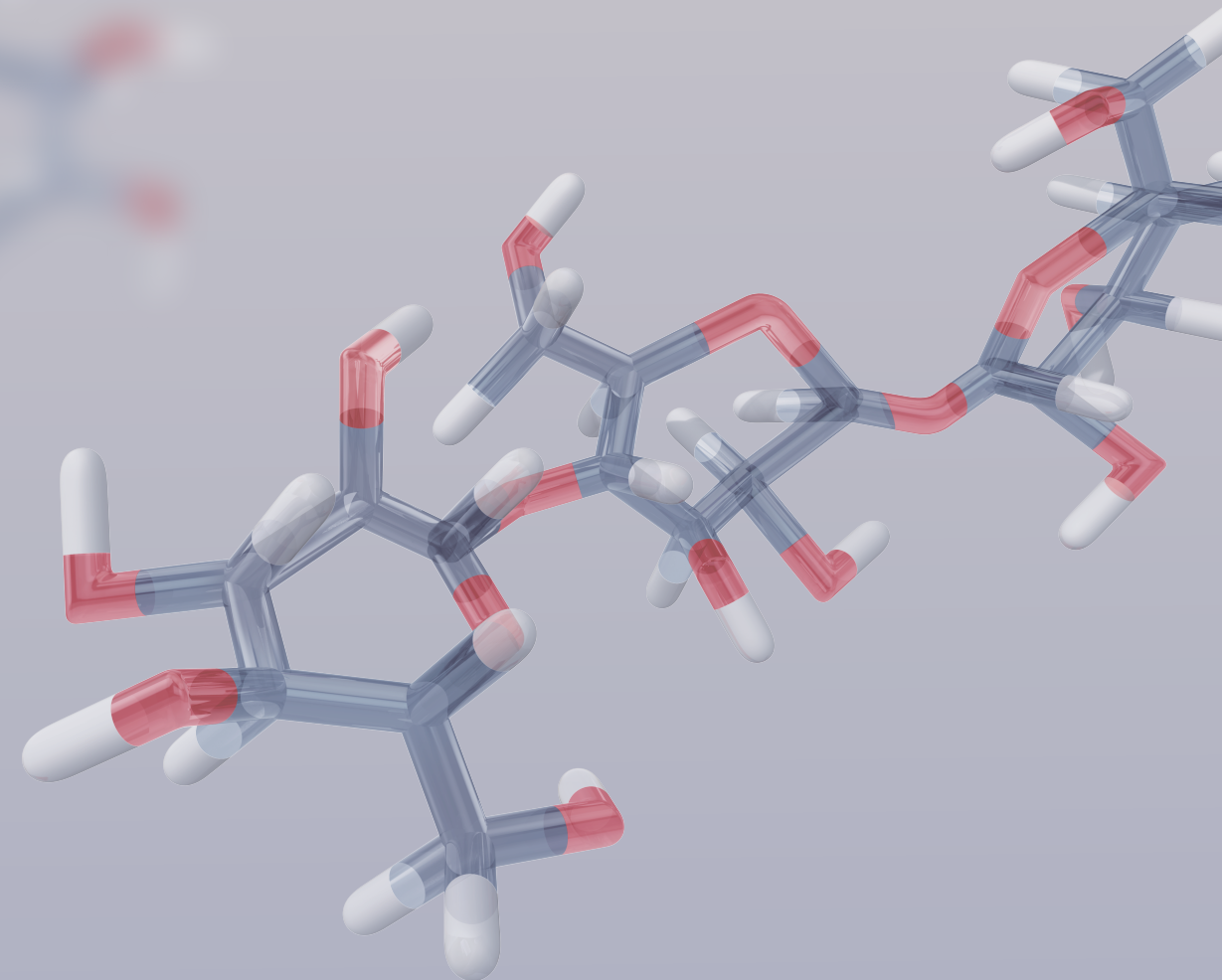
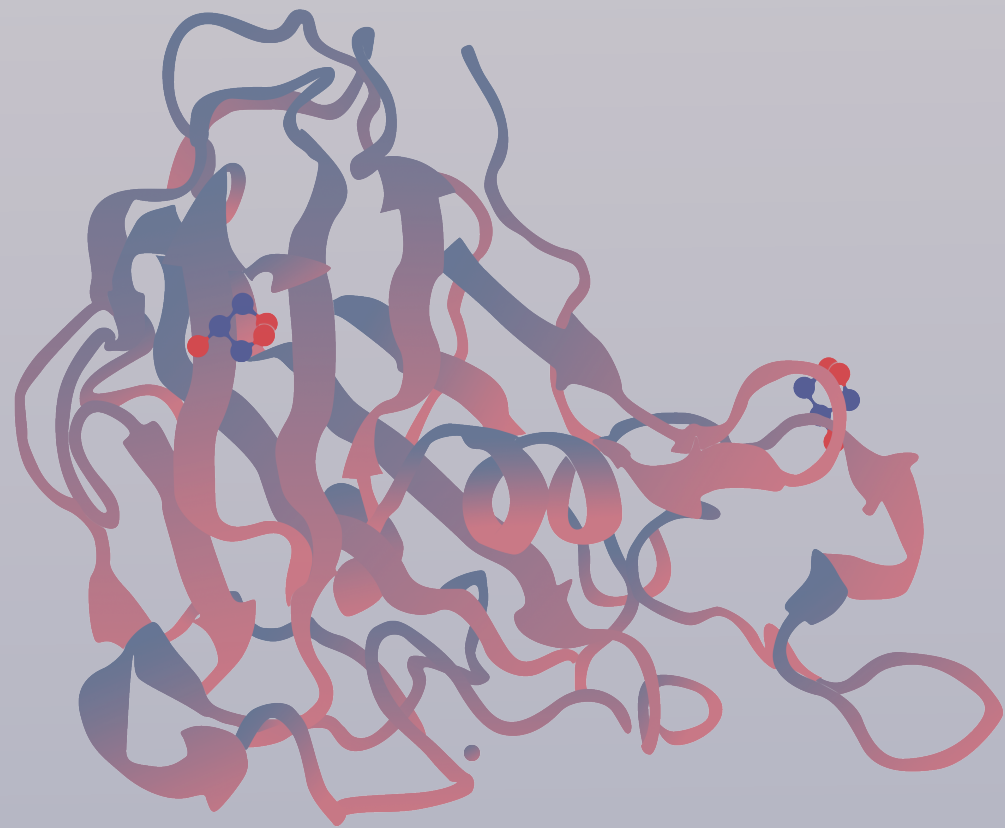


HIGH RESOLUTION PRODUCT PROFILING OF LYTIC POLYSACCHARIDE MONOOXYGENASES



PEICHENG SUN
孙培成

Propositions

1. C4-oxidizing LPMOs favor C6-hydroxylation over C6-aldehyde formation.
(this thesis)
2. A copper-histidine brace active site is not predictive for oxidative carbohydrate cleavage.
(this thesis)
3. The way to sustainable chemical engineering is a narrow path.
4. The function of "Schmutzdecke" (*dirty layer*) in slow sand filtration for drinking water is not reflected by its name.
5. Completion of a research is as complex as Chinese cooking.
6. Working from home increases productivity and efficiency.

Propositions belonging to the thesis, entitled:

High resolution product profiling of lytic polysaccharide monooxygenases

Peicheng Sun

Wageningen, 13 May 2022

High resolution product profiling of lytic polysaccharide monooxygenases

Peicheng Sun

Thesis committee

Promotors

Prof. Dr W.J.H. van Berkel
Personal Chair at the Laboratory of Food Chemistry
Wageningen University & Research

Dr M.A. Kabel
Associate Professor, Laboratory of Food Chemistry
Wageningen University & Research

Other members

Prof. Dr R. van Kranenburg, Wageningen University & Research
Prof. Dr K. Salomon Johansen, University of Copenhagen, Denmark
Prof. Dr M. Wuhler, Leiden University Medical Center (LUMC)
Prof. Dr D. Haltrich, University of Natural Resources and Life Sciences (BOKU), Austria

This research was conducted under the auspices of the Graduate School VLAG
(Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health
Sciences).

High resolution product profiling of lytic polysaccharide monooxygenases

Peicheng Sun

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 13 May 2022
at 1:30 p.m. in the Aula.

Peicheng Sun

High resolution product profiling of lytic polysaccharide monooxygenases
306 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summary in English

ISBN: 978-94-6447-101-4

DOI: <https://doi.org/10.18174/563451>

Abstract

Biorefinery of non-food lignocellulosic plant biomass to be valorized as biofuels, biochemicals and biomaterials is considered as a sustainable alternative for fossil resource-based production. In biorefinery, a key step is enzymatic degradation of plant cell wall polysaccharides into (fermentable) monomeric building blocks, which is typically driven by (hemi-)cellulases and essential lytic polysaccharide monooxygenases (LPMOs). Although some LPMOs have already been incorporated into industrial enzyme cocktails, only a limited number of LPMOs have been biochemically characterized. Therefore, this PhD thesis aimed to further shed light on the mode-of-action, regioselectivity, substrate cleavage profiles and specificity of AA9 LPMOs from the fungi *Myceliophthora thermophila* C1 (*Mt*LPMOs) and *Neurospora crassa* (*Nc*LPMOs).

We demonstrated that AA9 LPMOs show distinct cleavage profiles towards different types of cellulose; from bacterial cellulose mainly oxidized DP2–4 products were generated, while from the regenerated amorphous cellulose larger oxidized products prevailed. We developed a new method, making use of NaBD₄-reduction in combination with hydrophilic interaction chromatography–mass spectrometric analysis, for separation and identification of LPMO-generated non-, C1- and C4-oxidized oligosaccharides. We found that oxidized (and reduced) oligosaccharides differ in their mass spectrometric fragmentation behaviors and patterns. Using this developed methodology, we discovered a series of novel double, C4 and C6, oxidized cello-oligosaccharides. Again using this methodology, we identified and distinguished for the first time (isomeric) LPMO-oxidized xylogluco-oligosaccharides, and characterized two distinct oxidative xyloglucan degradation profiles. These xyloglucan substitution tolerant and intolerant cleavage profiles were correlated to specific active site segment configurations of AA9 LPMOs. Additionally, the active site segment configuration predicted for two (other) *Mt*LPMOs, confirmed the aforementioned correlation with their xyloglucan substitution (in)tolerant cleavage profiles. Lastly, we studied two novel AA16 copper-dependent oxidoreductases, which were predicted to function as LPMOs. We showed their inactivity towards carbohydrate substrates, hence, disproved them being LPMOs, but we discovered that these AA16s can boost AA9 *Mt*LPMOs to oxidatively degrade cellulose.

Table of contents

Chapter 1	General introduction	1
Chapter 2	Oxidized product profiles of AA9 lytic polysaccharide monooxygenases depend on the type of cellulose	31
Chapter 3	Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidized cello-oligosaccharides from their non-oxidized and reduced forms	59
Chapter 4	Regioselective C4 and C6 double oxidation of cellulose by lytic polysaccharide monooxygenases	87
Chapter 5	Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity	115
Chapter 6	Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation	165
Chapter 7	Extending the diversity of <i>Myceliophthora thermophila</i> lytic polysaccharide monooxygenases: two different xyloglucan cleavage profiles	211
Chapter 8	Copper-dependent AA16 oxidoreductases boost cellulose active AA9 lytic polysaccharide monooxygenases from <i>Myceliophthora thermophila</i> C1	235
Chapter 9	General discussion	263
Summary		291
Acknowledgements		295
About the author		301

Chapter 1

General introduction

1. Relevance of this research

With the rapidly growing world population and industrial development, the demand of energy, chemicals and materials have increased sharply in the last century (1, 2). Currently, this demand is largely met by exploiting fossil resources (3), although being responsible for release of greenhouse gasses, such as carbon dioxide, which has been recognized as one of the major factors causing the global climate changes (2, 3). Therefore, worldwide, a change in mindset is strived at, in which a circular economy is centralized. In this context, sustainable biorefineries are promoted, valorizing non-food lignocellulosic plant biomass and forestry side-streams into, particularly, chemicals and materials (1, 2). Eventual emitted carbon dioxide is recycled via assimilation to new plant biomass.

To overcome the natural resistance of plant biomass for its degradation in (fermentable) smaller building blocks needed for production of chemicals and materials, enzymes are applied in the biorefinery process. Enzymes are considered as green catalysts and many current industrial enzyme cocktails are rich in plant cell wall polysaccharide degrading hydrolases (4, 5). Since 2010, however, lytic polysaccharide monooxygenases (LPMOs) have been discovered boosting the activity of hydrolytic enzymes (6-8). Thus, addition of such LPMOs into industrial cocktails is expected to improve the saccharification efficiency and to decrease enzyme costs. Although in the last decade many new insights have led to a better understanding of how LPMOs operate (9-13), still, only few LPMOs have been biochemically characterized limiting their industrial implementation.

Therefore, this thesis aims to provide further insights into the mode-of-action, regioselectivity, substrate cleavage profiles and specificity of five (new) LPMOs from the fungus *Myceliophthora thermophila* C1 (MtLPMOs). Seven (other) MtLPMOs were already characterized in our laboratory (14). In addition, AA9 LPMOs from the fungus *Neurospora crassa* (NcLPMOs) were studied to strengthen the LPMO characterization portfolio.

2. Plant cell wall

2.1. Composition and architecture of plant cell wall

Plant cell walls provide plant tissues with strength and rigidity (15, 16), and are built of three main layers: a middle lamella, a primary cell wall and a secondary cell wall (**Fig. 1**) (17, 18). The most abundant cell wall components are polysaccharides including cellulose, hemicelluloses and pectin, the aromatic polymer lignin and structural proteins (15, 18, 19). The middle lamella is the outer layer and mainly composed of pectin (20, 21). The primary cell wall is typically a (relatively) thin, extensible and highly hydrated layer, which generally consists of 15–40% (w/w) cellulose, 30–50% (w/w) pectin and 20–30% (w/w) hemicellulosic xyloglucan and arabinoxylan (22, 23). In non-food lignocellulosic plant biomass and forestry side-streams, the largest part of the dry matter consists of secondary cell wall material and, thus, is the most relevant for biorefinery. Secondary cell walls are rich in cellulose,

hemicelluloses and lignin, which are intertwined in a complex network (**Fig. 1**) (24, 25). The exact composition of secondary cell walls depends on the botanical origin, species, age, place or year of harvest and part of the plant (e.g., root or leaf) (26). For example, secondary cell walls of grasses are generally composed of around 35–45% (w/w) cellulose, 20–30% (w/w) hemicellulosic xylan and 15–30% (w/w) lignin (24, 25).

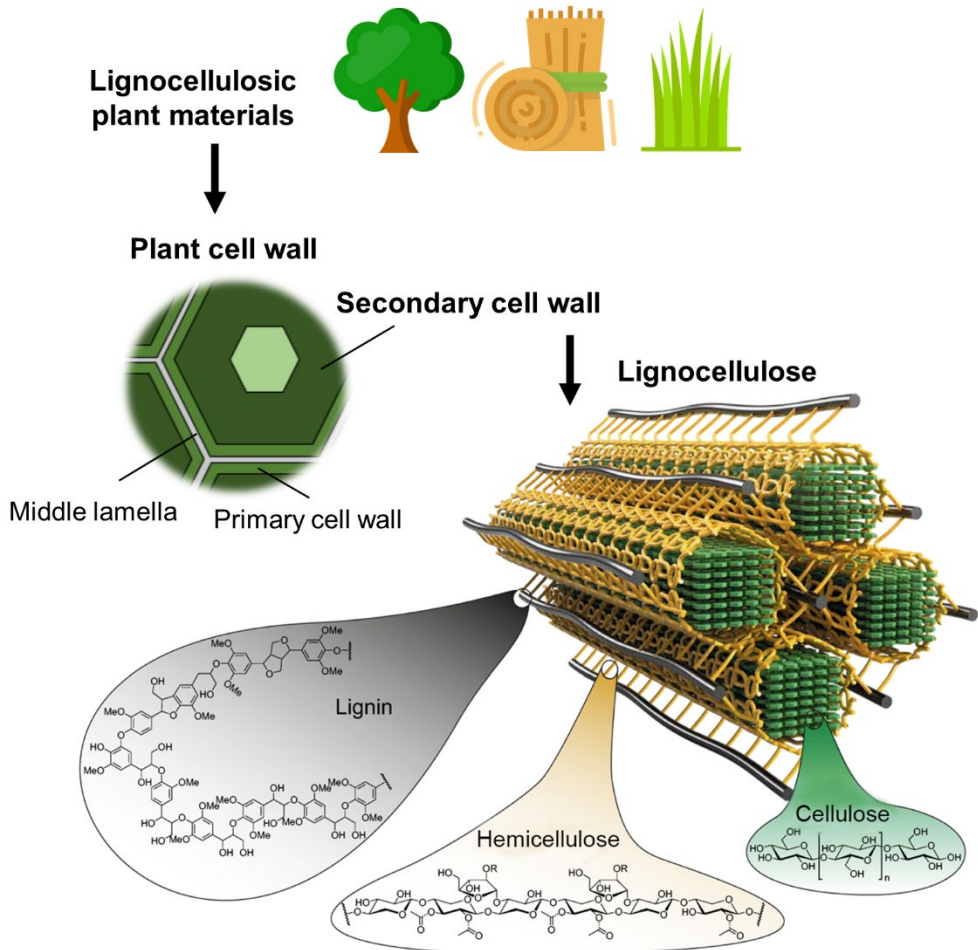


Fig. 1. Schematic representation of lignocellulosic plant biomass, the architecture of plant cell walls, and the structure and network of most relevant lignocellulose polymers (e.g., lignin, hemicellulosic xylan and cellulose). Adapted from Brethauer et al. (27).

