper ion is represented as a sphere (colored orange). Note the prolonged loop of MaLPMO10B (colored red) which turns away from the substrate-binding surface of the protein.*

5.1 Concluding remarks and future work

The present study has demonstrated that residues W82 and N85 play an important role in the determination of oxidative regioselectivity in *Ma*LPMO10B. Analyses of the oxidative activity of *Ma*LPMO10B double-mutant W82Y/N85F have shown that the C4 oxidizing activity was almost abolished compared to that of the wild type LPMO. The C1 oxidizing activity of the double-mutant was also reduced, but the relative reduction was far from that of the C4 oxidizing activity.

In further research of oxidative regioselectivity in *Ma*LPMO10B, a next step may be to successfully produce a triple-mutant of W82Y/N85F/Q141G, as all three positions were shown to be highly conserved in strict C1 oxidizing cellulose-active AA10 LPMOs (B. Bisarro, personal communication, unpublished work). Further research is necessary to identify, produce and analyze the activity of additional mutants in order to possibly eliminate the C4 oxidizing activity completely.

Since both positions of the double-mutant (i.e. 82 and 85 in the sequence of *Ma*LPMO10B) are highly conserved among cellulose-active C1/C4 oxidizing AA10 LPMOs, one might assume that similar mutations in the corresponding positions of other cellulose-active C1/C4 oxidizing AA10 LPMOs might lead to the same reduction of C4 oxidizing activity.

6 References

Aachmann, F. L., Sørlie, M., Skjåk-Bræk, 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. *PNAS Journal, 109*(46), 18779-18784.

Agger, J. W., Isaksen, T., Várnai, 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. *PNAS Journal, 111*(17), 6287-6292.

Alvarez, F. J. (2014). The Effect of Chitin Size, Shape, Source and Purification Method on Immune Recognition. *Molecules, 19*, 4433-4451.

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. *Proc Natl Acad Sci U S A, 111*(25), 9109-9114.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., & Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res., 34*, 604-608.

Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M., & Marletta, M. A. (2015). Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem., 84*, 923-946.

Beier, S., & Bertilsson, S. (2013). Bacterial chitin degradation - mechanisms and ecophysiological strategies. *Front Microbiol, 4*.

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. *Biochem. J, 331*, 775-781.

Book, A. J., Yennamalli, R. M., Takasuka, T. E., Currie, C. R., Jr, G. N. P., & Fox, B. G. (2014). Evolution of substrate specificity in bacterial AA10 lytic polysaccharide monooxygenases. *Biotechnology for Biofuels, 7*(109).

Boraston, A. B., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J., 382*, 769-781.

Brown, R. M. (2004). Cellulose structure and biosynthesis: What is in store for the 21st century? *Journal of Polymer Science, 42*(3), 487-495.

Busk, P. K., & Lange, L. (2015). Classification of fungal and bacterial lytic polysaccharide monooxygenases. *BMC Genomics, 16*(368).

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 Res, 37*, 233-238.

Chen, X., Chew, S. L., Kerton, F. M., & Yan, N. (2014). Direct conversion of chitin into a N-containing furan derivative. *Green Chem, 16*, 2204-2212.

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. *J Biol Chem, 291*(14), 7439-7449. doi: 10.1074/jbc.M115.702365

Delano, W. L., & Lam, J. W. (2005). PyMOL: A communications Tool For Computational Models. *Abstr Pap Am Chem S*(230), 1371-1372.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.-M., & Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research, 36*, 465-469.

Ding, S.-Y., & Himmel, M. E. (2006). The Maize Primary Cell Wall Microfibril: A New Model Derived from Direct Visualization. *J. Agric. Food Chem, 54*(3), 597-606.

Doblin, M. S., Kurek, I., Jacob-Wilk, D., & Delmer, D. P. (2002). Cellulose Biosynthesis in Plants: from Genes to Rosettes. *Plant Cell Physiol, 43*(12), 1407-1420.

Eibinger, M., Ganner, T., Bubner, P., Rosker, S., Kracher, D., Haltrich, D., Ludwig, R., Plank, H., & Nidetzky, B. (2014). Cellulose Surface Degradation by a Lytic Polysaccharide Monooxygenase and Its Effect on Cellulase Hydrolytic Efficiency. *Journal of Biological Chemistry, 289*(52), 35929-35938. doi: 10.1074/jbc.M114.602227

Forsberg, Z., Mackenzie, A. K., Sørlie, M., Røhr, Å. K., Helland, R., Arvai, A. S., Vaaje-Kolstad, G., & Eijsink, V. G. H. (2014b). Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A, 111*(23), 8446-8451.

Forsberg, Z., Nelson, C. E., Dalhus, B., Mekasha, S., Loose, J. S., Crouch, L. I., Rohr, A. K., Gardner, J. G., Eijsink, V. G., & Vaaje-Kolstad, G. (2016). Structural and Functional Analysis of a Lytic Polysaccharide Monooxygenase Important for Efficient Utilization of Chitin in Cellvibrio japonicus. *J Biol Chem, 291*(14), 7300-7312. doi: 10.1074/jbc.M115.700161

Forsberg, Z., Røhr, Å. K., Mekasha, S., Andersson, K. K., Eijsink, V. G. H., Vaaje-Kolstad, G., & Sørlie, M. (2014a). Comparative Study of Two Chitin-Active and Two Cellulose-Active AA10-Type Lytic Polysaccharide Monooxygenases. *Biochemistry, 53*, 1647-1656.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., Mackenzie, A., Sørlie, M., Horn, S. J., & Eijsink, V. G. H. (2011). Cleavage of cellulose by a CBM33 protein. *Protein Science, 20*, 1479-1483.

Hall, M., Bansal, P., Lee, J. H., Realff, M. J., & Bommarius, A. S. (2010). Cellulose crystallinity - a key predictor of the enzymatic hydrolysis rate. *FEBS Journal, 277*(6), 1571-1582.

Hallac, B. B., & Ragauskas, A. J. (2011). Analyzing cellulose degree of polymerization and its relevancy to cellulosic ethanol. *Biofuels, Bioprod. Bioref., 5*, 215-225.

Harris, P. V., Welner, D., Mcfarland, K. C., Re, E., Poulsen, J.-C. N., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., & Leggio, L. L. (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.

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., Henrissat, B., Davies, G. J., & Walton, P. H. (2014). Discovery and characterization of a new family of lytic polysaccharide mono-oxygenases. *Nat Chem Biol, 10*(2), 122-126.

Hemsworth, G. R., Johnston, E. M., Davies, G. J., & Walton, P. H. (2015). Lytic Polysaccharide Monooxygenases in Biomass Conversion. *Trends in Biotechnology, 33*(12), 747-761.

Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J., & Walton, P. H. (2013). The Copper Active Site of CBM33 Polysaccharide Oxygenases. *J. Am. Chem. Soc., 135*, 6069-6077.

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. doi: 10.1073/pnas.0608909103

Horn, S. J., Sørbotten, A., Synstad, B., Sikorski, P., Sørlie, M., Vårum, K. M., & Eijsink, V. G. H. (2006). Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. *FEBS Journal, 273*, 491-503.

Horn, S. J., Vaaje-Kolstad, G., Westereng, B., & Eijsink, V. G. H. (2012). Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels, 5*(45).

Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttilä, 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. *The Journal of Biological Chemistry, 289*(5), 2632-2642.