2.2. Cellulose

Cellulose is referred to as the most abundant renewable macromolecule on Earth (28). In addition to be part of cell walls in vascular plant, cellulose is synthesized in most groups of algae and by a number of bacterial species (29). Overall, cellulose is a homogenous linear polysaccharide built of β -(1 \rightarrow 4)-linked glucosyl residues (**Fig. 2**).

The length of these glucan chains varies per plant source, origin and location in the cell wall (22, 29-32). For instance, in the plant cell wall, it has been observed that each cellulose glucan chain contains approximately 8000–14000 glucosyl residues (22, 30). In secondary cell walls, approximately 30 to 36 glucan chains of cellulose are assembled and twisted to microfibrils via van der Waals forces and hydrogen bonds (30, 33, 34). Especially the extensively present hydrogen bonds amongst oxygen atoms (O2, O3, O5 and O6) within and between glucan chains give the microfibrils their ordered and crystalline structure (33-37).

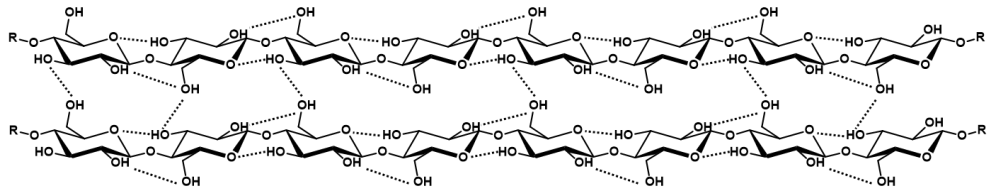


Fig. 2. Chemical structure glucan chains intertwined in fibrils. Dotted lines indicate hydrogen bonds between glucosyl units.

Moreover, these cellulose microfibrils are formed by parallelly stacked and side-by-side packed glucan chains, which have been defined as cellulose I (29, 38). Cellulose I is further divided into two sub-allomorphs, namely I_α and I_β , differing in the alternation of glucosyl units, the way of hydrogen bonds formed, and the direction of glucan chains arranged (29, 38). I_β has been predominately found in plant cell walls (34, 36, 39). Apart from the crystalline microfibrils, cellulose also comprises amorphous regions, due to less ordered and inconsistently packed glucan chains (32).

For different applications, native cellulose has been modified to obtain optimal functional properties (31). For example, different alkaline or liquid ammonia extraction conditions change the I_β structure of native cellulose to different allomorphs of type II, III and IV (31, 39-41), which is more susceptible for enzyme degradation. In addition, partial acid hydrolysis of native cellulose generates so-called microcrystalline cellulose, which is widely used as cellulose model for studying, for example, activity and catalytic performance of cellulases (42-46). The microcrystalline cellulose differs to native cellulose, for instance, having a shorter chain length (degree of polymerization (DP) 100–300) and a smaller particle size (30–50 μm) (43, 46-53). Microcrystalline cellulose can be further modified to prepare other types of cellulose, such as phosphoric acid swollen cellulose and regenerated amorphous cellulose (42, 54-56).

2.3. Hemicellulose

Hemicellulose is traditionally referred to as a group of heterogeneous matrix polysaccharides being associated with cellulose in plant cell walls (57). Hemicelluloses comprise polysaccharides varying in types of building block and having various side-chains, which differ amongst botanical types and tissues (57, 58). For instance, grasses, cereals and hardwoods comprise mainly xylans (up to 50% (w/w)), while softwood hemicelluloses largely consist of glucomannans (57, 59).

2.3.1. Xylan

Xylans have a backbone of β -(1 \rightarrow 4)-linked xylosyl residues. Hardwoods mainly consist of glucuronoxylan, which has a xylan backbone with *O*-2 substituted 4-*O*-methyl-glucuronosyl units (57). In cereals and grasses, xylan backbones are substituted by arabinosyl residues at either the *O*-2, *O*-3, or both positions, and, to a lesser extent, by glucuronosyl or 4-*O*-methyl-glucuronosyl residues at the *O*-2 positions (59-62). These xylans are often referred to as glucuronoarabinoxylans (**Fig. 3**). In addition, glucuronoarabinoxylans can further contain acetyl groups at the *O*-2, *O*-3 or both positions of the xylosyl backbone (57, 62). Softwoods contain similar types of xylan as in grasses, though the amount and pattern of substitution are different (57).

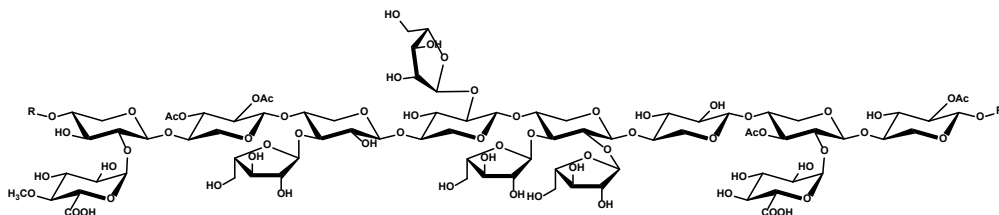


Fig. 3. Representation of a typical glucuronoarabinoxylan structure. Ac = acetyl group; R = continuation of the xylan structure.

2.3.2. Xyloglucan

Xyloglucan is a major hemicellulosic component in primary cell walls (57, 63, 64). Xyloglucan has a cellulose-like β -(1 \rightarrow 4)-linked glucosyl backbone, which is substituted by xylosyl residues via α -(1 \rightarrow 4)-linkages (65). Further substitution is also observed, such as by β -(1 \rightarrow 2)-linked galactosyl units and, rarely, by α -(1 \rightarrow 3)-linked arabinosyl residues (66). In addition, the galactosyl units can be further substituted by fucosyl residues via α -(1 \rightarrow 2) linkages (67). The exact type and amount of substitution varies for different plant species and tissues (65). These complicated trivial names of xyloglucan substitution have been simplified by an unambiguous “one-letter code” nomenclature developed by Fry et al. (66). As shown in **Fig. 4**, the glucosyl, glucosyl-xylosyl, glucosyl-xylosyl-galactosyl, and glucosyl-xylosyl-galactosyl-fucosyl residues are referred to as G, X, L and F, respectively. Other, less frequently found types of substitution with their corresponding codes are not described here, but are reviewed by Park et al. (68).

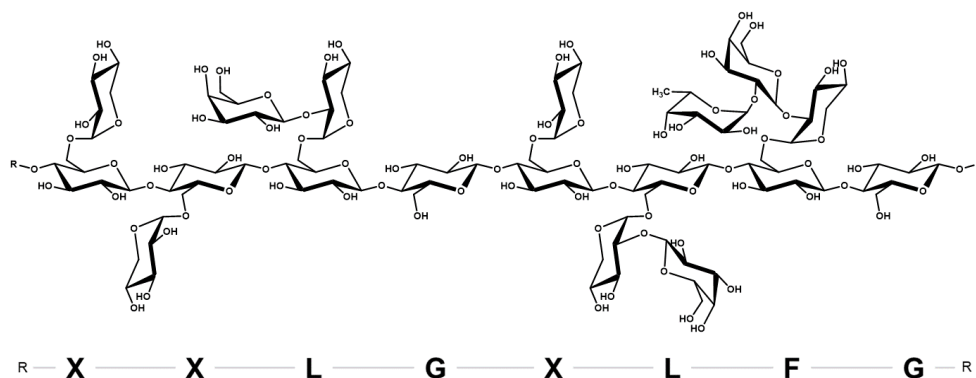


Fig. 4. Representation of a typical xyloglucan structure (e.g., from black currant berries), and the corresponding “one-letter” code of each structural unit according to Fry et al. (66). R = continuation of the xyloglucan structure.

In addition, xyloglucan can vary in the type of building blocks. Most xyloglucan structures have been defined as “XXXG”- (e.g., tamarind seed) and “XXGG”-type (e.g., potato) (69, 70). To be more specific, “XXXG”-types also includes partial galactosyl substitution, such as in tamarind xyloglucan (XXLG, XLXG and XLGL) (70). Other types of building blocks, such as “XXG”- (e.g., sycamore (71)), “XGGG”- (e.g., tobacco (72)), “XXGGG”- (e.g., barely (73)) and “XXXX”- (e.g., *Helipterum eximium* (74)) have been proposed, but studied to a much lesser extent.

2.3.3. Other hemicelluloses

The simplest mannan structure is composed of a β -(1 \rightarrow 4)-linked mannosyl backbone. Glucomannan consists of a linear backbone of β -(1 \rightarrow 4)-linked mannosyl and glucosyl residues (75-77). Glucomannan can be substituted by galactosyl units at the O-6 position, subsequently named galactoglucomannan (78). In addition, O-acetylation has been found on glucomannan and galactoglucomannan in some plant species (79, 80). Other hemicelluloses, such as mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan, (arabino-)galactan, arabinan and laminarin, are not described further in this section. Their composition, structure, building units and property can be found elsewhere (12, 57-59, 61, 81-83).

2.3.4. Hemicellulose interaction with other cell wall components

In plant cell walls, hemicelluloses, especially the unbranched region, have been reported to cross-link to each other and to be associated with cellulose via non-covalent interactions (e.g., hydrogen bonding) (58, 82-84). In grasses, xylans adsorb to cellulose to form a xylan-cellulose network (61, 62, 85) (**Fig. 1**). Similarly, mannan, glucomannan and galactomannan have been shown to adsorb to cellulose in the secondary cell wall of softwoods (78, 86). Xyloglucan also has been described to bind to cellulose (microfibrils) in the primary cell wall (68, 87). Notably, several studies suggested that xyloglucan may interact with cell wall components via covalent bonds, in addition to hydrogen bonding, though it is not known if this also applies to its interaction with cellulose (64, 65, 88, 89).

Apart from their association with cellulose, some hemicelluloses are known to covalently interact with the aromatic heterogeneous polymer lignin (**Fig. 1**). Lignin is an aromatic polymer built up from oxidative coupling of 4-hydroxyphenylpropanoids (90). In softwoods (e.g., Japanese red pine), some mannosyl residues have been elucidated to link to lignin via an ether bond (91). In grasses, glucuronoarabinoxylan links to lignin indirectly via ferulate and diferulate “bridges” (92-94). The above described hemicellulose-lignin structures are referred to as lignin-carbohydrate complexes (LCC) (95), which will not be further discussed in this thesis. Hemicellulose and lignin are cross-linked and intertwined via these non-covalent and covalent interactions, which further entwines on the crystalline cellulose fibrils to build up a three-dimensional network (**Fig. 1**).

3. Enzymatic degradation of plant cell wall polysaccharides

3.1. Classification of carbohydrate-active enzymes

As described above, the lignocellulosic plant biomass has a complex and rigid structure. Therefore, the conversion of plant polysaccharides to their corresponding monomeric sugar units requires a wide range of different enzymes. With the increasing number of identified carbohydrate degrading enzymes at gene level, classification of these enzymes became essential, which has led to the establishment of the Carbohydrate-Active enZymes (CAZy) database (<http://www.cazy.org>) (96). CAZy recognizes glycoside hydrolase (GH) families, which includes (hemi-)cellulases, in addition to polysaccharide lyases (PL), carbohydrate esterases (CE), glycosyltransferases (GT), auxiliary activities (AA) and carbohydrate binding module (CBM) families. Although this CAZy classification system is based on amino acid sequence identity, and not on function similarity, it still provides some prediction of general carbohydrate substrate specificities (96). Initially, lignin- and other non-carbohydrate-active enzymes that potentially assist carbohydrate degradation were classified into AA families (96). Later, other carbohydrate-active redox enzymes, such as LPMOs and cellobiose dehydrogenases (CDHs), were added to these AA families (97).

In the following text only fungal enzymes are discussed, as fungal LPMOs were studied in this thesis.

3.2. Cellulose degrading enzymes

The classical model for cellulose degradation defines the catalytic action of three different groups of hydrolases: endoglucanases (endo- β -1,4-glucanases, EGs), cellobiohydrolases (CBHs) and β -1,4-glucosidases (BGs) (**Fig. 5**) (5, 98, 99). EGs cleave the internal β -(1 \rightarrow 4)-linked glycosidic bonds in cellulose, and fungal members are found in GH5, GH7, GH12 and GH45. CBHs release cellobiose from β -(1 \rightarrow 4)-linked glucan chain from either the reducing (CBH I) or non-reducing (CBH II) ends (100). Fungal CBH I and CBH II genes are classified in GH7 and GH6, respectively (100). BGs hydrolyze soluble cello-oligosaccharides into glucose from the non-reducing end. Fungal BGs are mostly found in GH1 and GH3 (100). Notably, this classification of cellulose degrading enzymes into three groups is oversimplified. It has been shown

that EGs can be further distinguished to cleave cellulose in a random or processive mode, and CBHs also differ in their cleavage towards cellulose from strictly exo- to a highly processive action (5, 98-107). This might explain that usually multiple EGs, CBHs and BGs are present in fungal genomes, which can be produced to work synergistically in degrading cellulose (5, 108).

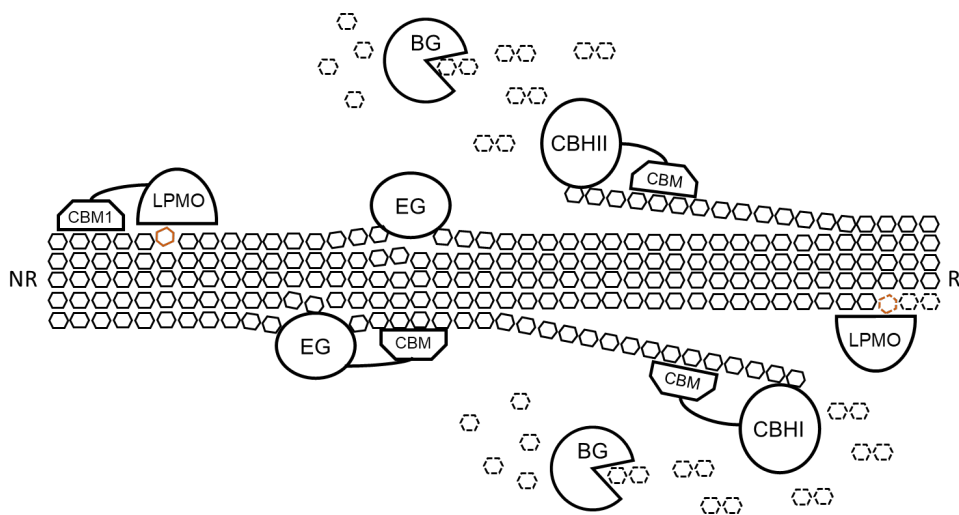


Fig. 5. Current schematic model of fungal enzymatic cellulose degradation adapted from Horn et al. (6). The mode-of-action of EG (endoglucanase), CBH (cellobiohydrolase) and BG (β -1,4-glucosidase) are described in the main text. Fungal LPMOs can cleave cellulose and release cello-oligosaccharides with either oxidized reducing end (R) or non-reducing end (NR) (orange hexagons). Full line hexagons: insoluble glucosyl units; Dotted line hexagons: soluble glucosyl units.

More recent, other (non-hydrolytic) enzymes or proteins involved in cellulose degradation, such as expansins, swollenins and LPMOs, have been categorized as Non-Hydrolytic Cellulose Active Proteins (NHCAPs) (109). In particular, LPMOs abundantly present in (hemi-)cellulose degrading Ascomycetes and Basidiomycetes (110, 111) have been shown to boost the cellulose degradation. This LPMO-boosting effect is suggested to relate to the formation of more chain ends and disrupting crystalline cellulose regions, which increases the substrate availability and accessibility for EGs and CBHs (6, 13, 111). Therefore, LPMOs cannot be ignored in our current understanding of cellulose degradation (**Fig. 5**), and the mode-of-action of these unique enzymes are further explained in Section 4.1 and 4.2.

3.3. Hemicellulose degrading enzymes

Hemicelluloses highly differ in their types of backbone and substitution, and thus, their degradation requires various enzymes having different substrate specificities. In this section, xylan and xyloglucan degrading enzymes are described. Other hemicellulose degrading enzymes can be found elsewhere (112).

3.3.1. Xylan degrading enzymes

The xylan backbone can be cleaved by endo- β -1,4-xylanases (e.g., GH10, GH11 and GH30), and the resulting xylo-oligosaccharides are further hydrolyzed by β -xylosidases (e.g. GH3, GH43 or GH54) to xylose (113-115). Certain endo- β -1,4-xylanases are more tolerant to degrade a substituted xylan backbone (e.g., GH10; see **Fig. 3**), while others are less tolerant for substitutions (e.g., GH11) (116, 117). In addition, fungi also produce a variety of xylan-debranching enzymes, such as α -arabinofuranosidases (e.g., GH43, GH51, GH54, etc.), α -glucuronidases (e.g., GH67 and GH115), and acetylxylan esterases (Carbohydrate Esterase (CE1)) (108, 118-120). Their detailed mode-of-action and the resulting product profiles can be found elsewhere (112).

So far only one fungal LPMO (*LsAA9A* from *Lentinus similis*) has been shown to cleave the xylan backbone when only xylan is present (121). Some fungal LPMOs were reported to cleave the xylan backbone only when associated with cellulose (122-124). The LPMO activity towards xylan is further discussed in **Chapter 9** of this thesis.

3.3.2. Xyloglucan degrading enzymes

The β -(1 \rightarrow 4)-linked-xyloglucan backbone can be cleaved by endoglucanases, of which fungal candidates are found in GH5, GH12, GH16 and GH74 (125). Notably, endoglucanases from GH12 and GH74 have been characterized as xyloglucan specific enzymes, which did not show activity towards other substrates such as cellulose or β -glucan (126, 127). These xyloglucan specific endoglucanases have been specified as xyloglucanases (XEGs) (127-132). XEGs from GH12 display an endo-active cleavage of the xyloglucan backbone, hydrolyzing the glycosidic linkages at the "reducing end" site of unsubstituted glucosyl units, releasing "XXXG-type" oligosaccharides (133, 134). XEGs from GH74 have a similar endo-type of action as the GH12 XEGs, but follow a processive-behavior (128, 131, 135, 136). In addition, some GH74 XEGs are able to cleave within the xyloglucan building blocks XXXG and XXLG to generate XX, XG and LG units. However, GH74 XEGs do not cleave at the "reducing end" site of L units in XLXG and XLLG (131, 135-138). Furthermore, β -glucosidases (e.g., GH1 and GH3) can cleave terminal unsubstituted glucosyl units from xylogluco-oligosaccharides. These β -glucosidases are often more active after treatment with xyloglucan-debranching enzymes such as α -xylosidases (GH31), α -galactosidases (GH27 and GH36) and α -fucosidases (GH95) (108).

Several fungal AA9 LPMOs have been reported to cleave within xyloglucan backbones (12), which is further discussed in Section 4.5.

4. Lytic polysaccharide monooxygenases (LPMOs)

Lytic polysaccharide monooxygenases (LPMOs) are introduced here based on studies published until the early 2018, the time when this project started. Recently published literature (2018-2021) is discussed in **Chapter 9**, together with the major findings from this project.

4.1. Discovery and current classification of LPMOs

Already back in 1950, Reese and coworkers have suggested that some unknown non-hydrolytic proteins may play a role in cellulose degradation in addition to hydrolytic enzymes (139). In 2005, it was reported that a protein (CBP21) produced by a bacterium could enhance the hydrolytic degradation of chitin, which is a cellulose-like crystalline polysaccharide built of β -(1 \rightarrow 4)-linked *N*-acetylglucosamine residues (140). Later, this protein class was added into the CAZy database as CBM33 family (141). Putative CBM33 candidates are widely distributed in genomes of bacteria and viruses, but absent in fungi. Instead, fungi produce another type of proteins, which was classified into the GH61 family, having a similar structure as CBM33 proteins (142, 143). At that time, it was shown that supplementation of GH61 members to cellulolytic enzyme cocktails increased cellulose degradation, although activity of sole GH61 proteins towards cellulosic substrates was not detected (6, 142, 143). Until 2010, the catalytic role and function of CBM33 and GH61 proteins was still unknown, even after the first structure of GH61 protein (Cel61A from *Trichoderma reesei*) was published (142). Finally, in 2010, Vaaje-Kolstad and coworkers demonstrated for the first time that a CBM33 protein cleaves the glycosidic bonds in chitin via oxidative mechanisms using molecular oxygen (O₂) and external electron donors (8). Subsequently, both CBM33 and GH61 proteins were demonstrated to oxidatively cleave cellulose (7, 144, 145), and renamed as polysaccharide monooxygenases (PMOs) or lytic polysaccharide monooxygenases (LPMOs). In 2013, the GH61 and CBM33 were reclassified as AA9 and AA10, respectively (96, 146).

Subsequently, new LPMO families were identified (e.g., AA11 and AA13). AA11 LPMOs are found in fungi (except one putative protein from bacteria), and characterized to be only active towards chitin (147). AA13 members are only found in fungi, and characterized to be starch-active LPMOs (148). In the beginning of 2018, AA14 LPMOs were identified in fungi, showing the activity towards xylan, but only when associated with cellulose (149). One month later, AA15 LPMOs were discovered from the genomes of marine and terrestrial invertebrates, fungi, algae and oomycetes (150). AA15 LPMOs were shown to be active towards both cellulose and chitin (150). The newest AA16 has also been described as LPMO family, of which one protein was reported to cleave cellulose (151). In the rest of section, the information focuses on the so far largest AA9 LPMOs. Other classes of LPMOs will also be included due to the similarity in structural features and catalytic mechanisms.

4.2. Structure of LPMO catalytic domains

The first structures of an AA9 (PDB ID: 2VTC) and an AA10 LPMO (PDB ID: 2BEM) have been elucidated in 2008 and 2005, respectively (142, 152). Subsequently, more structures of AA9 and AA10, as well as of AA11, AA13, AA14 and AA15 LPMOs, have been elucidated as extensively reviewed previously (13, 146, 149, 150, 153, 154). Overall, though generally LPMOs exhibit a low amino acid sequence similarity, both in between and within families, they display a common fibronectin-/immunoglobulin-like β -sandwich core structure (**Fig. 6**) (153, 154). Different from GH enzymes, which has

a surface cleft or groove, the LPMO active site is located at a relatively flat, solvent (and substrate) exposed, surface (**Fig. 6**) (146, 149, 150, 155-157).

4.2.1. Conserved mono-copper active site and unique “histidine-brace” motif

LPMOs have a single copper atom in the active site, which is coordinated by a N-terminal histidine (His1), a second histidine and a third aromatic amino acid (**Fig. 6**). In AA9, AA11, AA13 and AA14 LPMOs, the third amino acid is a tyrosine, while a phenylalanine is found in AA10 (90% conserved) and chitin-specific AA15 LPMOs (149, 150, 154). This structural coordination is recognized as a histidine-brace motif, which is conserved in all LPMOs to date (145, 154, 158, 159). The histidine-brace motif has been shown to be key for the LPMO catalytic reaction. It has been reported that mutation of either His1 or the second His led to a complete loss of catalytic action in a AA9 LPMO (GH61E from *Thielavia terrestris*) (143). Although mutation of the tyrosine in the “histidine-brace” motif of GH61E did not result in complete inactivation, its activity decreased sharply (143). In addition, the N-terminal His1 has been shown to be methylated when LPMOs are homologously produced in fungal hosts (145, 157). However, this unusual post-translational modification was not observed when LPMOs are produced in bacteria or heterologous hosts such as the methylotrophic yeast *Pichia pastoris* (160-162). Nevertheless, LPMOs with or without the N-terminal methylation show similar activity (161, 162), and the function of this modification remains unclear.

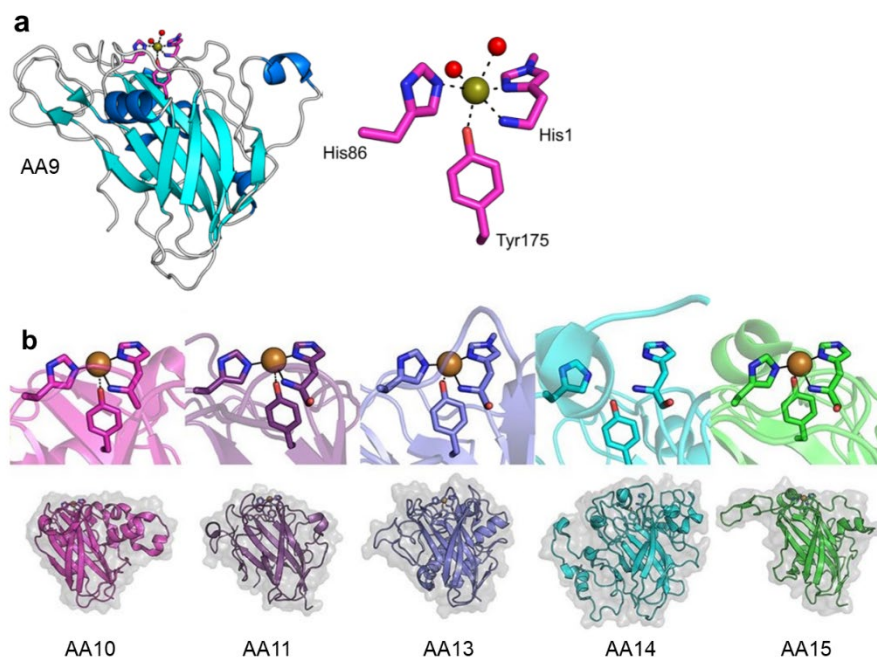


Fig. 6. (a) Example of an AA9 LPMO three-dimensional structure (*TaGH61A*, PDB ID: 2YET)(145) and a close look of the active site. (b) Examples of three-dimensional structures and active sites from LPMOs of AA10 (PDB ID: 5OPF), AA11 (PDB ID: 4MAI), AA13 (PDB ID: 4OPB), AA14 (PDB ID: 5NO7) and AA15 (PDB ID: 5MSZ). Adapted from Eijnsink et al. (11) and Tandrup et al. (158).

4.2.2. Diverse configuration of substrate binding sites

Structural elements around the active site have been shown to be diverse. These structural elements consist of α -helices and loops connected the core β -sandwich structure, and are suggested to drive the substrate binding (142, 143, 153, 157). For early published AA9 LPMO structures, Li et al. (157) noticed that a loop named L2 is the most diverse region comprising different numbers and types of amino acid, and different secondary structures (**Fig. 7a**). In the same study, PMO-3 from *N. crassa* (NcLPMO9M) was found to have an extended L2 loop (**Fig. 7b**), while PMO-2 (NcLPMO9C) was shown to obtain a shorter L2 loop (**Fig. 7b**) (157). In addition, *Hj*Cel61B (142) and *Ta*GH61A (145) were characterized to have a similar L2 configuration as for NcLPMO9M, and *Tt*GH61E (143) was shown to lack the L2 loop as for NcLPMO9C.

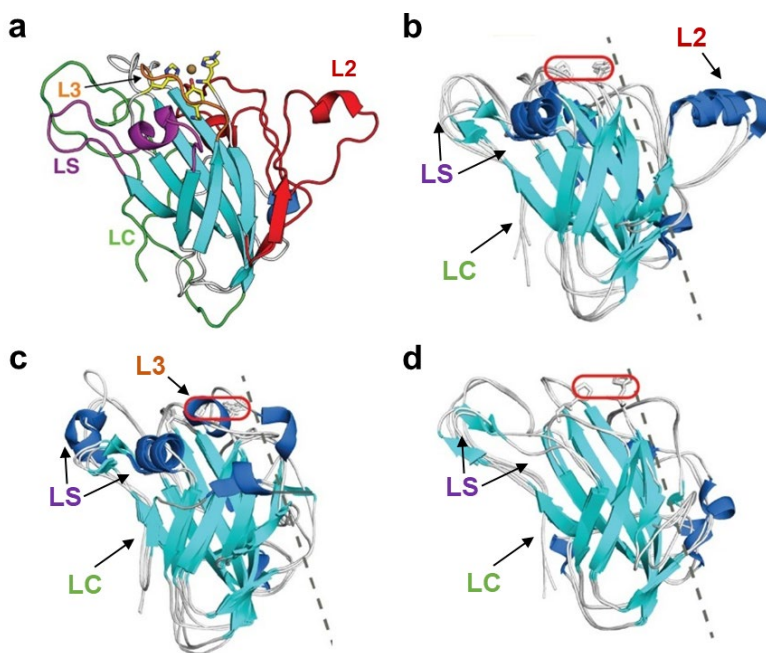


Fig. 7. (a) Schematic illustration of AA9 LPMO loop configuration. Note: no AA9 LPMO has been found to include all four loops so far. (b) Overlapped three dimensional structures of *Hj*LPMO9B (PDB ID: 2VTC), *Ta*LPMO9A (PDB ID: 2YET) and *Nc*LPMO9M (PDB ID: 4EIS). (c) Overlapped three dimensional structures of *Nc*LPMO9C (PDB ID: 4D7U), *Nc*LPMO9D (PDB ID: 4EIR) and *Ls*LPMO9A (PDB ID: 5ACF). (d) Overlapped three dimensional structures of *Tt*LPMO9E (PDB ID: 3EII), *Nc*LPMO9F (PDB ID: 4QI8) and *Pc*LPMO9D (PDB ID: 4B5Q). The active site is indicated by a red oval. The core β -sandwich and the L2 loop is separated by the dashed line. Adapted from Vaaje-Kolstad et al. (153).

The L2 loop has been found to contain several surface located aromatic amino acids (e.g., tyrosine), which can influence the binding to the polysaccharide substrate (153). In a later study by Wu et al. (155), for *Pch*GH61D two other loop regions have been defined: a long C-terminal (LC) and a short (LS) loop (**Fig. 7a**). LC and LS loops are exclusively found for AA9 LPMOs, though only a little difference has been found within

AA9 LPMOs (153). Finally, an L3 loop has been identified by Borisova and coworkers (161), in NcLPMO9C (**Fig. 7a, 7c**), which is a short insertion between β -strands 3 and 4. In contrast, NcLPMO9M, HjCel61B and TaGH61A lack this L3 loop (161). Interestingly, Tan and coworkers (163) found that NcLPMO9F obtain similar LS and LC loops, but lacks both the L2 and the L3 loop (**Fig. 7d**). Likewise, TtLPMO9E and PcLPMO9D lacked these L2 and L3 loops, which was concluded based on sequence alignment analysis and structural simulation (153, 161).

The configuration of loops have been suggested to steer the substrate binding, and further influence the regioselectivity and substrate specificity, which is explained in detail in Section 4.5.

4.3. Carbohydrate-binding modules (CBMs)

As explained above, surface binding sites of LPMOs vary. A further diversification of LPMOs is the presence or absence of auxiliary modules, for example of CBMs (6, 143). CBMs are often found to attach to glycoside hydrolases to help the catalytic sites bind to their substrate. A CBM is defined as a contiguous amino acid sequence, having carbohydrate-binding activity, but without catalytic activity (96, 110). CBMs were originally named cellulose-binding domains (CBDs), based on the finding that initially discovered modules bound particularly to cellulose (96, 164). Later, other modules showed affinity for other substrates, such as chitin or starch, justifying the current terminology (e.g., CBM) (96). Currently, the CAZy database recognizes 88 CBM families (96). To AA9 LPMOs, based on sequence analysis, mainly cellulose-specific binding modules (e.g., CBM1 and CBM2) are linked (153). Further, it has been reported that roughly one fifth of AA9 LPMOs are appended to a C-terminal CBM1 through a flexible region named linker (110, 143, 165).