Jalak, J., Kurašin, M., Teugjas, H., & Väljamäe, P. (2012). Endo-exo Synergism in Cellulose Hydrolysis Revisited. *J Biol Chem, 287*(34), 28802-28815.

Karlsson, J., Saloheimo, M., Siika-Aho, M., Penttilä, M., & Tjerneld, F. (2001). Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur. J. Biochem., 268*, 6498-6507.

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

Kuipers, R. K. P., Joosten, H.-J., Verwiel, E., Paans, S., Akerboom, J., Oost, J. V. D., Leferink, N. G. H., Berkel, W. J. H. V., Vriend, G., & Schaap, P. J. (2009). Correlated mutation analyses on super-family alignments reveal functionally important residues. *Proteins: Structure, Function, and Bioinformatics, 67*(3), 608-616.

Kumirska, J., Czerwicka, M., Kaczyński, Z., Bychowska, A., Brzozowski, K., Thöming, 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.

Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., & Sweeney, M. D. (2011). Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Appl Environ Microbiol., 77*(19), 7007-7015.

Leggio, L. L., Simmons, T. J., Poulsen, J.-C. N., Frandsen, K. E. H., Hemsworth, G. R., Stringer, M. A., Freiesleben, P. V., Tovborg, M., Johansen, K. S., Maria, L. D., Harris, P. V., Soong, C.-L., Dupree, P., Tryfona, T., Lenfant, N., Henrissat, B., Davies, G. J., & Walton, P. H. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nature Communications, 6*.

Letunic, I., & Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Research, 39*, 475-478.

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*(41).

Li, S., Bashline, L., Lei, L., & Gu, Y. (2014). Cellulose Synthesis and Its Regulation. *Arabidopsis Book, 12*.

Li, X., William T. Beeson, I., Phillips, C. M., Marletta, M. A., & Cate, J. H. D. (2012). Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure, 20*(6), 1051-1061.

Loose, J. S. M., Forsberg, Z., Fraaije, M. W., Eijsink, V. G. H., & Vaaje-Kolstad, G. (2014). A rapid quantitative activity assay shows that the Vibrio cholerae colonization factor GbpA is an active lytic polysaccharide monooxygenase. *Febs Letters, 588*(18), 3435-3440. doi: 10.1016/j.febslet.2014.07.036

Matthews, J. F., Beckham, G. T., BergenstråHle-Wohlert, M., Brady, J. W., Himmel, M. E., & Crowley, M. F. (2012). Comparison of Cellulose I*β* Simulations with Three CarbohydrateForce Fields. *J. Chem. TheoryComput, 8*, 735-748.

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.

Nakagawa, Y. S., Kudo, M., Loose, J. S., Ishikawa, T., Totani, K., Eijsink, V. G., & Vaaje-Kolstad, G. (2015). A small lytic polysaccharide monooxygenase from Streptomyces griseus targeting alpha- and beta-chitin. *FEBS J, 282*(6), 1065-1079. doi: 10.1111/febs.13203

Orłowski, A., Róg, T., Paavilainen, S., Manna, M., Heiskanen, I., Backfolk, K., Timonen, J., & Vattulainen, I. (2015). How endoglucanase enzymes act on cellulose nanofibrils: role of amorphous regions revealed by atomistic simulations. *Cellulose, 22*(5), 2911-2925.

Park, S., Baker, J. O., Himmel, M. E., Parilla, P. A., & Johnson, D. K. (2010). Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol Biofuels, 3*(10).

Phillips, C. M., William T. Beeson, I., Cate, J. H., & Marletta, M. A. (2011). Cellobiose Dehydrogenase and a Copper-Dependent Polysaccharide Monooxygenase Potentiate Cellulose Degradation by *Neurospora crassa*. *ACS Chem. Biol., 6*, 1399-1406.

Quinlan, R. J., Sweeney, M. D., Leggio, L. L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., Krogh, K. B. R. M., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., & Walton, P. H. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A, 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 Environ. Sci., 6*, 1858-1867.

Sorensen, T. H., Cruys-Bagger, N., Windahl, M. S., Badino, S. F., Borch, K., & Westh, P. (2015). Temperature Effects on Kinetic Parameters and Substrate Affinity of Cel7A Cellobiohydrolases. *Journal of Biological Chemistry, 290*(36), 22193-22202. doi: 10.1074/jbc.M115.658930

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. *Biosci. Biotechnol. Biochem., 66*(5), 1075-1083.

Vaaje-Kolstad, G., Bøhle, L. A., Gåseidnes, S., Dalhus, B., Bjørås, 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., Aalten, D. M. F. V., Synstad, B., & Eijsink, V. G. H. (2005b). The Non-catalytic Chitin-binding Protein CBP21 from Serratia marcescens Is Essential for Chitin Degradation. *The Journal of Biological Chemistry, 280*(31), 28492-28497.

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. doi: 10.1111/febs.12181

Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H. K., Eijsink, V. G. H., & Aalten, D. M. F. V. (2005a). Crystal Structure and Binding Properties of the Serratia marcescens Chitin-binding Protein CBP21. *The Journal of Biological Chemistry, 280*(12), 11313-11319.

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., & Eijsink, V. G. H. (2010). An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science, 330*(8), 219-222.

Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D., & Marletta, M. A. (2014b). Determinants of Regioselective Hydroxylation in the Fungal Polysaccharide Monooxygenases. *J. Am. Chem. Soc., 136*, 562-565.

Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R., & Marletta, M. A. (2014a). A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci U S A, 111*(38), 13822-13827.

Watanabe, T., Kimura, K., Sumiya, T., Nikaidou, N., Suzuki, K., Suzuki, M., Taiyoji, M., Ferrer, S., & Regue, M. (1997). Genetic Analysis of the Chitinase System of Serratia marcescens 2170. *Journal of Bacteriology, 179*(22), 7111-7117.

Westereng, B., Cannella, D., Agger, J. W., Jørgensen, H., Andersen, M. L., Eijsink, V. G. H., & Felby, C. (2015). Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Scientific Reports, 5*.

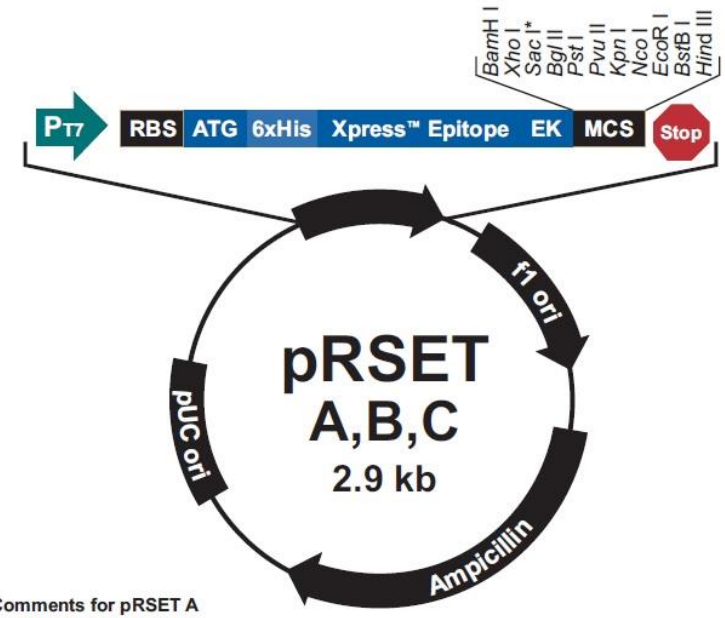
Wu, M., Beckham, G. T., Larsson, A. M., Ishida, T., Kim, S., Payne, C. M., Himmel, M. E., Crowley, M. F., Horn, S. J., Westereng, B., Igarashi, K., Samejima, M., Ståhlberg, J., Eijsink, V. G. H., & Sandgren, M. (2013). Crystal Structure and Computational Characterization of the Lytic Polysaccharide Monooxygenase GH61D from the Basidiomycota Fungus Phanerochaete chrysosporium. *The Journal of Biological Chemistry, 288*(18), 12828-12839.