In particular studied for GHs, CBMs have been reported to be involved in substrate interaction and recognition, and thereby, they can influence the enzyme stability, activity and catalytic performance (107, 141, 166-168). Similar to LPMOs, CBMs have aromatic amino acids (e.g., tyrosine) in the substrate binding surface that are critical for the recognition of substrates like cellulose (165, 169). In context of LPMO activity, various studies have suggested that appended CBMs increase the LPMO substrate binding affinity, resulting more stable LPMO activity (165, 170-173). Although results were not conclusive, it has been suggested that the presence of a CBM altered the substrate preference of the LPMO, for instance from crystalline to amorphous structure, although no change in actual substrate specificity was observed. Further, CBMs might slightly alter the regioselectivity of the LPMO, reflected by a changed ratio of C1- and C4-oxidized products (Section 4.5 for more details) (161, 162, 168, 170-174).

4.4. LPMO catalytic pathways: O₂ vs H₂O₂

The first LPMO activity has been reported for CBP21 from the bacterium *Serratia marcescens* (later recoded as SmLPMO10A) by Vaaje-Kolstad and coworkers in 2010, as mentioned in the above text (8). Using isotopically labeled dioxygen (¹⁸O₂) and

mass spectrometry, Vaaje-Kolstad and coworkers demonstrated that the LPMO-induced oxygen atom in the oxidized ends of chito-oligosaccharides originated from dioxygen, confirming an overall monooxygenase reaction (8). In addition, it was also shown that oxidative cleavage of chitin by *Sm*LPMO10A requires small-molecule reducing agent (ascorbic acid here) and a metal cofactor (sodium was incorrectly assigned at that time) (8, 152). Later, Quinlan and coworkers reported that a AA9 LPMO *Ta*GH61A is a mono copper-dependent monooxygenase, which can release oxidized cello-oligosaccharides from cellulose in the presence of O₂ and reductant (gallate was used in this case) (145). Other studies further confirmed these results, and also showed that extracellular redox enzyme-based electron donating system (e.g., cellobiose dehydrogenases (CDHs)) can also activate the AA9 LPMOs in addition to small molecular reductants (175-178)

Subsequent work and studies focused on insights into the mechanism of the LPMO-catalyzed reaction, and several routes were proposed (6, 13, 145, 146, 179-181). It is well accepted that the LPMO catalytic cycle starts with a priming one-electron reduction, meaning that a Cu(II) is reduced to Cu(I) by external electron donors (**Fig. 8, route a**). Subsequently, O₂ binds to the Cu(I) to form a Cu(II)-superoxo species (Cu(II)-OO⁻) (181, 182). The following steps are still under debate. Overall, it was proposed that Cu(II)-OO⁻ obtains in total two hydrogen atoms and a second electron to eventually convert to Cu(II)-O[•] (not shown in **route a**). The different intermediates formed during this conversion are still under debate and not further discussed here. The proposed Cu(II)-O[•] subsequently catalyzes hydroxylation on the C1- and/or C4-carbon position of a carbohydrate substrate to form a Cu(II) (or Cu(I) as suggested by others) for a next catalytic cycle (**Fig. 8, route a**). The glycosidic bond next to the hydroxylated carbon atom in the carbohydrate substrate is destabilized due to the inserted oxygen atom (144, 181, 182). This destabilization results in a subsequent elimination reaction, further leading to glycosidic bond cleavage. Consequently, both oxidized and non-oxidized carbohydrates are generated (144, 181, 182) (**Fig. 8**). C1-oxidation results in δ -lactones which further convert to aldonic acids, while C4-oxidation generates 4-ketoaldoses that are present in equilibrium with their geminal diols (pH dependent) in aqueous solution (144, 183, 184) (**Fig. 8**). The selective oxidation of the C1- and/or C4-carbon positions determines LPMO regioselectivity, which is further discussed in Section 4.5.

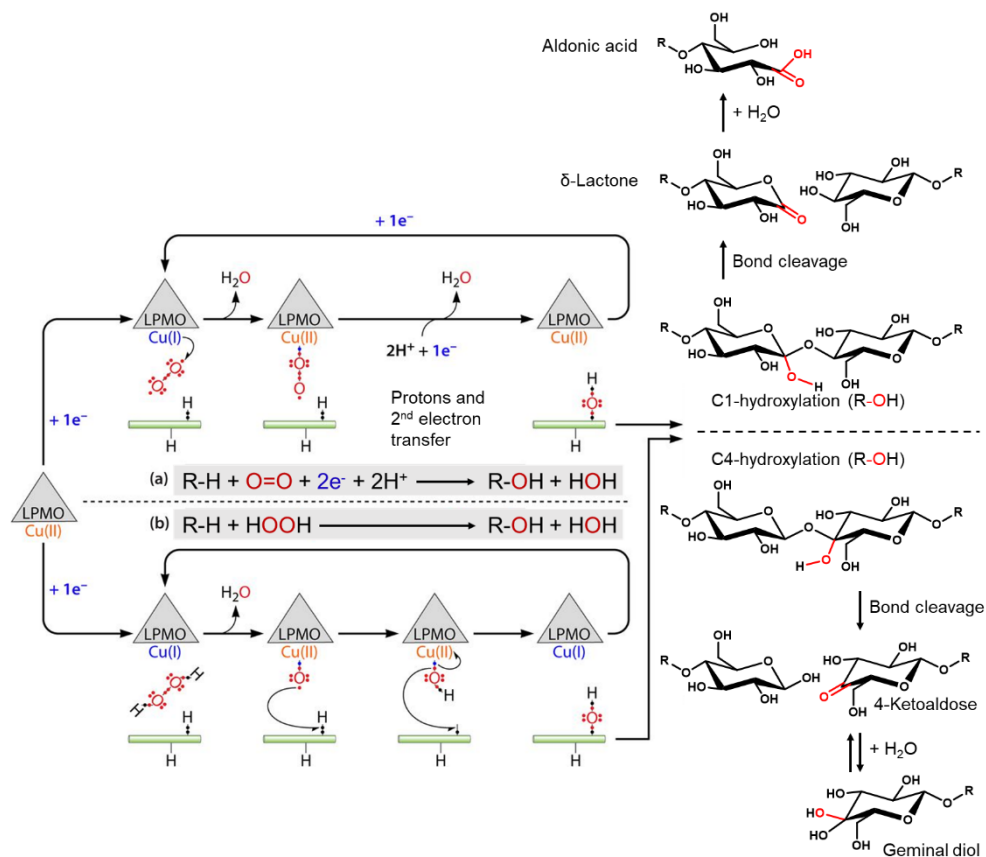


Fig. 8. Simplified schemes of LPMO reaction mechanisms. (a) O₂-dependent and (b) H₂O₂-dependent pathways both result in the hydroxylation of glycosidic bonds and, subsequently, bond cleavage. AA9 LPMOs have been shown to oxidatively cleave C1-, C4-, or C1-/C4-carbon positions of carbohydrate substrate (see more detail in Section 4.5). The schemes are based on Bissaro et al. (9) and Kabel et al. (112).

For a long time, the monooxygenase route (**Fig. 8, route a**) was well accepted based on the observation that LPMO catalysis relies on O₂. However, this was challenged by Bissaro and coworkers who proposed that H₂O₂ is the catalytically relevant co-substrate for LPMO-driven carbohydrate oxidation (185). In the H₂O₂-dependent route (**Fig. 8, route b**), after the priming reduction, H₂O₂ binds to Cu(I) to form a Cu(I)⋯H₂O₂ intermediate. Subsequently, Cu(I)⋯H₂O₂ converts to Cu(II)-O⁻, and the rest of the reaction is the same as for the O₂-dependent route. Bissaro and coworkers showed that addition of 200 μM of H₂O₂ to a ScLPMO10C-cellulose incubation yielded 26-fold more soluble oxidized cello-oligosaccharides after the first 2 min of incubation, compared to the incubation without the addition of H₂O₂ (185). Although H₂O₂ has been shown to enhance the LPMO catalytic reaction, over dosed H₂O₂ results in oxidative damage of the LPMO active site and further leads to the self-inactivation of LPMOs (185).

4.5. AA9 LPMO regioselectivity and substrate specificity

In the first published LPMO activity paper, Vaaje-Kolstad and coworkers demonstrated that CBP21 oxidizes the C1-carbon position of chitin (8). Regarding AA9 LPMOs, Quinlan and coworkers were the first to determine that *TaGH61A* oxidatively cleaved cellulose at both C1- and C4-carbon position (145). Afterwards, more and more AA9 LPMOs have been characterized for their regioselectivity, which was extensively reviewed by Frommhagen and coworkers (12). Oxidation at the C6-carbon position was suggested to occur for *PaGH61B* from *Podospira anserina* (186), however, the C6-oxidation was not confirmed by other studies (13, 144). Interestingly, many fungal species, such as *P. anserina*, *N. crassa* or *M. thermophila* produce a variety of LPMOs differing in C1- or C4- or both C1- and C4-oxidizing abilities (12). The reason why (AA9) LPMOs exhibit regioselectivities has not been elucidated yet. It is proposed that the regioselectivity reflects how the LPMOs interact with the substrates (e.g., exposed cellulose-fiber surface) (12). Borisova and coworkers suggested that differences in the accessibility of the solvent-facing axial position around the active site may explain the regioselectivity of LPMOs (161). In brief, C1-oxidizing AA9 LPMOs possess a tyrosine in this axial position, the mixed C1-/C4-oxidizing ones have a proline, and the C4-oxidizing ones have an alanine or an aspartate residue (161). In addition to look at differences in single amino acid residues, Vu and coworkers performed amino acid sequence analysis of 497 (putative) AA9 LPMOs with their published regioselectivity, and four groups (PMO1, PMO2, PMO3 and PMO3*) were classified (187). It was shown that the group of PMO1 and PMO2 are C1- and C4-oxidizing LPMOs, respectively, and the major difference is that PMO2s have an extra L3 loop (described in Section 4.2.2) compared to PMO1s. LPMOs from the groups of PMO3 and PMO3* do not have a L3, but a L2 loop. However, PMO3s are both C1- and C4-oxidizers, while PMO3*s are C1-oxidizers (187). This regioselectivity difference was further explained by the different amino acids present in these LPMOs. The study by Vu and coworkers promotes the understanding of correlation between LPMO (amino acid sequence based) structure and its regioselectivity. However, with more and more published AA9 LPMO data regarding regioselectivity and amino acid sequences, an increased number of exceptions was shown in each PMO group, indicating an still inadequate understanding of LPMO structure-regioselectivity correlation.

In context of substrate-specificity, initially, AA9 LPMOs were shown to only oxidize cellulose. In 2014, Isaksen and coworkers showed that *NcLPMO9C* is able to oxidatively cleave soluble cello-oligosaccharides (183). Later, Borisova and coworkers found that *NcLPMO9C* is also active towards xyloglucan and mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan (161). From this time onwards, oxidative cleavage of xyloglucan has been observed for various AA9 LPMOs (121, 162, 188-195). In addition, Frommhagen and coworkers reported that *MtLPMO9A* is able to oxidatively cleave xylan, but only when associated with cellulose (124). A detailed overview of AA9 LPMOs and their substrate specificity can be found in the review paper of Frommhagen and coworkers (12). The discovery that certain LPMOs showed activity towards soluble carbohydrate substrates, offers possibilities to study the LPMO-substrate interaction. Using NMR spectroscopy and X-ray crystallography, several studies revealed that various amino acid residues (e.g.,

asparagine and tyrosine) in the substrate binding surface are indeed involved in the interaction with substrates (121, 196, 197).

5. AA9 LPMOs from *Myceliophthora thermophila* and *Neurospora crassa*

As described previously, the degradation of (hemi-)cellulose requires the preferably synergistic collaboration of various enzymes. So far, widely used industrial (hemi-)cellulase cocktails are produced by several filamentous fungi such as *Trichoderma reesei* and *Aspergillus niger* (198). Another potential (hemi-)cellulase producer is *M. thermophila* C1, which shows more than 50 putative GHs and 26 putative LPMOs, of which 22 from AA9 (e.g., *MtLPMOs*), in its genome (199, 200). To understand why this fungus has 22 distinct AA9 LPMOs in its genome, and to learn whether these LPMOs differed in e.g., regioselectivity and substrate specificity, our Laboratory of Food Chemistry aimed to characterize all *MtLPMOs*. Till the end of 2017, 9 out of 22 LPMOs were (partially) characterized, summarized in **Table 1**.

Table 1. (Partially) characterized LPMOs from *M. thermophila*.

Enzyme name	Production host	Presence of CBM	Regio-selectivity	Substrate specificity
<i>MtLPMO9A</i> ¹	<i>M. thermophila</i> C1	No	C1/C4	Cellulose (RAC), xylan associated to cellulose, xyloglucan ¹⁰ , mixed β -(1→3, 1→4)-linked glucan ¹⁰
<i>MtLPMO9B</i> ²	<i>M. thermophila</i> C1	CBM1	C1	Cellulose (RAC)
<i>MtLPMO9C</i> ³	<i>M. thermophila</i> C1	No	C4	Cellulose (RAC), xyloglucan ¹⁰ , mixed β -(1→3, 1→4)-linked glucan ¹⁰
<i>MtLPMO9D</i> ⁴	<i>M. thermophila</i> C1	No	C1	Cellulose (RAC)
	<i>P. pastoris</i>	No	C1	Cellulose (PASC)
<i>MtLPMO9E</i> ⁵	<i>M. thermophila</i> C1	No	C4	Cellulose (RAC), xyloglucan ¹⁰ , cello-oligosaccharides (DP≥5)
<i>MtLPMO9F</i> ⁶	<i>M. thermophila</i> C1	CBM1	C4	Cellulose (RAC) ¹¹
<i>MtLPMO9G</i> ⁷	<i>M. thermophila</i> C1	No	C4	Cellulose (RAC) ¹¹
<i>MtLPMO9H</i> ⁸	<i>P. pastoris</i>	CBM1	C1/C4	Cellulose (RAC) ¹¹
MYCTH_112089 ⁹	<i>P. pastoris</i>	No	C1	Cellulose (PASC) ¹¹

¹MYCTH_85556, UniProt ID: G2QNT0 (124, 190); ²MYCTH_80312, UniProt ID: G2QCJ3 (190);

³MYCTH_100518, UniProt ID: G2QA92 (201); ⁴MYCTH_92668, UniProt ID: G2QAB5 (202) (187);

⁵MYCTH_79765, UniProt ID: G2Q7A5 (14); ⁶MYCTH_111088, UniProt ID: G2Q9F7 (14);

⁷MYCTH_55803, UniProt ID: n.d. (14); ⁸MYCTH_46583, UniProt ID: G2Q9T3 (203);

⁹MYCTH_112089, UniProt ID: G2QI82 (187)

¹⁰Minor activity

¹¹Only tested on cellulose

In addition, and in context of this research, LPMOs from the filamentous fungus *N. crassa* were also of interest, in particular to broaden the AA9 LPMO toolbox. This fungus has 14 AA9 LPMOs annotated in its genome (96), and is a well-known model eukaryotic organism for study of genetics, bioinformatics, molecular biology and biochemistry (204-207). A summary of (partially) characterized AA9 NcLPMOs is shown in **Table 2**.

Table 2. (Partially) characterized LPMOs from *N. crassa*.