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

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

Appendix A

a)



Comments for pRSET A
2897 nucleotides

*Version C does not contain *Sac* I

T7 promoter: bases 20-39
6xHis tag: bases 112-129
T7 gene 10 leader: bases 133-162
Xpress™ epitope: bases 169-192
Multiple cloning site: bases 202-248
T7 reverse priming site: bases 295-314
T7 transcription terminator: bases 256-385
f1 origin: bases 456-911
bla promoter: bases 943-1047
Ampicillin (*bla*) resistance gene (ORF): bases 1042-1902
pUC origin: bases 2047-2720 (C)

b)

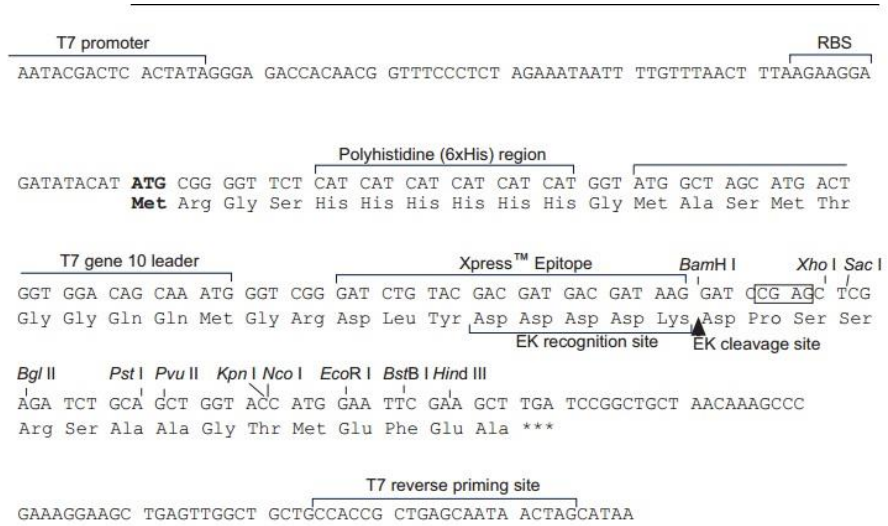


Figure A1: Overview of the pRSET-B expression vector. *A vector map of pRSET A, B and C is shown in panel a), whereas the sequence of the multiple cloning site of pRSET-B is shown in panel b)). Figure sources: Invitrogen*

Appendix B

Codon optimized sequence of *Malpmo*10B gene (Optimized Sequence Length:1101, GC%:60.16). The sequence includes a native signal sequence.

ATGTCAACGCCGTATCGTCGCCCGCTGCCGCTGGCCGCCGCAATCCTGGGTGTTTGTGCCGTC
GTCGCTGCTCTGCTGACCACCGCTTTTAGCGGTCCGGCCAGCGCGCATGGCAGCGTGGTTGAT
CCGGCCAGCCGTTCTTATAGTTGCTGGCAGCGCTGGGGCGGTGATTTCCAAAACCCGGCAAT
GGCTACCCAGGACCCGATGTGTTGGCAGGCGTGGCAAGCCGATCCGAACGCAATGTGGAACT
GGAATGGCCTGTTTCGTGAAGGTGTGGCGGGCAATCATCAAGGCGCCATTCCGGATGGTCAG
CTGTGCTCGGGCGGTCGTACCCAGAGCGGCCGTTATAATGCACTGGATACCGTGGGTGCTTG
GAAAACGGTTCCGGTCACCAACAACTTCCGTGTTAAATTTTTCGATCAAGCGAGCCACGGCG
CCGACTATATTCGCGTGTACGTTACGAAACAGGGTTACAACGCACTGACCAGTCCGCTGCGT
TGGTCCGACCTGGAACTGGTGGGTCAAATTGGTAATACGCCGGCCTCCCAGTGGACCCGCGA
AGTTGATGGTGTCTCAATTCAGATTCCGGCCAATGCACCGGGTCGTACCGGCCGCCATGTCGT
GTATACCATCTGGCAGGCGTCACACCTGGACCAATCGTATTACCTGTGCAGCGATGTTGACTT
TGGCGGTTCTGGTCCGACCACGCCGCCGACGTCTACCCCGCCGACGAGTACCCCGCCGACGT
CCACCCCGCCGACCAGCCCGCCGCCGACCACGCCGAATCCGGCGGGCGGTTGTACCGCAACC
TACGCTATTACCGGTAGCTGGGGCGGTGGCTTCCAGGCAGATGTCAAAGTGACCAATGGTTC
CGGCTCACCGATCCGTGGCTGGTCGGTTAGCTGGAACTATCAGAATGGTCAGCAAGTTAACT
CTGCGTGGAATGCCACCGTCACCACGAGTGGCACGCTGGTCACCGCACGCAACGTGGCTTAC
AATGGCTCGCTGGCACCGGGTGCTTCTACGAGTTTTGGTTTCACCGGCAGCGCAGGTGCTACC
AATCCGGTGCCGTCCATTGTTTCATGTACCACGACCAGTTAA

Codon optimized sequence of the *Malpmo*10D gene (Optimized Sequence Length:963, GC%:60.05). The sequence does not include a signal sequence.

CATGGCAGTGTTACGAATCCGCCGACCCGCAATTATGGCTGTTGGGAACGCTGGGGCAGTGA
CCACCTGAACCCGACGATGGCACAAACCGATCCGATGTGCTGGCAGGCGTGGCAAGCCAAC
CCGAATACCATGTGGAACTGGAATGGCCTGTATCGTGAAAACGTGGGCGGTAATCATCAGGC
GGCCGTTCCGGATGGTCAACTGTGTAGTGGCGGTCGCACCCAGGGCGGTCTGTATGCCTCCC
TGGACGCAGTGGGTGCTTGGACCGCGAAACCGATGCCGAACAATTTTACCCTGACGCTGACC
GATGGCGCCAAACACGGTGCAGACTATATGCTGATTTACATCACCAAACAGGGTTTCGATCC
GACCACGCAACCGCTGACGTGGAACTCACTGGAACTGGTCCTGCGTACCGGCTCGTATCCGA
CCACGGGTCTGTACGAAGCTCAGGTGAACGCGGGCAATCGTACGGGTCGCCATGTGGTTTAT
ACCATTTGGCAGGCAAGCCACCTGGATCAACCGTATTACCTGTGCTCTGACGTTATCTTTGGC
GGTGGCGGTAACCCGCCGCCGACCACCCCGCCGCCGACGACCCCGCCGCCGACTACCCCGCC
GCCGACAACCCCGCCGCCGACCACCCCGCCGCCGGCGGGCAACGGTGCCTGTACGGCAACCT
ATCGTAAAACGAATGAATGGAGTGGCGGTTTCGGTGCCGAAGTCACCGTGCGTGCCGGTAAC
GCAGCTATTTCTGGTTGGACGGTTGCGTGGACCTGGCCGAATGGCCAGAGCATCACCAACGC
TTGGAATGCGACGGTCACCAGCTCTGGTAGTTCCGTTACCGCCAGCAACGTCGGCTACAATG
GTTCACTGTCGGCTAACGGCAGCACGTCTTTTGGTTTCAATGCAAGTTGGAACGGCACCAAC
AGCACGCCGACCCTGACCTGCACCGCCCGCTAA

Appendix C

Amino acid sequence of *Ma*LPMO10B. The sequence includes a signal peptide (yellow). *Ma*LPMO10B^{cd} does not include the CBM module (grey).