Enzyme name	Production host	Presence of CBM	Regio-selectivity	Substrate specificity
NcLPMO9A ¹	<i>P. pastoris</i>	CBM1	C4	Cellulose (RAC) ¹⁰
NcLPMO9C ²	<i>P. pastoris</i>	CBM1	C4	Cellulose (RAC), xyloglucan, cello-oligosaccharides (DP \geq 4), glucomannan ¹¹ , mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan ¹¹ , carboxymethylcellulose ¹¹
NcLPMO9D ³	<i>P. pastoris</i> / <i>N. crassa</i>	No	C4	Cellulose (RAC) ¹⁰
NcLPMO9E ⁴	<i>P. pastoris</i> / <i>N. crassa</i>	CBM1	C1	Cellulose (RAC)
NcLPMO9F ⁵	<i>P. pastoris</i>	No	C1	Cellulose (PASC) ¹⁰
NcLPMO9M ⁶	<i>P. pastoris</i> / <i>N. crassa</i>	No	C1/C4	Cellulose (PASC) ¹⁰
NCU07760 ⁷	<i>P. pastoris</i>	CBM1	C1/C4	Cellulose (PASC) ¹⁰
NCU00836 ⁸	<i>P. pastoris</i>	CBM1	C1	Cellulose (PASC) ¹⁰
NCU01867 ⁹	<i>P. pastoris</i>	CBM1	C1	Cellulose (PASC) ¹⁰

¹NCU2240, UniProt ID: Q7S439 (184, 187); ²NCU02916, UniProt ID: Q7SHI8 (161, 177, 183, 187, 188, 196, 208); ³NCU01050, UniProt ID: Q1K8B6 (157, 182, 187); ⁴NCU08760, UniProt ID: Q7RWN7 (157, 182, 187, 208); ⁵NCU03328, UniProt ID: Q1K4Q1 (177, 187, 208); ⁶NCU07898, UniProt ID: Q7SA19 (157, 182, 187); ⁷UniProt ID: Q7S111 (187); ⁸UniProt ID: Q7SCJ5 (187); ⁹UniProt ID: Q7SHD9 (177, 208)

¹⁰Only tested on cellulose

¹¹Minor activity

6. Analytical approaches and assays to monitor LPMO activity

Mostly, a combination of HPAEC-PAD and MALDI-TOF-MS has been used to determine oxidized products formed from cellulose or chitin by LPMOs (**Table 3**) (124, 145, 162, 182, 186, 188, 190, 209). In addition to MALDI-TOF-MS, direct infusion electrospray-mass spectrometry (ESI-MS) was also applied for structural determination of LPMO-generated oxidized oligosaccharides (162, 188). Although these methods are useful to monitor product profiles, further method development is required for quantification of products formed, and corresponding kinetic analysis. So far HPAEC-PAD is the most widely used approach to perform quantification or kinetic analysis. However, oxidized (mono-)oligo-saccharides are not commonly available as standards, which limits the

structural elucidation and quantification of oxidized products. Other indirect approaches not based on the direct detection of LPMO-generated products have been developed, which are mainly used for fast screening of LPMO activity, as summarized in **Table 3**. These indirect approaches (e.g., Amplex Red assay) do not provide information of LPMO mode-of-action and substrate cleavage profiles, which is the main aim in this thesis, and thus are not discussed in detail. Although development of new assays for kinetic studies of LPMOs is of great interest, it is not involved in this thesis. In this thesis, we aimed at further developing methods not only to monitor the product profiles, but also unambiguously determine the structures of LPMO-generated oxidized products of cellulose and xyloglucan.

As mentioned above, mass spectrometric analysis such as MALDI-TOF-MS and direct infusion ESI-MS allow structural determination. However, currently, these approaches cannot distinguish more complicated isomers, for example, the oxidized xyloglucan oligosaccharides (162, 192, 193). Furthermore, depending on the pH and buffer conditions, C1-oxidized oligosaccharides can be present in δ -lactone and/or aldonic acid forms, while C4-oxidized ones can be present in 4-ketoaldose or geminal diol forms. The same DP of δ -lactone and 4-ketoaldose has the same molecular weight, and same DP of aldonic acid and geminal diols shares the same molecular weight as well. Occurrence of these same masses, together with the presence of non-oxidized oligosaccharides complicates and challenges the unambiguous identification of oxidized products.

In order to overcome the above mentioned mass spectrometric analysis limitations, other techniques have been developed, including reversed phase ultra-high performance liquid chromatography-ultraviolet detection-electrospray ionization-mass spectrometry (RP-UHPLC-UV-ESI-MS), porous graphitic carbon-liquid chromatography-charged aerosol detection/mass spectrometry (PGC-LC-CAD/MS) and hydrophilic interaction liquid chromatography-mass spectrometry (HILIC-MS) (184, 201). These techniques have the major advantage that separation of (isomeric) oligomers precedes mass identification, which can help identify their structures. However, so far, full separation of complex oxidized oligosaccharide mixtures have only been achieved to a certain extent (**Table 3**). Furthermore, systematic studies to understand how mass fragmentation spectra can unambiguously be translated in structural characteristics of different (oxidized) oligosaccharides have not been published at the end of 2017 or before.

Table 3. Analytical approaches and assays for the determination of LPMO activity till the end of 2017.

No. ^a	Analytical technique ^b	Type of analysis	Separation	Major advantage	Major limitation
1	Direct infusion ESI-MS	Qualitative	No	Direct detection	Cannot distinguish isomers
2	MALDI-TOF-MS	Qualitative	No	High-throughput direct detection	Cannot distinguish isomers
3	NMR spectroscopy	Qualitative	No	Accurate structural determination	Low-throughput
4	Glycan microarray	Qualitative / Semi-quantitative	No	Fast screening of substrate specificity	Standards for target compounds are required
5	HPAEC-PAD	Qualitative / Quantitative	Yes	Good separation of (oxidized) oligosaccharides	No structural determination
6	HPAEC-PAD & Post column	Qualitative / Quantitative	Yes	Good separation of (oxidized) monosaccharides	No structural determination
7	HPLC-RI-UV	Qualitative / Quantitative	Yes	Good separation of (oxidized) monosaccharides	No structural determination
8	PGC-LC-CAD/MS	Qualitative/Quantitative ^c	Yes	Moderate separation and structural determination	Only applied for DP<4
9	RP-UHPLC-UV-ESI-MS	Qualitative/Quantitative ^c	Yes	Good separation and structural determination	Signal intensity relatively low
10	HILIC-MS	Qualitative/Quantitative ^c	Yes	Moderate separation and structural determination	Co-elution of C4- and non-oxidized oligosaccharides; broad elution time
11	PACE	Qualitative/ Semi-quantitative	Yes	Fast screening of substrate specificity	No structural determination; no separation of the same DP of oxidized and non-oxidized products
12	Amplex Red Assay	Qualitative/Quantitative	No	Fast screening of substrate specificity; high-throughput	No detection of oxidized products
13	XPS	Qualitative	No	Suitable for insoluble substrates	No detection of oxidized products; low-throughput
14	Fluorescence spectroscopy	Qualitative/Quantitative ^c	No	Fast screening of substrate specificity; high-throughput	No detection of oxidized products; Side reaction occurs
15	EPR	Qualitative/ Semi-quantitative ^c	No	Suitable for insoluble substrate	No detection of oxidized products; low-throughput

^aReferences: 1 (162, 183); 2 (8, 124); 3 (183); 4, (188); 5 (124); 6 (210); 7 (210); 8 (184); 9 (201); 10 (184); 11 (121); 12 (208); 13 (211); 14 (212); 15 (213)

^bAbbreviations: ESI-MS, Electrospray ionization-mass spectrometry; MALDI-TOF-MS, Matrix-assisted laser desorption/ionization-time of flight-mass spectrometry; NMR, Nuclear magnetic resonance; HPAEC-PAD, High performance anion-exchange chromatography-pulsed amperometric detection; HPLC-RI-UV, High performance liquid chromatography-refractive index-ultraviolet detection; RP-UHPLC-UV-ESI-MS, Reversed phase ultra-high performance liquid chromatography-ultraviolet detection-electrospray ionization-mass spectrometry; PGC-LC-CAD/MS, Porous graphitic carbon-liquid chromatography-charged aerosol detection/mass spectrometry; HILIC-MS, Hydrophilic interaction liquid chromatography-mass spectrometry; PACE, Polysaccharide analysis by carbohydrate gel electrophoresis; XPS, X-ray photoelectron spectroscopy; EPR, Electron paramagnetic resonance

^cMethodology for quantitative analysis needs to be developed and the corresponding standards need to be produced

7. Project aim

This project aims to study (new) AA9 *Mt*LPMOs and *Nc*LPMOs to provide further insights into the mode-of-action, substrate cleavage profile and structure-substrate specificity/regioselectivity correlation. Overall, it is hypothesized that AA9 LPMO substrate specificity and cleavage profiles correlate to their structural configuration.

8. Thesis outline

In **Chapter 1**, the background and aim of this project is described. The plant cell wall architecture and the cell wall polysaccharide degrading enzymes are thoroughly introduced. The existing studies and knowledge of LPMOs at the time of the beginning of this thesis (early 2018) are introduced.

In **Chapter 2**, cellulose specificity and the resulting product profiles are studied for four AA9 LPMOs. We find that four AA9 LPMOs generate mainly oxidized cellobiose from bacterial cellulose, while more evenly distributed mixtures of oxidized cello-oligosaccharides (DP2–8) are observed with microcrystalline cellulose and regenerated amorphous cellulose.

In **Chapter 3**, mass spectrometric fragmentation behaviors and patterns of (NaBD₄-reduced) non-, C4- and C1-oxidized cello-oligosaccharides are studied by using hydrophilic interaction coupled to (negative ion mode) mass spectrometry.

In **Chapter 4**, we discover a series of novel LPMO products: double, C4 and C6, oxidized cello-oligosaccharides, aided by NaBD₄ reduction, and hydrophilic interaction chromatography coupled to mass spectrometric analysis.

In **Chapter 5**, segments around the LPMO active site involved in substrate recognition are defined. This study also investigates the changes in activity induced by shortening the second active site segment or removing the CBM1 from *Neurospora crassa* LPMO9C.

In **Chapter 6**, We elucidate multiple structures of oxidized xyloglucan degradation products released by LPMOs and reveal a correlation between the configuration of active site segments and xyloglucan specificity of LPMOs. In particular, oxidative xyloglucan-active LPMOs, which are tolerant and intolerant to xyloglucan substitutions are structurally and phylogenetically distinguished from xyloglucan-inactive LPMOs.

In **Chapter 7**, two new AA9 LPMOs from *M. thermophila* C1 are characterized. Their xyloglucan specificity and substitution (in)tolerance are as predicted according to the active site segment structure-based phylogenetic grouping of AA9 LPMOs described in Chapter 6. In addition, novel double, C4 and C6, oxidized xylogluco-oligosaccharides are identified.

In **Chapter 8**, new AA16 enzymes are found to not have typical LPMO-like catalytic action towards carbohydrates. However, AA16 enzymes are shown to boost cellulose active AA9 LPMOs from *M. thermophila* C1.

In the final **Chapter 9**, the relevance of this research for future study and application of LPMOs in the enzymatic degradation of lignocellulosic biomass is discussed. Furthermore, we discuss new findings in the LPMO field (2018-2021) together with results obtained in this thesis.

9. References

1. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, et al. The path forward for biofuels and biomaterials. *Science*. 2006;311:484-9.
2. Nanda S, Mohammad J, Reddy SN, Kozinski JA, Dalai AK. Pathways of lignocellulosic biomass conversion to renewable fuels. *Biomass Conver Biorefin*. 2014;4:157-91.
3. Vanholme B, Desmet T, Ronse F, Rabaey K, van Breusegem F, De Mey M, et al. Towards a carbon-negative sustainable bio-based economy. *Front Plant Sci*. 2013;4:174.
4. Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuel Bioprod Biorefin*. 2007;1:119-34.
5. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, et al. Fungal cellulases. *Chem Rev*. 2015;115:1308-448.
6. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.
7. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, et al. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20:1479-83.
8. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørle M, et al. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219-22.
9. Bissaro B, Varnai A, Rohr AK, Eijsink VGH. Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. *Microbiol Mol Biol Rev*. 2018;82:4.
10. Chylenski P, Bissaro B, Sørle M, Rohr AK, Varnai A, Horn SJ, et al. Lytic polysaccharide monooxygenases in enzymatic processing of lignocellulosic biomass. *ACS Catal*. 2019;9:4970-91.
11. Eijsink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnol Biofuels*. 2019;12:58.
12. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol*. 2018;9:1080.
13. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33:747-61.
14. Frommhagen M. Lytic polysaccharide monooxygenase from *Myceliophthora Thermophila* C1 [PhD thesis]: Wageningen University and Research; 2017.
15. Pauly M, Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J*. 2008;54:559-68.
16. Cosgrove DJ, Jarvis MC. Comparative structure and biomechanics of plant primary and secondary cell walls. *Front Plant Sci*. 2012;3.
17. Ralph J, Lapierre C, Boerjan W. Lignin structure and its engineering. *Curr Opin Biotechnol*. 2019;56:240-9.
18. Cassab GI, Varner JE. Cell-Wall Proteins. *Annu Rev Plant Physiol Plant Mol Biol*. 1988;39:321-53.

19. Showalter AM. Structure and function of plant-cell wall proteins. *Plant Cell*. 1993;5:9-23.
20. Reiter WD. Biosynthesis and properties of the plant cell wall. *Curr Opin Plant Biol*. 2002;5:536-42.
21. Mohnen D. Pectin structure and biosynthesis. *Curr Opin Plant Biol*. 2008;11:266-77.
22. McNeil M, Darvill AG, Fry SC, Albersheim P. Structure and function of the primary cell walls of plants. *Annu Rev Biochem*. 1984;53:625-63.
23. Cosgrove DJ. Wall structure and wall loosening. A look backwards and forwards. *Plant Physiol*. 2001;125:131-4.
24. Kabel MA, Bos G, Zeevalking J, Voragen AG, Schols HA. Effect of pretreatment severity on xylan solubility and enzymatic breakdown of the remaining cellulose from wheat straw. *Bioresour Technol*. 2007;98:2034-42.
25. del Rio JC, Rencoret J, Prinsen P, Martinez AT, Ralph J, Gutierrez A. Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *J Agric Food Chem*. 2012;60:5922-35.
26. Mccann MC, Wells B, Roberts K. Complexity in the spatial localization and length distribution of plant cell-wall matrix polysaccharides. *J Microsc-Oxford*. 1992;166:123-36.
27. Brethauer S, Shahab RL, Studer MH. Impacts of biofilms on the conversion of cellulose. *Appl Microbiol Biotechnol*. 2020;104:5201-12.
28. Brown Jr RM. Cellulose structure and biosynthesis: what is in store for the 21st century? *J Polym Sci, Part A: Polym Chem*. 2004;42:487-95.
29. Saxena IM, Brown RM. Cellulose biosynthesis: Current views and evolving concepts. *Ann Bot*. 2005;96:9-21.
30. Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, et al. Toward a systems approach to understanding plant cell walls. *Science*. 2004;306:2206-11.
31. Klemm D, Heublein B, Fink HP, Bohn A. Cellulose: Fascinating biopolymer and sustainable raw material. *Angew Chem Int Edit*. 2005;44:3358-93.
32. Chen H. Chemical composition and structure of natural lignocellulose. *Biotechnology of lignocellulose*. Dordrecht: Springer; 2014. p. 25-71.
33. Parthasarathi R, Bellesia G, Chundawat S, Dale B, Langan P, Gnanakaran S. Insights into hydrogen bonding and stacking interactions in cellulose. *J Physic Chem A*. 2011;115:14191-202.
34. Nishiyama Y, Langan P, Chanzy H. Crystal structure and hydrogen-bonding system in cellulose I_β from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc*. 2002;124:9074-82.
35. Atalla RH, Vanderhart DL. Native cellulose: a composite of two distinct crystalline forms. *Science*. 1984;223:283-5.
36. Nishiyama Y, Sugiyama J, Chanzy H, Langan P. Crystal structure and hydrogen bonding system in cellulose I_α from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc*. 2003;125:14300-6.
37. Viētor RJ, Newman RH, Ha MA, Apperley DC, Jarvis MC. Conformational features of crystal-surface cellulose from higher plants. *Plant J*. 2002;30:721-31.
38. Jarvis M. Cellulose stacks up. *Nature*. 2003;426:611-2.
39. Mittal A, Katahira R, Himmel ME, Johnson DK. Effects of alkaline or liquid-ammonia treatment on crystalline cellulose: changes in crystalline structure and effects on enzymatic digestibility. *Biotechnol Biofuels*. 2011;4:1-16.
40. Zugenmaier P. Conformation and packing of various crystalline cellulose fibers. *Prog Polym Sci*. 2001;26:1341-417.
41. Finkenstadt V, Millane R. Crystal structure of *Valonia* cellulose I_β. *Macromolecules*. 1998;31:7776-83.
42. Valenzuela SV, Valls C, Schink V, Sanchez D, Roncero MB, Diaz P, et al. Differential activity of lytic polysaccharide monooxygenases on celluloses of different crystallinity. Effectiveness in the sustainable production of cellulose nanofibrils. *Carbohydr Polym*. 2019;207:59-67.
43. Hubbell CA, Ragauskas AJ. Effect of acid-chlorite delignification on cellulose degree of polymerization. *Bioresour Technol*. 2010;101:7410-5.
44. Zhao T, Chen ZZ, Lin XR, Ren ZY, Li B, Zhang YY. Preparation and characterization of microcrystalline cellulose (MCC) from tea waste. *Carbohydr Polym*. 2018;184:164-70.
45. Bondeson D, Mathew A, Oksman K. Optimization of the isolation of nanocrystals from microcrystalline cellulose by acid hydrolysis. *Cellulose*. 2006;13:171-80.
46. Elazzouzi-Hafraoui S, Nishiyama Y, Putaux JL, Heux L, Dubreuil F, Rochas C. The shape and size distribution of crystalline nanoparticles prepared by acid hydrolysis of native cellulose. *Biomacromolecules*. 2008;9:57-65.