MSTPYRRPLPLAAAILGVCAVVAALLTTAFSGPASAHGSVVDPASRSYSCWQRWGGDFQNPAMATQDPM
CWQAWQADPNAMWNWNGLFREGVAGNHQGAIPDGQLCSGGRTQSGRYNALDTVGAWKTVPVTNNFRV
KFFDQASHGADYIRVYVTKQGYNALTSPLRWSDLELVGQIGNTPASQWTREVDGVSIQIPANAPGRTGRHV
VYTIWQASHLDQSYYLCSDVDFGGSGPTTPPTSTPPTSTPPTSTPPTSPPPTTPNPAGGCTATYAITGSWGGGF
QADVKVTNGSGSPIRGWSVSWNYQNGQQVNSAWNATVTTSGTLVTARNVAYNGSLAPGASTSFGFTGSA
GATNPVPSIVSCTTTS

Amino acid sequence of *Ma*LPMO10D. The sequence does not include a signal peptide. The CBM module and associated linker is highlighted (grey).

HGSVTNPPTRNYGCWERWGSDHLNPTMAQTDPMCWQAWQANPNTMWNWNGLYRENVGGNHQAAVPD
GQLCSGGRTQGGLYASLDAVGAWTAKPMPNNFTLTLTDGAKHGADYMLIYITKQGFDPTTQPLTWNSLEL
VLRTGSYPTTGLYEAQVNAGNRTGRHVVYTIWQASHLDQPYYLCSDVIFGGGGNPPPTTPPPTTPPPTTPPP
TTPPPTTPPPAGNGACTATYRKTNEWSGGFGAEVTVRAGNAAISGWTVAWTWPNGQSITNAWNATVTSSG
SSVTASNVGYNGSLSANGSTSFGFNASWNGTNSTPTLTCTAR

Amplification of *Ma*LPMO10D (see section 3.4.1) resulted in two additional variants of the gene; both missing several amino acids in the linker region. A multiple sequence alignment of *Ma*LPMO10D variants is shown in figure C.1, illustrating the locations of the missing amino acids.

```
MaLPMO10D (-20 aa)      HGSVTNPPTRNYGCWERWGSDHLNPTMAQTDPMCWQAWQANPNTMWNWNGLYRENVGGNH
MaLPMO10D               HGSVTNPPTRNYGCWERWGSDHLNPTMAQTDPMCWQAWQANPNTMWNWNGLYRENVGGNH
MaLPMO10D[a1]           HGSVTNPPTRNYGCWERWGSDHLNPTMAQTDPMCWQAWQANPNTMWNWNGLYRENVGGNH
                        ************************************************************

MaLPMO10D (-20 aa)      QAAVPDGQLCSGGRTQGGLYASLDAVGAWTAKPMPNNFTLTLTDGAKHGADYMLIYITKQ
MaLPMO10D               QAAVPDGQLCSGGRTQGGLYASLDAVGAWTAKPMPNNFTLTLTDGAKHGADYMLIYITKQ
MaLPMO10D[a1]           QAAVPDGQLCSGGRTQGGLYASLDAVGAWTAKPMPNNFTLTLTDGAKHGADYMLIYITKQ
                        ************************************************************

MaLPMO10D (-20 aa)      GFDPTTQPLTWNSLELVLRTGSYPTTGLYEAQVNAGNRTGRHVVYTIWQASHLDQPYYLC
MaLPMO10D               GFDPTTQPLTWNSLELVLRTGSYPTTGLYEAQVNAGNRTGRHVVYTIWQASHLDQPYYLC
MaLPMO10D[a1]           GFDPTTQPLTWNSLELVLRTGSYPTTGLYEAQVNAGNRTGRHVVYTIWQASHLDQPYYLC
                        ************************************************************

MaLPMO10D (-20 aa)      SDVIFGGGGNPPPT--------------------TPPPAGNGACTATYRKTNEWSGGFGA
MaLPMO10D               SDVIFGGGGNPPPTTPPPTTPPPTTPPPTTPPPTTPPPAGNGACTATYRKTNEWSGGFGA
MaLPMO10D[a1]           SDVIFGGGGNPPPTTPPPT----------TPPPTTPPPAGNGACTATYRKTNEWSGGFGA
                        **************                  **************************

MaLPMO10D (-20 aa)      EVTVRAGNAAISGWTVAWTWPNGQSITNAWNATVTSSGSSVTASNVRYNGSLSANGSTSF
MaLPMO10D               EVTVRAGNAAISGWTVAWTWPNGQSITNAWNATVTSSGSSVTASNVGYNGSLSANGSTSF
MaLPMO10D[a1]           EVTVRAGNAAISGWTVAWTWPNGQSITNAWNATVTSSGSSVTASNVGYNGSLSANGSTSF
                        ******************************************** *************

MaLPMO10D (-20 aa)      GFNASWNGINSTPTLTCTAR
MaLPMO10D               GFNASWNGTNSTPTLTCTAR
MaLPMO10D[a1]           GFNASWNGTNSKPTLTCTAR
                        ********** ********
```

Figure C1: Multiple sequence alignment of MaLPMO10D variants. *The MSA shows the sequence of each MaLPMO10D variant, and the positions of the missing amino acids. Expression of MaLPMO10D (-20 aa) was not attempted.*

Appendix D

Examples of chromatograms obtained following protein purification. Purification of *Ma*LPMO10Dsl by ion exchange chromatography is representative for purification of LPMOs that did not bound, or that bound with low affinity, to the ion exchange column. Ion exchange chromatography (Figure D1) and size-exclusion chromatography (Figure D2) of *Ma*LPMO10D^{sl}.