47. Agblevor FA, Ibrahim MM, El-Zawawy WK. Coupled acid and enzyme mediated production of microcrystalline cellulose from corn cob and cotton gin waste. *Cellulose*. 2007;14:247-56.
48. Okiyama A, Motoki M, Yamanaka S. Bacterial cellulose. 3. Development of a new form of cellulose. *Food Hydrocoll*. 1993;6:493-501.
49. Shibazaki H, Kuga S, Okano T. Mercerization and acid hydrolysis of bacterial cellulose. *Cellulose*. 1997;4:75-87.
50. Mormann W, Demeter J, Wagner T. Partial silylation of cellulose with predictable degree of silylation–stoichiometric silylation with hexamethyldisilazane in ammonia. *Macromol Chem Phys*. 1999;200:693-7.
51. Kumar V, Kothari S, Banker GS. Effect of the agitation rate on the generation of low-crystallinity cellulose from phosphoric acid. *J Appl Polym Sci*. 2001;82:2624-8.
52. Matsuoka S, Kawamoto H, Saka S. What is active cellulose in pyrolysis? An approach based on reactivity of cellulose reducing end. *J Anal Appl Pyrolysis*. 2014;106:138-46.
53. Kambli ND, Mageshwaran V, Patil PG, Saxena S, Deshmukh RR. Synthesis and characterization of microcrystalline cellulose powder from corn husk fibres using biochemical route. *Cellulose*. 2017;24:5355-69.
54. Zhao H, Jones CL, Baker GA, Xia S, Olubajo O, Person VN. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. *J Biotechnol*. 2009;139:47-54.
55. Ciolacu D, Ciolacu F, Popa VI. Amorphous cellulose: Structure and characterization. *Cellul Chem Technol*. 2011;45:13-21.
56. Zhang YHP, Cui JB, Lynd LR, Kuang LR. A transition from cellulose swelling to cellulose dissolution by *O*-phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules*. 2006;7:644-8.
57. Scheller HV, Ulvskov P. Hemicelluloses. *Annu Rev Plant Biol*. 2010;61:263-89.
58. Peng F, Peng P, Xu F, Sun R-C. Fractional purification and bioconversion of hemicelluloses. *Biotechnol Adv*. 2012;30:879-903.
59. McMillan JD. Pretreatment of lignocellulosic biomass. 1994.
60. Brillouet JM, Joseleau JP, Utile JP, Lelievre D. Isolation, purification and characterization of a complex heteroxylan from industrial wheat bran. *J Agric Food Chem*. 1982;30:488-95.
61. Carpita NC. Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Biol*. 1996;47:445-76.
62. Kabel MA, van den Borne H, Vincken J-P, Voragen AG, Schols HA. Structural differences of xylans affect their interaction with cellulose. *Carbohydr Polym*. 2007;69:94-105.
63. Hsieh YS, Harris PJ. Xyloglucans of monocotyledons have diverse structures. *Molecular plant*. 2009;2:943-65.
64. Zabotina OA. Xyloglucan and its biosynthesis. *Front Plant Sci*. 2012;3:134.
65. Fry SC. The structure and functions of xyloglucan. *J Exp Bot*. 1989;40:1-11.
66. Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, et al. An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plant*. 1993;89:1-3.
67. Hilz H, de Jong LE, Kabel MA, Schols HA, Voragen AG. A comparison of liquid chromatography, capillary electrophoresis, and mass spectrometry methods to determine xyloglucan structures in black currants. *J Chromatogr*. 2006;1133:275-86.
68. Park YB, Cosgrove DJ. Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol*. 2015;56:180-94.
69. Vincken J-P, Wijsman AJ, Beldman G, Niessen WM, Voragen AG. Potato xyloglucan is built from XXGG-type subunits. *Carbohydr Res*. 1996;288:219-32.
70. Vincken JP, York WS, Beldman G, Voragen AG. Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiol*. 1997;114:9-13.
71. Kiefer LL, York WS, Albersheim P, Darvill AG. Structural characterization of an arabinose-containing heptadecasaccharide enzymically isolated from sycamore extracellular xyloglucan. *Carbohydr Res*. 1990;197:139-58.
72. York WS, Kolli VK, Orlando R, Albersheim P, Darvill AG. The structures of arabinoxyloglucans produced by solanaceous plants. *Carbohydr Res*. 1996;285:99-128.
73. Kato Y, Iki K, Matsuda K. Cell-wall polysaccharides of immature barley plants. II. Characterization of a xyloglucan. *Agric Biol Chem*. 1981;45:2745-53.
74. Mabusela WT, Stephen AM, Rodgers AL, Gerneke DA. A highly substituted glucan that coats the seeds of *Helipterum eximium*. *Carbohydr Res*. 1990;203:336-40.
75. Nishinari K. Konjac glucomannan. *Developments in food science*. 41: Elsevier; 2000. p. 309-30.
76. Maeda M, Shimahara H, Sugiyama N. Detailed examination of the branched structure of konjac glucomannan. *Agric Biol Chem*. 1980;44:245-52.

77. Dave V, McCarthy SP. Review of konjac glucomannan. *J Environ Polymer Degradation*. 1997;5:237.
78. Moreira L. An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol*. 2008;79:165-78.
79. Timell T. Recent progress in the chemistry of wood hemicelluloses. *Wood Sci Technol*. 1967;1:45-70.
80. Ishurd O, Kermagi A, Elghazoun M, Kennedy JF. Structural of a glucomannan from *Lupinus varius* seed. *Carbohydr Polym*. 2006;65:410-3.
81. Saha BC. Hemicellulose bioconversion. *J Ind Microbiol Biotechnol*. 2003;30:279-91.
82. Cosgrove DJ. Growth of the plant cell wall. *Nat Rev Mol Cell Biol*. 2005;6:850-61.
83. Bauer WD, Talmadge KW, Keegstra K, Albersheim P. The structure of plant cell walls: II. The hemicellulose of the walls of suspension-cultured sycamore cells. *Plant Physiol*. 1973;51:174-87.
84. Qaseem MF, Shaheen H, Wu A-M. Cell wall hemicellulose for sustainable industrial utilization. *Renew Sustain Energy Rev*. 2021;144:110996.
85. Henriksson Å, Gatenholm P. Controlled assembly of glucuronoxylans onto cellulose fibres. *Holzforschung*. 2001;55:494-502.
86. Yu L, Lyczakowski JJ, Pereira CS, Kotake T, Yu X, Li A, et al. The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage. *Plant Physiol*. 2018;178:1011-26.
87. Vincken J-P, de Keizer A, Beldman G, Voragen AGJ. Fractionation of xyloglucan fragments and their interaction with cellulose. *Plant Physiol*. 1995;108:1579-85.
88. Nishitani K, Masuda Y. Acid pH-induced structural changes in cell wall xyloglucans in *Vigna angularis* epicotyl segments. *Plant Sci Lett*. 1982;28:87-94.
89. Eckardt NA. Role of xyloglucan in primary cell walls. *Plant Cell*. 2008;20:1421-2.
90. Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, et al. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev*. 2004;3:29-60.
91. Nishimura H, Kamiya A, Nagata T, Katahira M, Watanabe T. Direct evidence for a ether linkage between lignin and carbohydrates in wood cell walls. *Sci Rep*. 2018;8:1-11.
92. Ralph J, Bunzel M, Marita JM, Hatfield RD, Lu F, Kim H, et al. Peroxidase-dependent cross-linking reactions of *p*-hydroxycinnamates in plant cell walls. *Phytochem Rev*. 2004;3:79-96.
93. Ishii T. Structure and functions of feruloylated polysaccharides. *Plant Sci*. 1997;127:111-27.
94. Iiyama K, Lam TB-T, Stone BA. Covalent cross-links in the cell wall. *Plant Physiol*. 1994;104:315.
95. Giummarella N, Pu Y, Ragauskas AJ, Lawoko M. A critical review on the analysis of lignin carbohydrate bonds. *Green Chem*. 2019;21:1573-95.
96. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490-D495.
97. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6:1-14.
98. Reese ET. Enzymatic hydrolysis of cellulose. *Appl Microbiol*. 1956;4:39.
99. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb Cell Fact*. 2016;1:1-13.
100. Lynd LR, Weimer PJ, Van Zyl WH, Pretorius IS. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev*. 2002;66:506-77.
101. Horn SJ, Sikorski P, Cedervist JB, Vaaje-Kolstad G, Sørle M, Synstad B, et al. Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci U S A*. 2006;103:18089-94.
102. Horn SJ, Sørle M, Varum KM, Valjamae P, Eijsink VGH. Measuring processivity. *Method Enzymol*. 2012;510:69-95.
103. Jeon SD, Yu KO, Kim SW, Han SO. The processive endoglucanase EngZ is active in crystalline cellulose degradation as a cellulosomal subunit of *Clostridium cellulovorans*. *New Biotechnol*. 2012;29:365-71.
104. Payne CM, Bomble Y, Taylor CB, McCabe C, Himmel ME, Crowley MF, et al. Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation. *J Biol Chem*. 2011;286:41028-35.
105. Vuong TV, Wilson DB. Processivity, synergism, and substrate specificity of *Thermobifida fusca* Cel6B. *Appl Environ Microbiol*. 2009;75:6655-61.
106. Wu SS, Wu SF. Processivity and the mechanisms of processive endoglucanases. *Appl Biochem Biotechnol*. 2020;190:448-63.

107. Zhou W, Irwin DC, Escovar-Kousen J, Wilson DB. Kinetic studies of *Thermobifida fusca* Cel9A active site mutant enzymes. *Biochemistry*. 2004;43:9655-63.
108. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol*. 2011;91:1477-92.
109. Ekwe E, Morgenstern I, Tsang A, Storms R, Powlowski J. Non-hydrolytic cellulose active proteins: research progress and potential application in biorefineries. *Ind Biotechnol*. 2013;9:123-31.
110. Lenfant N, Hainaut M, Terrapon N, Drula E, Lombard V, Henrissat B. A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res*. 2017;448:166-74.
111. Corrêa TL, dos Santos LV, Pereira GA. AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol*. 2016;100:9-16.
112. Kabel MA, Frommhagen M, Sun P, Schols HA. Modification of plant carbohydrates using fungal enzymes. In: Zaragoza Ó, Casadevall A, editors. *Encyclopedia of Mycology*. Oxford: Elsevier; 2021. p. 370-384.
113. Kormelink F, Searle-Van Leeuwen M, Wood T, Voragen A. Purification and characterization of three endo-(1,4)- β -xylanases and one β -xylosidase from *Aspergillus awamori*. *J Biotechnol*. 1993;27:249-65.
114. Kormelink FJ, Gruppen H, Viëtor RJ, Voragen AG. Mode of action of the xylan-degrading enzymes from *Aspergillus awamori* on alkali-extractable cereal arabinoxylans. *Carbohydr Res*. 1993;249:355-67.
115. Saha BC, Bothast RJ. *Enzymology of xylan degradation*. ACS Publications; 1999.
116. Kolenová K, Vršanská M, Biely P. Mode of action of endo-(1,4)- β -xylanases of families 10 and 11 on acidic xylooligosaccharides. *J Biotechnol*. 2006;121:338-45.
117. Pell G, Taylor EJ, Gloster TM, Turkenburg JP, Fontes CM, Ferreira LM, et al. The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. *J Biol Chem*. 2004;279:9597-605.
118. Sørensen HR, Jørgensen CT, Hansen CH, Jørgensen CI, Pedersen S, Meyer AS. A novel GH43 α -L-arabinofuranosidase from *Humicola insolens*: mode of action and synergy with GH51 α -L-arabinofuranosidases on wheat arabinoxylan. *Appl Microbiol Biotechnol*. 2006;73:850-61.
119. Ryabova O, Vršanská M, Kaneko S, van Zyl WH, Biely P. A novel family of hemicellulolytic α -glucuronidase. *FEBS Lett*. 2009;583:1457-62.
120. Biely P, Westereng B, Puchart V, De Maayer P, Cowan DA. Recent progress in understanding mode of action of acetylxyloxy esterases. *J Appl Glycosci*. 2014;jag. JAG-2013_018.
121. Simmons TJ, Frandsen KEH, Ciano L, Tryfona T, Lenfant N, Poulsen JC, et al. Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat Commun*. 2017;8:1064.
122. Hüttner S, Várnai A, Petrović DM, Bach CX, Anh DTK, Thanh VN, et al. Specific xylan activity revealed for AA9 lytic polysaccharide monooxygenases of the thermophilic fungus *Malbranchea cinnamomea* by functional characterization. *Appl Environ Microbiol*. 2019;85: e01408-19.
123. Couturier M, Ladeveze S, Sulzenbacher G, Ciano L, Fanuel M, Moreau C, et al. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol*. 2018;14:306.
124. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
125. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annu Rev Phytopathol*. 2014;52:427-51.
126. Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, et al. A xyloglucan-specific endo-(1,4)- β -glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification and characterization of the recombinant enzyme. *Glycobiol*. 1999;9:93-100.
127. Grishutin SG, Gusakov AV, Markov AV, Ustinov BB, Semenova MV, Sinitsyn AP. Specific xyloglucanases as a new class of polysaccharide-degrading enzymes. *Biochim Biophys Acta Gen Subj*. 2004;1674:268-81.
128. Martinez-Fleites C, Guerreiro CI, Baumann MJ, Taylor EJ, Prates JA, Ferreira LM, et al. Crystal structures of *Clostridium thermocellum* xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. *J Biol Chem*. 2006;281:24922-33.
129. Vitcosque GL, Ribeiro LFC, de Lucas RC, da Silva TM, Ribeiro LF, Damasio ARD, et al. The functional properties of a xyloglucanase (GH12) of *Aspergillus terreus* expressed in *Aspergillus nidulans* may increase performance of biomass degradation. *Appl Microbiol Biotechnol*. 2016;100:9133-44.