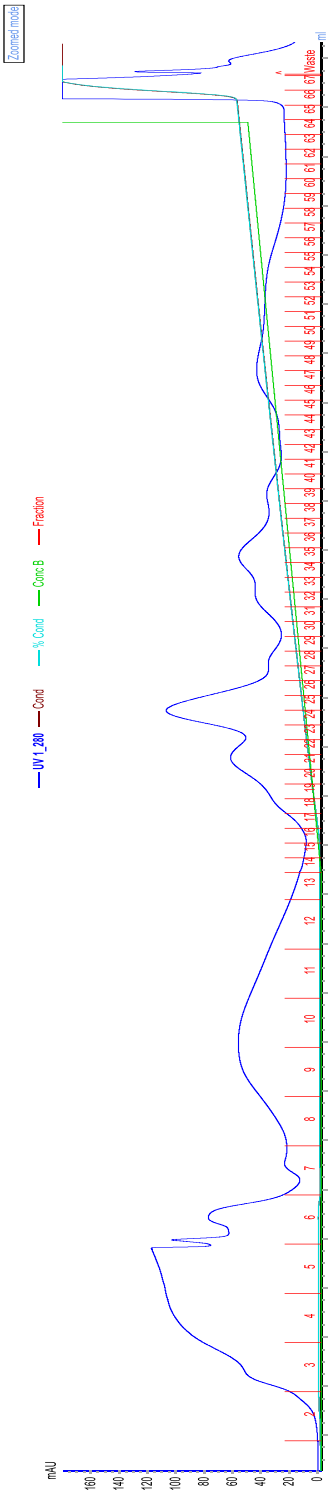


Figure D1. Ion exchange chromatography used for purification of MaLPMO10D^{sl}. *The chromatogram is representative for purification of LPMOs that did not bound, or that bound weakly, to the ion exchange column. The protein of interest eluted in flow-through fractions 3-6, as well as during applying of the linear salt gradient (i.e. fractions 21-27). Fraction numbers are shown in red. The X-axis represents the volume of buffer that eluted through the column. The blue line represents eluted proteins and show the intensity of UV absorbance measured at 280 nm. The green line represents measurement of conductivity used to monitor column equilibration and the linear salt gradient. The protein was purified using the HiTrap DEAE Sepharose FF, 5 ml column (GE Healthcare). Binding buffer: 25 mM Bis-Tris propane pH 9.5. Elution buffer: 25 mM Bis-Tris propane pH 9.5, 1 M NaCl.*

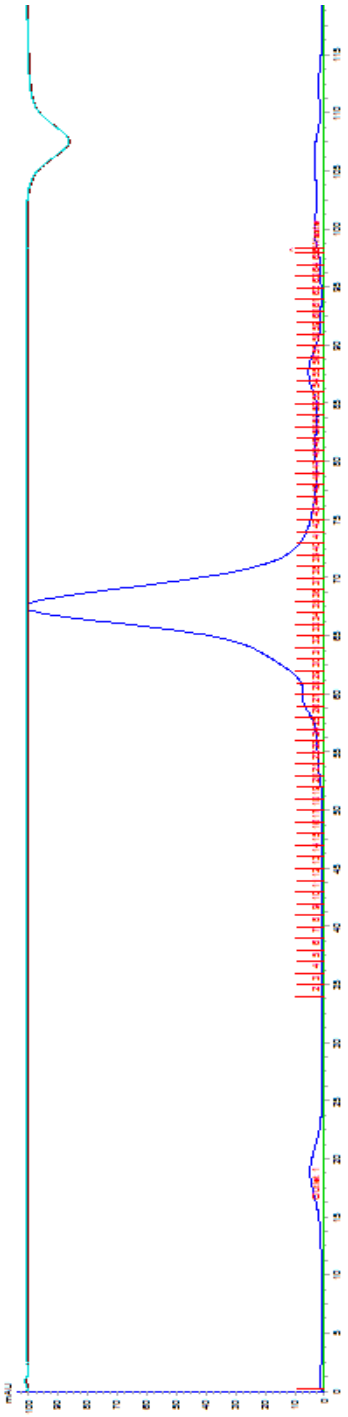


Figure D2. Size-exclusion chromatography used for purification of MaLPMO10D^{sl}. *The main peak represents eluted MaLPMO10D^{sl}. Fraction numbers are shown in red. The X-axis represents the volume of buffer passing through the column. The blue line represents eluted proteins as measured by UV absorbance at 280 nm. The light blue line represents measurement of conductivity, used to monitor equilibration of the column. The protein (1 ml concentrated protein sample) was purified using the HiLoad 16/60 Superdex 75 PG, 120 ml column (GE Healthcare) at a flow-rate of 0.8 ml/min.*

Ion exchange chromatography (Figure D3) and size-exclusion chromatography (Figure D4) of *Ma*LPMO10B[cd].

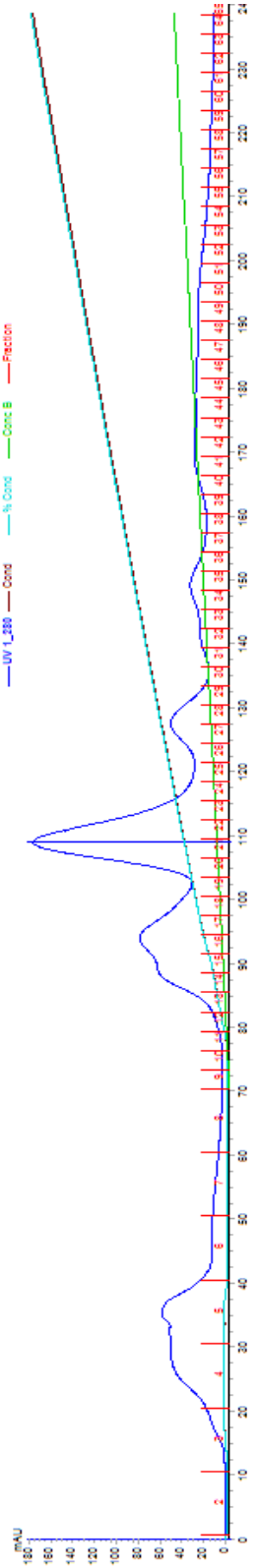


Figure D3. Ion exchange chromatography used for purification of MaLPMO10B[cd]. *The chromatogram is representative for LPMOs that bound with high affinity to the ion exchange column. The protein of interest eluted during applying of a linear salt gradient (i.e. fractions 18-25). Fraction numbers are shown in red. The X-axis represents the volume of buffer passing through the column. The blue line represents eluted proteins and show the intensity of UV absorbance measured at 280 nm. The green line represents measurement of conductivity used to monitor column equilibration and the linear salt gradient. The protein was purified using the HiTrap DEAE Sepharose FF, 5 ml column (GE Healthcare). Binding buffer: 25 mM Bis-Tris propane pH 9.5, 1 M NaCl. Elution buffer: 25 mM Bis-Tris propane pH 9.5. Elution buffer: 25 mM Bis-Tris propane pH 9.5, 1 M NaCl.*

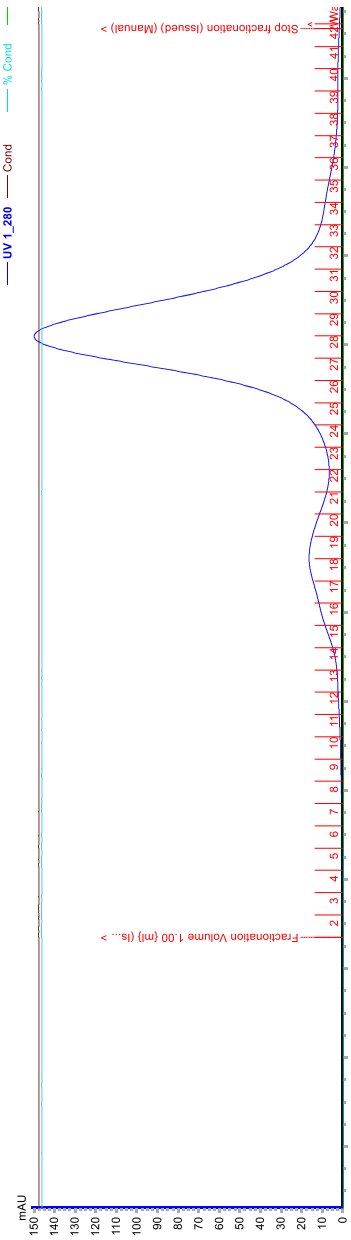


Figure D4. Size-exclusion chromatography used for purification of MaLPMO10B[cd]. *The main peak represents eluted MaLPMO10B[cd]. Fraction numbers are shown in red. The X-axis represents the volume of buffer passing through the column. The blue line represents eluted proteins as measured by UV absorbance at 280 nm. The light blue line represents measurement of conductivity, used to monitor equilibration of the column. The protein (1 ml concentrated protein sample) was purified using the HiLoad 16/60 Superdex 75 PG, 120 ml column (GE Healthcare) at a flow-rate of 0.8 ml/min.*

Appendix E

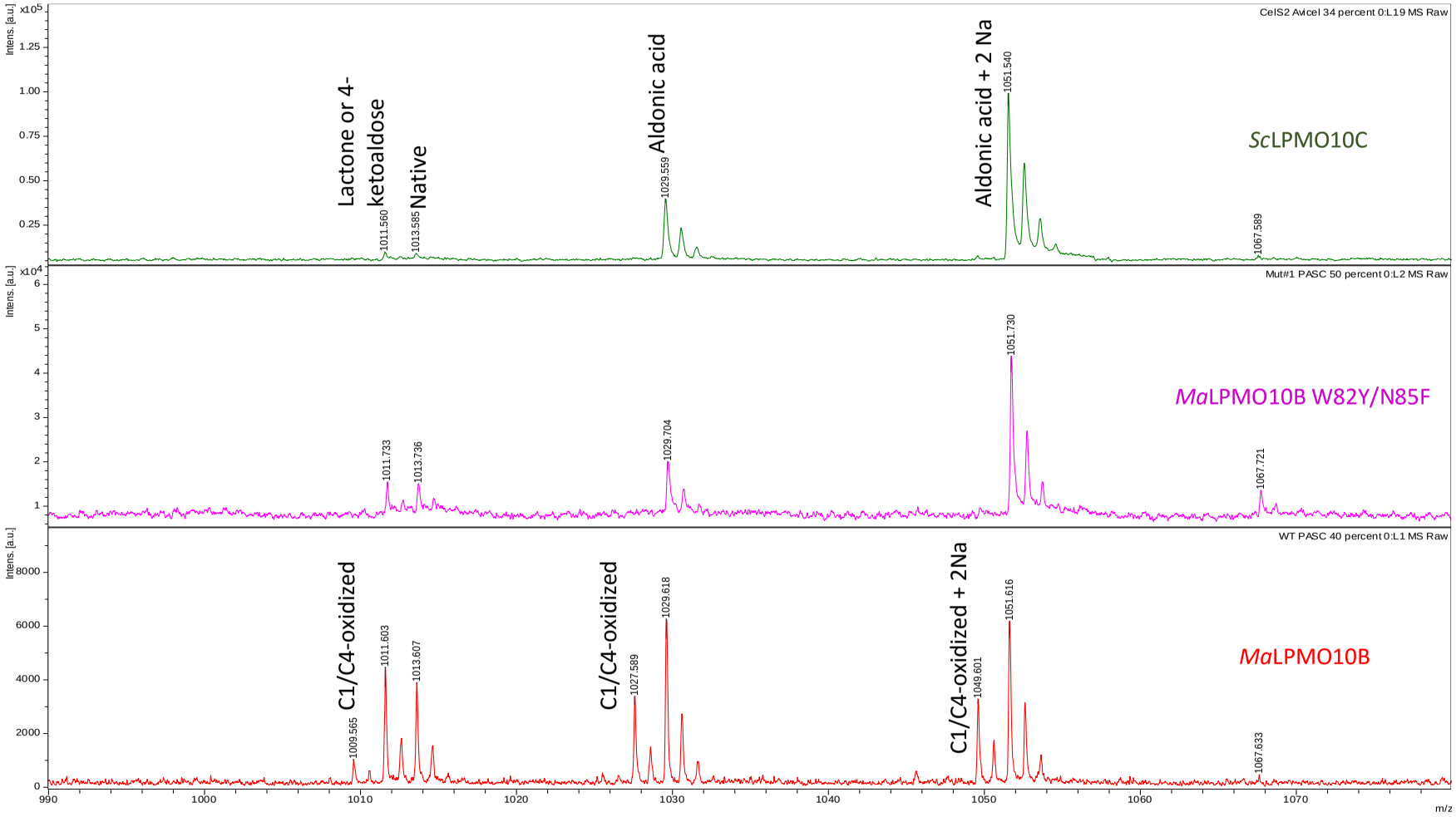


Figure F1. MALDI-TOF MS analysis of released oxidized cello-oligosaccharides from PASC by ScLPMO10C, MaLPMO10B and MaLPMO10B W82Y/N85F*. Overnight reactions were mixed with DHB in 1:2 ratios and air-dried prior to analysis by MALDI-TOF MS. The figure shows an enlargement of the DP6 ion cluster (highlighted in the top left corner) of each LPMO. Note the absence of C4/C1-oxidized products (labeled) in the MaLPMO10B W82Y/N85F specter.*

Appendix F

Bioinformatics analysis for selection of potential point mutations (performed by B. Bisarro)

A protein-protein BLAST (blastp) search was performed against non-redundant (nr) protein sequences of the NCBI database. The sequences of *Sc*LPMO10C (C1 oxidizer) and *Sc*LPMO10B (C1/C4 oxidizer) were used to create a pool of sequences of C1 or C1/C4 cellulose-active AA10 proteins (threshold identity values of 37 and 50 %). C1 and C1/C4 pools were compiled by selecting 27 and 28 sequences, respectively, from the blastp search results [the sequences corresponded to sequences identified by Book et al. (2014) in Clade II-B]. The sequence of the previously characterized *Hc*LPMO10-2 was manually added to the C1 pool, as the protein is known to be a C1 oxidizer although sequence identity is only 27 %. The aforementioned sequence identity threshold values were determined by making a phylogenetic tree and identifying characterized LPMOs of known oxidative regioselectivity as sequence boundaries. The sequences of *Tf*LPMO10B and *Cf*LPMO10C were selected as boundaries for the C1 and C1/C4 pools, respectively, and all sequences in between these and the query sequences (i.e. *Sc*LPMO10C or *Sc*LPMO10B) in the phylogenetic tree were selected to constitute each respective pool.

Using the T-Coffee Expresso online tool (Armougom et al., 2006), structural alignments of each representative dataset was performed to create a multiple-sequence alignment (MSA), adding the structures of *Sc*LPMO10C (PDB 4OY7) and *Sc*LPMO10B (PDB 4OY6) as structure input. The MSA was then used to create another phylogenetic tree (Figure F1) using PhyML [Phylogeny.fr; Dereeper et al. (2008)], and to perform the correlated mutation analysis (CMA). Correlation heat maps (Figure F4) and a correlation network (Figure F2) were created using the Comulator online tool [Bio-Predict 3DM; Kuipers et al. (2009)] which identified positions in the MSA with correlated mutations.