130. Ariza A, Eklof JM, Spadiut O, Offen WA, Roberts SM, Besenmatter W, et al. Structure and activity of *Paenibacillus polymyxa* xyloglucanase from glycoside hydrolase family 44. *J Biol Chem.* 2011;286:33890-900.
131. Desmet T, Cantaert T, Gualfetti P, Nerinckx W, Gross L, Mitchinson C, et al. An investigation of the substrate specificity of the xyloglucanase Cel74A from *Hypocrea jecorina*. *FEBS J.* 2007;274:356-63.
132. Ichinose H, Araki Y, Michikawa M, Harazono K, Yaoi K, Karita S, et al. Characterization of an endo-processive-type xyloglucanase having a β -(1 \rightarrow 4)-glucan-binding module and an endo-type xyloglucanase from *Streptomyces avermitilis*. *Appl Environ Microbiol.* 2012;78:7939-45.
133. Rykov SV, Kornberger P, Herlet J, Tsurin NV, Zorov IN, Zverlov VV, et al. Novel endo-(1,4)- β -glucanase *Bgh12A* and xyloglucanase *Xgh12B* from *Aspergillus cervinus* belong to GH12 subgroup I and II, respectively. *Appl Microbiol Biotechnol.* 2019;103:7553-66.
134. Matsuzawa T, Kameyama A, Nakamichi Y, Yaoi K. Identification and characterization of two xyloglucan-specific endo-(1,4)-glucanases in *Aspergillus oryzae*. *Appl Microbiol Biotechnol.* 2020;104:8761-73.
135. Yaoi K, Mitsuishi Y. Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. *J Biol Chem.* 2002;277:48276-81.
136. Arnal G, Stogios PJ, Asohan J, Attia MA, Skarina T, Viborg AH, et al. Substrate specificity, regiospecificity, and processivity in glycoside hydrolase family 74. *J Biol Chem.* 2019;294:13233-47.
137. Attia M, Stepper J, Davies GJ, Brumer H. Functional and structural characterization of a potent GH 74 endo-xyloglucanase from the soil saprophyte *Cellvibrio japonicus* unravels the first step of xyloglucan degradation. *FEBS J.* 2016;283:1701-19.
138. Arnal G, Stogios PJ, Asohan J, Skarina T, Savchenko A, Brumer H. Structural enzymology reveals the molecular basis of substrate regiospecificity and processivity of an exemplar bacterial glycoside hydrolase family 74 endo-xyloglucanase. *Biochem J.* 2018;475:3963-78.
139. Reese ET, Siu RG, Levinson HS. The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J Bacteriol.* 1950;59:485.
140. Vaaje-Kolstad G, Horn SJ, van Aalten DM, Synstad B, Eijsink VG. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem.* 2005;280:28492-7.
141. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 2009;37:D233-D8.
142. Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J Mol Biol.* 2008;383:144-54.
143. Harris PV, Welner D, McFarland K, Re E, Navarro Poulsen J-C, Brown K, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry.* 2010;49:3305-16.
144. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc.* 2012;134:890-2.
145. Quinlan RJ, Sweeney MD, Leggio LL, Otten H, Poulsen J-CN, Johansen KS, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A.* 2011;108:15079-84.
146. Hemsworth GR, Davies GJ, Walton PH. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol.* 2013;23:660-8.
147. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol.* 2014;10:122-6.
148. Leggio LL, Simmons TJ, Poulsen J-CN, Frandsen KE, Hemsworth GR, Stringer MA, et al. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun.* 2015;6:1-9.
149. Couturier M, Ladeveze S, Sulzenbacher G, Ciano L, Fanuel M, Moreau C, et al. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol.* 2018;14:306-10.
150. Sabbadin F, Hemsworth GR, Ciano L, Henrissat B, Dupree P, Tryfona T, et al. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun.* 2018;9:756.

151. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevret D, et al. AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels*. 2019;12.
152. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VG, van Aalten DM. Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem*. 2005;280:11313-9.
153. Vaaje-Kolstad G, Forsberg Z, Loose JS, Bissaro B, Eijsink VG. Structural diversity of lytic polysaccharide monooxygenases. *Curr Opin Struct Biol*. 2017;44:67-76.
154. Frandsen KE, Lo Leggio L. Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes. *IUCrJ*. 2016;3:448-67.
155. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, et al. Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem*. 2013;288:12828-39.
156. Davies G, Henrissat B. Structures and mechanisms of glycosyl hydrolases. *Structure*. 1995;3:853-9.
157. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure*. 2012;20:1051-61.
158. Tandrup T, Frandsen KEH, Johansen KS, Berrin J-G, Lo Leggio L. Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochem Soc Trans*. 2018;46:1431.
159. Meier KK, Jones SM, Kaper T, Hansson H, Koetsier MJ, Karkehabadi S, et al. Oxygen activation by Cu LPMOs in recalcitrant carbohydrate polysaccharide conversion to monomer sugars. *Chem Rev*. 2017.
160. Hemsworth GR, Taylor EJ, Kim RQ, Gregory RC, Lewis SJ, Turkenburg JP, et al. The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc*. 2013;135:6069-77.
161. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.
162. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels*. 2015;8:90.
163. Tan TC, Kracher D, Gandini R, Sygmond C, Kittl R, Haltrich D, et al. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat Commun*. 2015;6:1-11.
164. Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Warren RAJ. Domains in microbial β -(1 \rightarrow 4)-glycanases: sequence conservation, function, and enzyme families. *Microbiol Rev*. 1991;55:303-15.
165. Courtade G, Forsberg Z, Heggset EB, Eijsink VGH, Aachmann FL. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J Biol Chem*. 2018;293:13006-15.
166. Irwin D, Shin DH, Zhang S, Barr BK, Sakon J, Karplus PA, et al. Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. *J Bacteriol*. 1998;180:1709-14.
167. Hong J, Ye XH, Zhang YHP. Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications. *Langmuir*. 2007;23:12535-40.
168. Nakamura A, Tasaki T, Ishiwata D, Yamamoto M, Okuni Y, Visootsat A, et al. Single-molecule imaging analysis of binding, processive movement, and dissociation of cellobiohydrolase *Trichoderma reesei* Cel6A and its domains on crystalline cellulose. *J Biol Chem*. 2016;291:22404-13.
169. Taylor CB, Talib MF, McCabe C, Bu LT, Adney WS, Himmel ME, et al. Computational investigation of glycosylation effects on a family 1 carbohydrate-binding module. *J Biol Chem*. 2012;287:3147-55.
170. Crouch LI, Labourel A, Walton PH, Davies GJ, Gilbert HJ. The contribution of non-catalytic carbohydrate binding modules to the activity of lytic polysaccharide monooxygenases. *J Biol Chem*. 2016;291:7439-49.
171. Forsberg Z, Nelson CE, Dalhus B, Mekasha S, Loose JS, Crouch LI, et al. Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J Biol Chem*. 2016;291:7300-12.
172. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, et al. High-resolution structure of a lytic polysaccharide monooxygenase from *Hypocrea jecorina* reveals a

- predicted linker as an integral part of the catalytic domain. *J Biol Chem.* 2017;292:19099-109.
173. Kruer-Zerhusen N, Alahuhta M, Lunin VV, Himmel ME, Bomble YJ, Wilson DB. Structure of a *Thermobifida fusca* lytic polysaccharide monooxygenase and mutagenesis of key residues. *Biotechnol Biofuels.* 2017;10:243.
 174. Eibinger M, Ganner T, Bubner P, Rosker S, Kracher D, Haltrich D, et al. Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J Biol Chem.* 2014;289:35929-38.
 175. Loose JS, Forsberg Z, Kracher D, Scheiblbrandner S, Ludwig R, Eijsink VG, et al. Activation of bacterial lytic polysaccharide monooxygenases with cellobiose dehydrogenase. *Protein Sci.* 2016;25:2175-86.
 176. Loose JS, Forsberg Z, Kracher D, Scheiblbrandner S, Ludwig R, Eijsink VG, et al. Activation of bacterial lytic polysaccharide monooxygenases with cellobiose dehydrogenase. *Protein Sci.* 2016;25:2175-86.
 177. Kracher D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K, et al. Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science.* 2016;352:1098-101.
 178. Tan TC, Kracher D, Gandini R, Sygmund C, Kittl R, Haltrich D, et al. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat Commun.* 2015;6:7542.
 179. Span EA, Marietta MA. The framework of polysaccharide monooxygenase structure and chemistry. *Curr Opin Struct Biol.* 2015;35:93-9.
 180. Johansen KS. Lytic polysaccharide monooxygenases: The microbial power tool for lignocellulose degradation. *Trends Plant Sci.* 2016;21:926-36.
 181. Walton PH, Davies GJ. On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Curr Opin Chem Biol.* 2016;31:195-207.
 182. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol.* 2011;6:1399-406.
 183. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, et al. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem.* 2014;289:2632-42.
 184. Westereng B, Arntzen MO, Aachmann FL, Varnai A, Eijsink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monooxygenases acting on cellulose. *J Chromatogr A.* 2016;1445:46-54.
 185. Bissaro B, Rohr AK, Muller G, Chylenski P, Skaugen M, Forsberg Z, et al. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat Chem Biol.* 2017;13:1123-8.
 186. Bey M, Zhou S, Poidevin L, Henrissat B, Coutinho PM, Berrin JG, et al. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospora anserina*. *Appl Environ Microbiol.* 2013;79:488-96.
 187. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc.* 2014;136:562-5.
 188. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A.* 2014;111:6287-92.
 189. Cannella D, Mollers KB, Frigaard NU, Jensen PE, Bjerrum MJ, Johansen KS, et al. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. *Nat Commun.* 2016;7:11134.
 190. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels.* 2016;9:186.
 191. Jagadeeswaran G, Gainey L, Prade R, Mort AJ. A family of AA9 lytic polysaccharide monooxygenases in *Aspergillus nidulans* is differentially regulated by multiple substrates and at least one is active on cellulose and xyloglucan. *Appl Microbiol Biotechnol.* 2016;100:4535-47.
 192. Kojima Y, Varnai A, Ishida T, Sunagawa N, Petrovic DM, Igarashi K, et al. A lytic polysaccharide monooxygenase with broad xyloglucan specificity from the brown-rot fungus *Gloeophyllum trabeum* and its action on cellulose-xyloglucan complexes. *Appl Environ Microbiol.* 2016;82:6557-72.

193. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Varnai A, et al. FgLPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Lett.* 2016;590:3346-56.
194. Fanuel M, Garajova S, Ropartz D, McGregor N, Brumer H, Rogniaux H, et al. The *Podospira anserina* lytic polysaccharide monooxygenase PaLPMO9H catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnol Biofuels.* 2017;10:63.
195. Ladeveze S, Haon M, Villares A, Cathala B, Grisel S, Herpoel-Gimbert I, et al. The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol Biofuels.* 2017;10:215.
196. Courtade G, Wimmer R, Rohr AK, Preims M, Felice AKG, Dimarogona M, et al. Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc Natl Acad Sci U S A.* 2016;113:5922-7.
197. Frandsen KE, Simmons TJ, Dupree P, Poulsen JC, Hemsworth GR, Ciano L, et al. The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat Chem Biol.* 2016;12:298-303.
198. Stricker AR, Mach RL, De Graaff LH. Regulation of transcription of cellulases-and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*). *Appl Microbiol Biotechnol.* 2008;78:211-20.
199. Berka RM, Grigoriev IV, Otilar R, Salamov A, Grimwood J, Reid I, et al. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol.* 2011;29:922-7.
200. Hinz SW, Pouvreau L, Joosten R, Bartels J, Jonathan MC, Wery J, et al. Hemicellulase production in *Chrysosporium lucknowense* C1. *J Cereal Sci.* 2009;50:318-23.
201. Frommhagen M, van Erven G, Sanders M, van Berkel WJH, Kabel MA, Gruppen H. RP-UHPLC-UV-ESI-MS/MS analysis of LPMO generated C4-oxidized gluco-oligosaccharides after non-reductive labeling with 2-aminobenzamide. *Carbohydr Res.* 2017;448:191-9.
202. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol.* 2017;102:1281-95.
203. Karnaouri A, Muraleedharan MN, Dimarogona M, Topakas E, Rova U, Sandgren M, et al. Recombinant expression of thermostable processive MteG5 endoglucanase and its synergism with MtLPMO from *Myceliophthora thermophila* during the hydrolysis of lignocellulosic substrates. *Biotechnol Biofuels.* 2017;10:126.
204. Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, et al. Enabling a community to dissect an organism: overview of the *Neurospora* functional genomics project. *Adv Genet.* 2007;57:49-96.
205. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, et al. Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol Mol Biol Rev.* 2004;68:1-108.
206. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature.* 2003;422:859-68.
207. Davis RH, Perkins DD. *Neurospora*: a model of model microbes. *Nat Rev Genet.* 2002;3:397-403.
208. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels.* 2012;5:79.
209. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol.* 2011;77:7007-15.
210. Cannella D, Chia-wen CH, Felby C, Jørgensen H. Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol Biofuels.* 2012;5:1-10.
211. Selig MJ, Vuong TV, Gudmundsson M, Forsberg Z, Westereng B, Felby C, et al. Modified cellobiohydrolase-cellulose interactions following treatment with lytic polysaccharide monooxygenase CelS2 (ScLPMO10C) observed by QCM-D. *Cellulose.* 2015;22:2263-70.
212. Vuong TV, Liu B, Sandgren M, Master ER. Microplate-based detection of lytic polysaccharide monooxygenase activity by fluorescence-labeling of insoluble oxidized products. *Biomacromolecules.* 2017;18:610-6.
213. Liao S, Pan B, Li H, Zhang D, Xing B. Detecting free radicals in biochars and determining their ability to inhibit the germination and growth of corn, wheat and rice seedlings. *Environ Sci Technol.* 2014;48:8581-7.

Chapter 2

Oxidized product profiles of AA9 lytic polysaccharide monooxygenases depend on the type of cellulose

Peicheng Sun¹, Susana V. Valenzuela^{2,3}, Pimvisuth Chunkruea¹, Francisco. I. Javier Pastor^{2,3}, Christophe V.F.P. Laurent^{4,5}, Roland Ludwig⁴, Willem J.H. van Berkel¹ and Mirjam A. Kabel¹

¹Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

²Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain

³Institute of Nanoscience and Nanotechnology (IN2UB), University of Barcelona, Barcelona, Spain

⁴Biocatalysis and Biosensing Laboratory, Department of Food Science and Technology, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

⁵Institute of Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

Abstract

Lytic polysaccharide monooxygenases (LPMOs) are essential for enzymatic conversion of lignocellulose-rich biomass in the context of biofuels and platform chemicals production. Considerable insight into the mode of action of LPMOs has been obtained, but research on cellulose specificity of these enzymes is still limited. Hence, we studied product profiles of four fungal Auxiliary Activity family 9 (AA9) LPMOs during their oxidative cleavage of three types of cellulose: bacterial cellulose (BC), Avicel® PH-101 (AVI), and regenerated amorphous cellulose (RAC). We observed that attachment of a carbohydrate binding module 1 (CBM1) did not change the substrate specificity of LPMO9B from *Myceliophthora thermophila* C1 (*Mt*LPMO9B), but stimulated the degradation of all three types of cellulose. A detailed quantification of oxidized ends in both soluble and insoluble fractions, and characterization of oxidized cello-oligosaccharide patterns, suggested that *Mt*LPMO9B generates mainly oxidized cellobiose from BC, while producing oxidized cello-oligosaccharides from AVI and RAC ranged more randomly from DP2–8. Comparable product profiles, resulting from BC, AVI and RAC oxidation, were found for three other AA9 LPMOs. These distinct cleavage profiles highlight cellulose-specificity rather than an LPMO dependent mechanism, and may further reflect that the product profiles of AA9 LPMOs are modulated by different cellulose types.

Keywords

Biomass; Biorefinery; Cellulose; Auxiliary Activity (AA); Lytic Polysaccharide Monooxygenase (LPMO); Carbohydrate Binding Module (CBM); Oxidized cello-oligosaccharide; Product profile

1. Introduction

Lignocellulose-rich biomass has been recognized as a sustainable source to produce fuels, chemicals and materials, which will, eventually, contribute to replacement of non-renewable fossil based products (1). A key step in this biorefinery concept is the degradation of abundantly present cell wall polysaccharides (e.g., cellulose and hemicellulose) into fermentable monomeric sugars (2). This widely studied process is optimal when using an enzymatic cocktail of (hemi-)cellulases and lytic polysaccharide monooxygenases (LPMOs) (3). LPMOs are copper-dependent enzymes and currently classified into sequence-based "Auxiliary Activity" families AA9–11 and AA13–16 in the Carbohydrate-Active enZymes (CAZy) database (<http://www.cazy.org>) (4). In this study we focus on LPMOs from the largest AA family (e.g., AA9). So far AA9 LPMOs are all fungal enzymes and active on cellulose (5). Although in the last decade much research has been conducted to disclose the catalytic mechanisms and structural features of AA9 LPMOs (6), insight into cellulose specificity (e.g., towards different cellulose types) and corresponding product profiles is still limited. Assigning such properties to individual LPMOs, and highlighting their specific product profiles, is a prerequisite to find the most appropriate candidates for envisaged applications.

AA9 LPMOs catalyze the hydroxylation of either C1- , C4- , or both C1- and C4-carbon positions (i.e. defining their regioselectivity) in cellulose using O₂ and/or H₂O₂ as co-substrate and an external electron donor (e.g., ascorbic acid) (6). Several studies have proposed that the regioselectivity for C1- or C4-oxidation depends on how LPMOs bind to their substrate (7). In addition, LPMOs can be connected to a carbohydrate binding module (CBM), and, as suggested in several studies, this might influence the regioselectivity of oxidation (5). It has been reported that roughly one fifth of AA9 LPMOs are fused to a C-terminal cellulose specific CBM1 via a flexible linker (8).