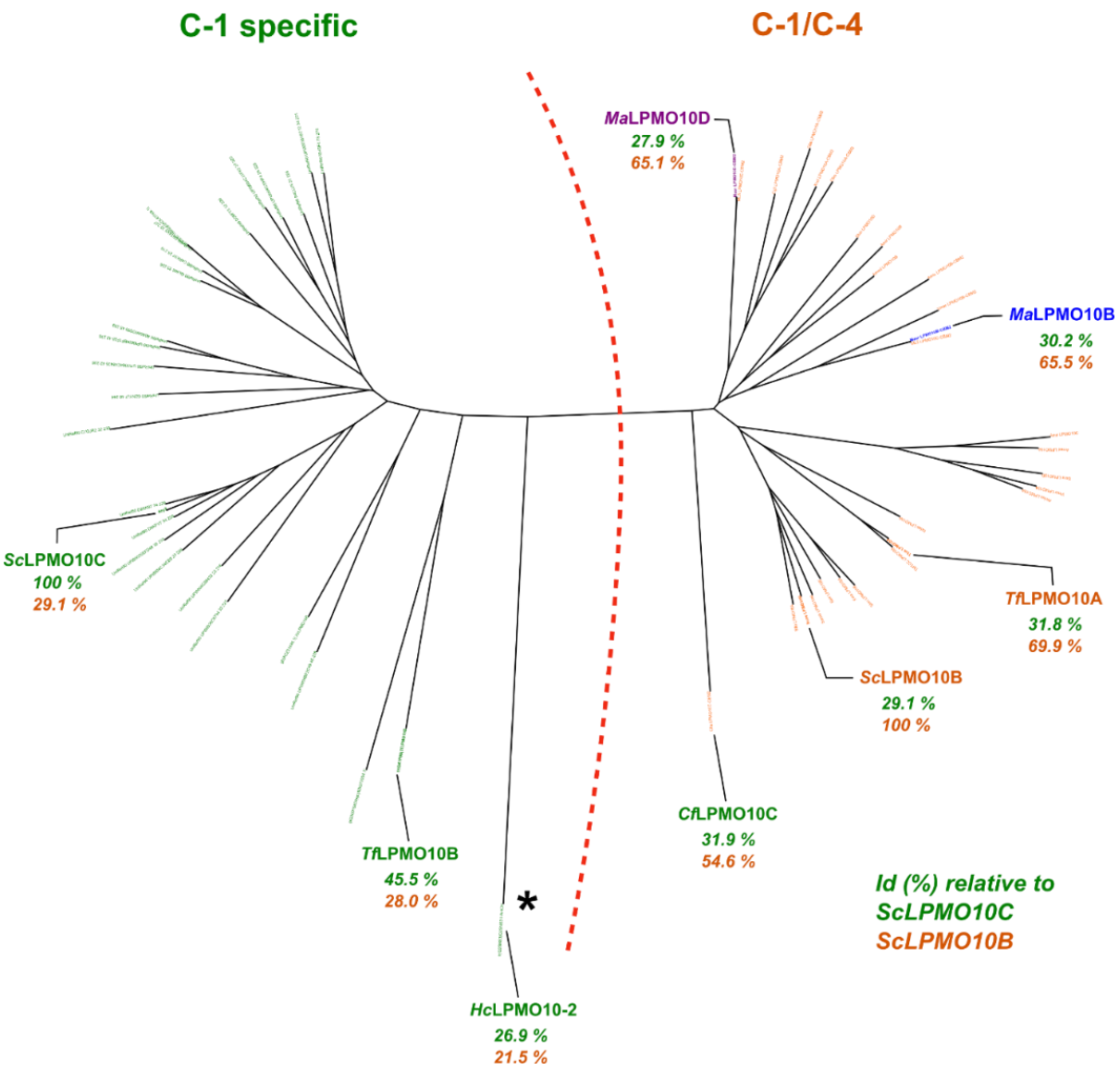


Figure F1. Phylogenetic tree illustrating the evolutionary divergence between cellulose-active AA10 proteins of C1 and C1/C4 oxidative regioselectivity. *PhyML was used to create the tree based on a MSA of the sequences of 27 and 28 C1 and C1/C4 oxidizing AA10 proteins, respectively. The iTOL online tool [http://itol.embl.de; Letunic and Bork (2011)] was used for visualization of the tree. The sequence identities relative to ScLPMO10C (green) and ScLPMO10B (orange) are shown on the figure, and are based on a comparison of the catalytic domain sequence of mature proteins. The previously characterized C1 oxidizing HcLPMO10-2 was manually added to the C1 sequence pool, and has a fairly low sequence identity compared with ScLPMO10C. This may be explained by the fact that all proteins in the tree, except HcLPMO10-2, are produced from Gram-negative bacteria, whereas HcLPMO10-2 is produced by the Gram-positive bacterium Hahella chejuensis. Figure source: B. Bisarro, unpublished work.*

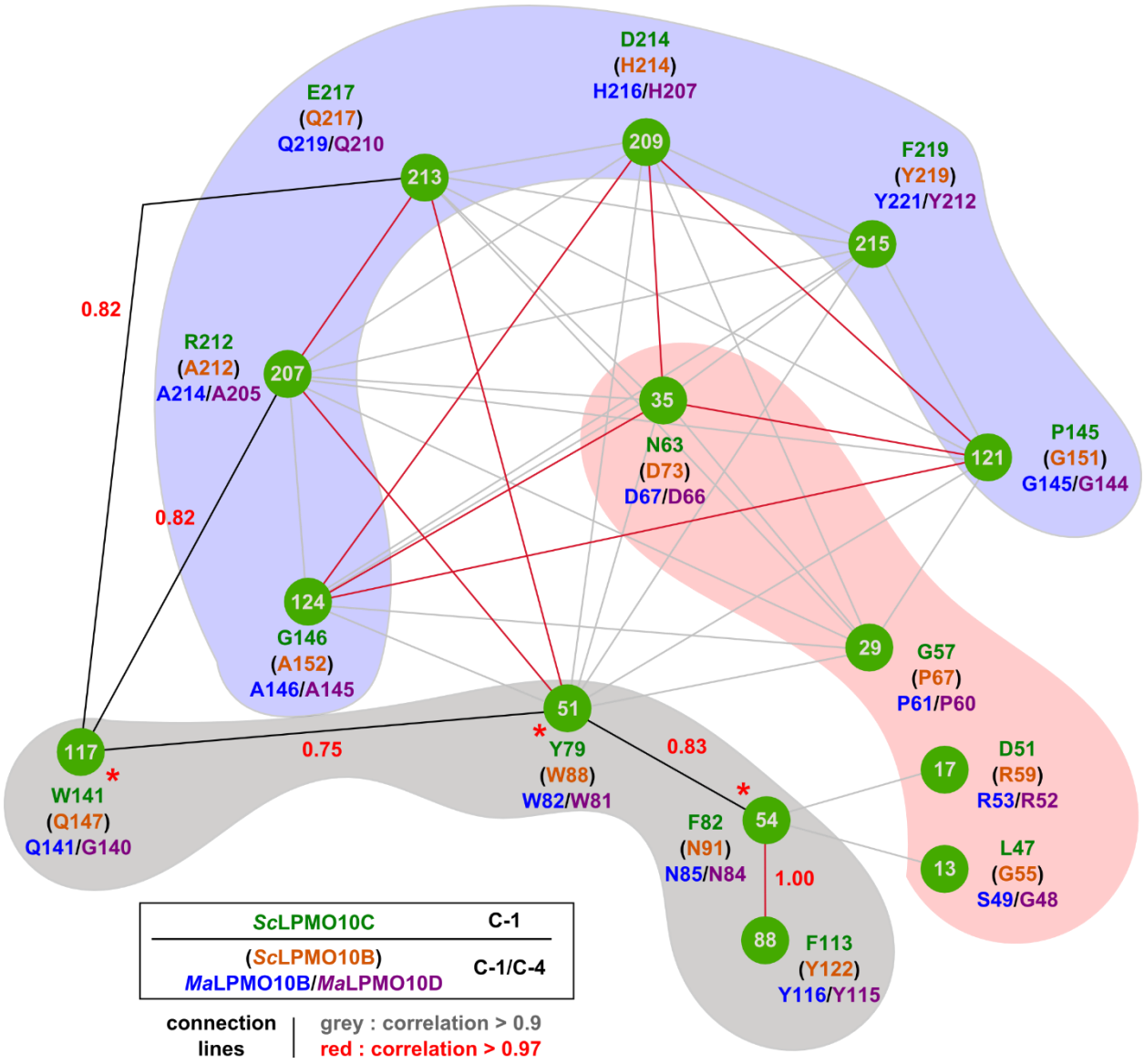


Figure F2. Correlated mutation analysis (CMA) of C1 and C1/C4 oxidizing cellulose-active AA10 proteins. *The position of each respective residue in the MSA is shown in the green circles. The amino acid found in the sequence of ScLPMO10C (green), ScLPMO10B (orange), MaLPMO10B (blue) or MaLPMO10D (purple) is shown below or next to each respective MSA position number. Correlations are illustrated by solid lines connecting various MSA numbers. Grey and red lines indicate correlation factors above 0.9 and 0.97, respectively. Based on structural analysis and previously performed mutagenesis studies, other relevant connections were added (black lines), even though their corresponding correlation factor is lower than 0.9. Residues that were targeted for site-directed mutagenesis (see section 3.4.3) are tagged with a red star. The colored zones represent three different clusters of residues that were identified after comparing all the pooled residues with the corresponding LPMO 3D-structures (using PyMOL; see Figure F3). Figure source: B. Bisarro, unpublished work.*

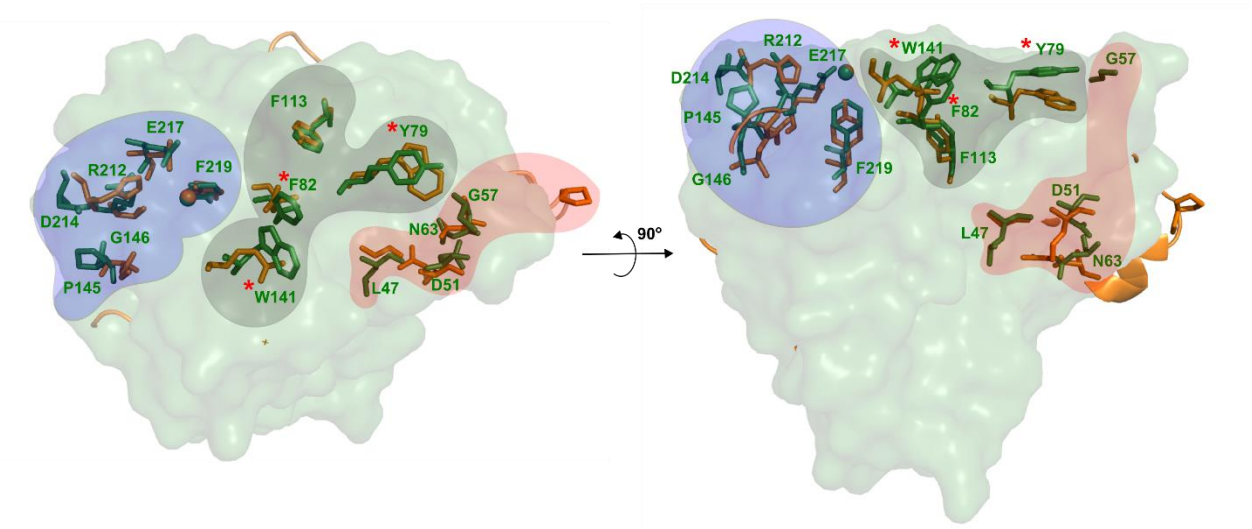


Figure F3. *Structures of ScLPMO10C (PDB 4OY7) and ScLPMO10B (4OY6) illustrating correlated residues from CMA analysis. Zones are shown with the same colors as were used in figure F2. Correlated residues are represented as sticks (ScLPMO10C, green; ScLPMO10B, orange) and residue numbering (green) are that of ScLPMO10C. Residues that were targeted for site-directed mutagenesis (see section 3.4.3) are tagged with a red star. Three clusters were identified: on the side of the copper center (blue zone); on the right side of the copper center (green zone); and a cluster belonging to a variable loop (red zone) located distant from the active site. Figure source: B. Bisarro, unpublished work.*

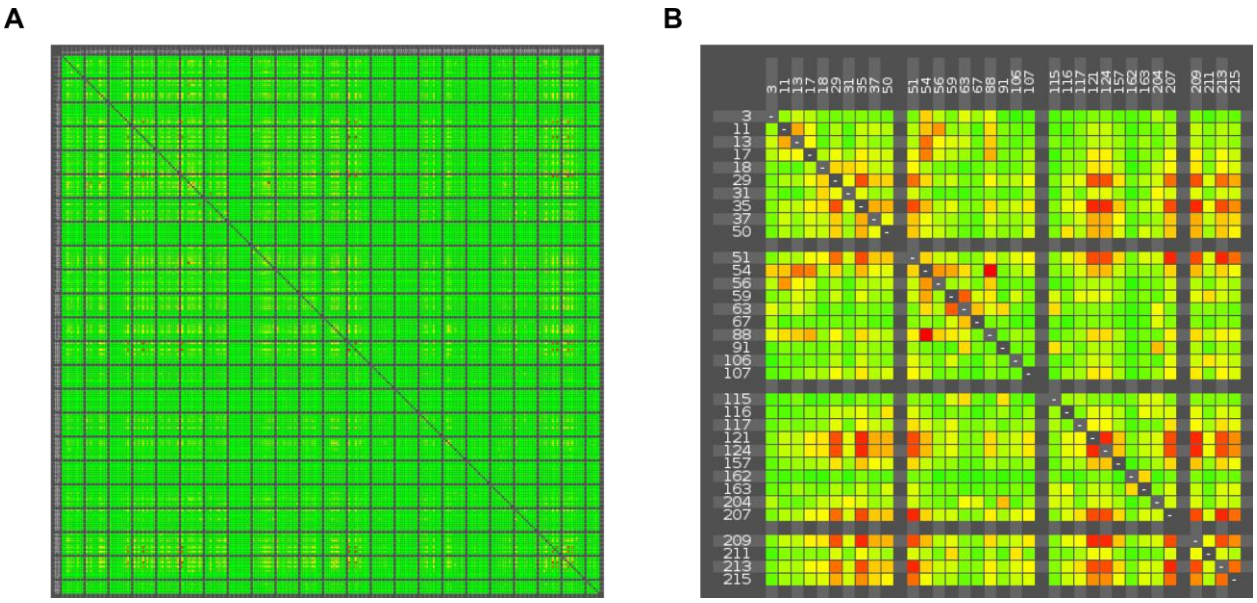
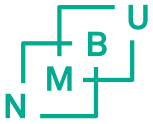


Figure F4. Heat map derived from correlated mutation analysis *(A). Full view and (B) condensed view displaying only highly correlated positions.*



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

Postboks 5003
NO-1432 Ås
Norway