Apart from regioselectivity, AA9 LPMOs exhibit substrate specificity. Some AA9 LPMOs were reported to cleave xyloglucan, glucomannan, mixed β -(1→3, 1→4)-linked glucan, (cellulose-associated) xylan and even soluble cello-oligosaccharides, all in addition to cellulose (5). To understand if substrate specificity correlates with AA9 LPMO structural elements surrounding the active sites, a structure-based multiple sequence alignment and phylogenetic analysis have been performed for some AA9 LPMOs (9). Five segments surrounding the active site were linked to substrate recognition (9). Moreover, through this structure-function analysis, oxidative xyloglucan-active LPMOs, being tolerant or intolerant to xyloglucan substitutions, could be distinguished from cellulose-specific LPMOs (10). So far, details about the specificity of AA9 LPMOs towards various types of cellulose have remained elusive.

Cellulose is a homopolymer consisting of β -(1→4)-linked linear glucan chains (11). In plant cell walls, the linear glucan chains of cellulose are assembled via hydrogen bonds and van der Waals forces to form crystalline microfibrils (11). The chain length of cellulose can be expressed by the degree of polymerization (DP) (11). Depending on the source, the treatment and the assays used, DP values of cellulose vary from 300 to 15000 (12). Commercial cellulose is usually extracted and purified from

lignocellulose-rich biomass and the most widely applied type is microcrystalline cellulose (e.g., Avicel® PH-101 (AVI)) (13). Microcrystalline cellulose can be used to prepare other cellulosic substrates with different properties (e.g., low crystallinity) such as phosphoric acid swollen cellulose (PASC) and regenerated amorphous cellulose (RAC) (14, 15). In addition, bacteria are known to synthesize a type of cellulose (e.g., bacterial cellulose (BC)), which differs from plant cell wall-derived cellulose in degree of polymerization, crystallinity and morphology, as described elsewhere (15, 16).

As mentioned above, AA9 LPMO cleavage profiles, or product profiles, of different types of cellulose have not been studied in detail. One of the main reasons is that most analytical techniques are only suitable to analyze soluble compounds and cannot be used to analyze insoluble cellulose directly. For that reason, the LPMO catalytic action towards cellulose has mainly been monitored with imaging techniques, for example, atomic force microscopy (17). Although this technique unravels interesting physical changes of the substrate, it does not provide data at the molecular level needed to study cleavage profiles.

In this study, we used three AA9 LPMOs from *Myceliophthora thermophila* C1 (MtLPMOs) and one AA9 LPMO from *Neurospora crassa* (NcLPMO9M), which differ in the presence or absence of CBM1, regioselectivity and substrate specificity (Table S1). We analyzed their cellulose degradation profiles and detailed specificity towards AVI, RAC, and BC by quantifying formed oxidized ends over time (in supernatant and residual cellulose) and profiling the soluble oxidized cello-oligosaccharides. We found that four AA9 LPMOs generated mainly oxidized cellobiose from BC, while more evenly distributed mixtures of oxidized cello-oligosaccharide (e.g., DP2–8) were observed for AVI and RAC. The cellulose specificity and product profiles of four AA9 LPMOs were modulated by the type of cellulose rather than being LPMO type dependent.

2. Materials and methods

2.1. Carbohydrate standards, cellulose type and other chemicals

D-Glucose, D-gluconic acid (GlcOx[#]_1) and β -glucosidase from almond (9.3 U/mg, lyophilized powder) were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). Ascorbic acid (Asc) was purchased from VWR International (Radnor, Pennsylvania, USA). D-Cellobionic acid (GlcOx[#]_2) ammonium salt was purchased from Toronto Research Chemicals (Toronto, Ontario, Canada). Water used in all experiments was produced via a Milli-Q system (Millipore, Molsheim, France). Bacterial cellulose (BC) was produced by *Komagataeibacter xylinus* and prepared as described by Valenzuela et al. (15). Regenerated amorphous cellulose (RAC) was prepared from AVI (Avicel® PH-101, Sigma-Aldrich) as described previously (18, 19). Other carbohydrates were purchased from either Sigma-Aldrich or Megazyme (Bray, Ireland). Cellulase cocktails of Celluclast® 1.5 L and Accellerase® BG were obtained from Novozymes (Bagsværd, Denmark) and Genencor (Palo Alto, California, USA), respectively.

2.2. Enzyme production and purification

The genes encoding *MtLPMO9B* (*Mt9B*⁺, MYCTH_80312; UniProt ID, G2QCJ3) and the one with the truncated linker and CBM1 domain (*Mt9B*⁻), were homologously expressed in a low protease/(hemi-) cellulose producing *Myceliophthora thermophila* C1 strain (IFF Nutrition & Biosciences, Leiden, The Netherlands), as described elsewhere (20, 21). *Mt9B*⁺ and *Mt9B*⁻ were purified in three subsequent chromatographic steps (see Supplementary Information for more detail). The production and purification of *MtLPMO9I*, *MtLPMO9H* and *NcLPMO9M* are also described in the Supplementary Information.

2.3. Incubation of different cellulose types with AA9 LPMOs

2.3.1. HPSEC analysis for molecular weight distribution of (degraded) TXG

Three cellulose types (BC, AVI and RAC) were suspended in 50 mM ammonium acetate buffer (pH 5.0) in the absence or presence of 1 mM ascorbic acid (Asc, final concentration). Subsequently, *Mt9B*⁺ and *Mt9B*⁻ were added to the corresponding mixtures at a final concentration of 0.75 μ M. The time-course incubations (1, 2, 4, 6, 16 and 24 h) with BC and AVI were prepared in separate 2 mL Eppendorf tubes (Hamburg, Germany) with final volumes of 1 mL, while the time-course incubations with RAC were prepared in 15 mL Greiner tubes (Lake Forest, Illinois, USA) with final volumes of 3 mL. BC and AVI samples were incubated in the Eppendorf Thermomixer® Comfort at 800 rpm, placed in an almost vertical direction, and RAC samples were incubated in a head-over-tail disk rotator under 25 rpm at 30 or 50 °C. Incubations performed in either a thermomixer or in a head-over-tail disk rotator resulted in comparable outcomes (not shown). Control samples containing either cellulose and enzymes without Asc, or with 1 mM Asc were incubated accordingly in the same way. At each time point, a 0.5 mL RAC sample was taken out from the tube and BC/AVI samples were removed from the incubators. The incubation was stopped by separation of supernatant (SUP) from the residue (RES) directly after centrifugation at 22000 \times g for 10 minutes at 4 °C. SUP and RES of all time points were stored at -20 °C for further analysis. Another batch of BC and RAC end-point (24 h) incubations with *MtLPMO9I*, *MtLPMO9H* and *NcLPMO9M* was performed in the same way (only at 30 °C) as described above. All incubations were performed in duplicate and SUP was diluted five times prior to high performance anion exchange chromatography (HPAEC) analysis for oligosaccharide profiling.

2.4. Quantification of gluconic acid and cellobionic acid in the sample supernatant

To investigate the amounts of C1-oxidized ends in the supernatant of the samples, (C1-oxidized) cello-oligosaccharides were hydrolyzed by β -glucosidase to GlcOx#_1 and GlcOx#_2 which were quantified by using HPAEC. β -Glucosidase hydrolysis was performed by following a previously described method with the following modifications (22). β -Glucosidase was firstly dissolved in 62.5 mM ammonium acetate (pH 5.0)

buffer to give a 2.5 U/mL stock solution. Subsequently, 400 μ L β -glucosidase stock solution was mixed with 100 μ L SUP from each sample of time-course incubation of BC, AVI and RAC with *Mt9B*⁺ or *Mt9B*⁻, and control samples to a concentration of 1 U/mL. The reaction was incubated in an Eppendorf Thermomixer® Comfort at 800 rpm at 37 °C for 24 h. The amount of released GlcOx[#]_1 and GlcOx[#]_2 in SUP were quantified by using HPAEC with calibration curves of known concentrations (0–50 μ g/mL each) of GlcOx[#]_1 and GlcOx[#]_2. β -glucosidase hydrolyzed samples were diluted five times for HPAEC analysis.

2.5. Quantification of gluconic acid released from residual cellulose

To investigate the amounts of C1-oxidized ends in the residual cellulose, RES was hydrolyzed by a commercial cellulase cocktail to GlcOx[#]_1, which was quantified by using HPAEC. Cellulase cocktail hydrolysis of residual cellulose was carried out based on a previously described method (22). Celluclast® 1.5 L and Accellerase® BG were fractionated by size-exclusion chromatography (SEC) to discard fractions with impurities which disturb the HPAEC quantification of gluconic acid. The fractionation of Celluclast® 1.5 L and Accellerase® BG is described in the Supplementary Information. Cellulase cocktail stock solution was firstly prepared by mixing purified Celluclast® 1.5 L (final concentration of 2.5 mg/mL; 1.25 mg protein/mg residue) and Accellerase® BG (final concentration of 1 mg/mL; 0.5 mg protein/mg residue) with 50 ammonium acetate (pH 5.0) buffer. Subsequently, 500 μ L cellulase cocktail stock solution was mixed with RES from each sample of time-course incubation of BC, AVI and RAC with *Mt9B*⁺ or *Mt9B*⁻, and control samples. Due to the removal of SUP after centrifugation and addition of 500 μ L Asc-free cellulase cocktail stock solution, only trace amounts of Asc remained in RES, which is too low to drive the *Mt9B*⁺ and *Mt9B*⁻ reactions to confound the results of RES hydrolysis. The hydrolysis was incubated in an Eppendorf Thermomixer® Comfort at 800 rpm at 50 °C for 48 h. The amount of released GlcOx[#]_1 in RES was quantified by using HPAEC with calibration curves of known concentrations (0–50 μ g/mL) of GlcOx[#]_1 standard. Samples hydrolyzed by the cellulase cocktail were diluted ten times prior to HPAEC analysis.

2.6. HPAEC analysis for gluconic acid and cellobionic acid quantification and for oligosaccharide profiling

GlcOx[#]_1, GlcOx[#]_2 and (oxidized) cello-oligosaccharides were analyzed by HPAEC. The analysis was performed on an ICS-5000 system (Dionex, Sunnyvale, California, USA) equipped with a CarboPac PA-1 column (2 mm ID \times 250 mm; Dionex) in combination with a CarboPac PA guard column (2 mm ID \times 50 mm; Dionex). The system was equipped with pulsed amperometric detection (PAD). Mobile phases were (A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH. The column temperature was set to 20 °C and two elution programs were used. For the quantification of GlcOx[#]_1 (and GlcOx[#]_2), a 35 min elution program was used as described previously (22).

For profiling the (oxidized) cello-oligosaccharides, a 65 min elution profile was applied, as also described previously (18, 23). After HPAEC-PAD profiling, the peak area of each

DP of C1-oxidized cello-oligosaccharide present in SUP from all time-point incubations was manually integrated and recorded. Total peak area (calculated by the sum of all DPs) in each 24 h incubation sample was set as 100% and the percentage of each DP of C1-oxidized cello-oligosaccharide in other time-course incubated samples was expressed accordingly.

3. Results and discussion

3.1. Substrate specificity screening of *Mt9B*⁺ and *Mt9B*⁻

The substrate specificity of the purified *Mt9B*⁺ and *Mt9B*⁻ (Fig. S1) was screened with a wide range of carbohydrates (the experimental set-up is described in the Supporting information), and results are shown in **Table 1**. *Mt9B*⁺ and *Mt9B*⁻ were free of cellulase side-activity as shown in Fig. S2. In the presence of Asc, *Mt9B*⁺ and *Mt9B*⁻ produced a range of detectable C1-oxidized cello-oligosaccharides from all four types of cellulose (limited activity towards carboxymethyl cellulose. Data not shown). Since both *Mt9B*⁺ and *Mt9B*⁻ released C1-oxidized cello-oligosaccharides, it was concluded that the CBM1 had no effect on the regioselectivity of oxidation of *MtLPMO9B*. No activity of *Mt9B*⁺ or

Table 1. Carbohydrate activity screening of *Mt9B*⁺ and *Mt9B*⁻, in the presence of Asc.

Substrate	Oxidized oligosaccharides observed (+) or not (-) (in the presence of 1 mM Asc)			
	<i>Mt9B</i> ⁺		<i>Mt9B</i> ⁻	
	GlcOx_n ^a	HemiOx_n ^b	GlcOx_n ^a	HemiOx_n ^b
Cellulose				
BC	+	-	+	-
AVI	+	-	+	-
RAC	+	-	+	-
Carboxymethyl cellulose	+	-	+	-
Hemicellulose				
Xyloglucan (tamarind)	-	-	-	-
β-glucan (barley)	-	-	-	-
β-glucan (oat spelt)	-	-	-	-
Xylan (oat spelt)	-	-	-	-
Xylan (birchwood)	-	-	-	-
Arabinoxylan (wheat)	-	-	-	-
Mannan (acacia)	-	-	-	-
Galactan (potato)	-	-	-	-
Glucomannan (konjac)	-	-	-	-
Arabinan (sugar beet)	-	-	-	-
Laminarin (<i>Laminaria digitata</i>)	-	-	-	-
RAC/Hemicellulose combination				
RAC + Xyloglucan (tamarind)	+	-	+	-
RAC + Xylan (birchwood)	+	-	+	-
Oligosaccharides				
Cello-oligosaccharides (DP2–6)	-	-	-	-
Xylo-oligosaccharides (DP2–6)	-	-	-	-

^aGlcOx_n: oxidized cello-oligosaccharides

^bHemiOx_n: oxidized hemicello-oligosaccharides from the corresponding hemicellulose

$Mt9B^-$ was detected for any of the hemicellulosic substrates tested, not even in mixtures with RAC. Given the lack of activity on soluble cello-oligosaccharides, both $Mt9B^+$ and $Mt9B^-$ were concluded to be specifically active towards polymeric cellulosic structures. Similar results for $Mt9B^+$ were described by Frommhagen et al. (23), although in the current study a more extensive substrate screening was performed.

The influence of a CBM1 on the regioselectivity of oxidation has previously been investigated for several AA9 LPMOs (17, 24-26). Danneels et al. and Laurent and Sun et al. reported that removal of a CBM1 from *HjLPMO9A* and *NcLPMO9C* did not alter their regioselectivity towards cellulose (9, 25). In another study, although the regioselectivity was not changed, the ratio between C1- and C4-oxidized cello-oligosaccharides released by *PaLPMO9H* with and without a CBM1 was different (17). Little is known about the influence of the presence of a CBM on the substrate specificity of AA9 LPMOs. Only for *NcLPMO9C* it was reported that the substrate specificity did not change after truncation of its CBM1 (24). Thus, more detailed characterization is required for a better understanding of the catalytic performance of AA9 LPMOs with different cellulose substrates.

3.2. CBM1 promotes the *MtLPMO9B* cellulose degradation

As $Mt9B^+$ or $Mt9B^-$ showed only oxidative cleavage of different cellulose types, we further investigated their binding affinity (the experimental set-up is described in the Supplementary Information) and oxidative cleavage towards BC, AVI and RAC. As shown in Table S2, the amount of cellulose-bound $Mt9B^+$ was higher compared to $Mt9B^-$ in all BC, AVI and RAC samples at both 30 °C and 50 °C. This is in line with other LPMO studies where the presence of CBM1 resulted in more protein binding per gram of substrate (9, 27-30).

To quantify the oxidative cleavage of $Mt9B^+$ or $Mt9B^-$ over time, soluble (C1-oxidized) cello-oligosaccharides in SUP, and the insoluble (oxidized) cellulose in RES, were hydrolyzed by β -glucosidase and a cellulase cocktail, respectively. Subsequently, GlcOx[#]_1 and GlcOx[#]_2, were quantified to indicate the level of oxidation. A previous study from our laboratory showed that both GlcOx[#]_1 and GlcOx[#]_2 were released from cello-oligosaccharides by β -glucosidase and that only GlcOx[#]_1 was released by the cellulase cocktail (22). In addition, full hydrolysis of RES from BC and RAC was achieved. Although only ~67% of AVI were hydrolyzed under the same conditions (Fig. S3), it is still able to compare the results of AVI hydrolysis to results obtained from BC and RAC hydrolysis.

The quantification of the time-dependent oxidative cleavage of the different cellulose types by $Mt9B^+$ or $Mt9B^-$ is shown in **Fig. 1** and Fig. S4 for incubations at 30 °C and 50 °C, respectively. At 30 °C (**Fig. 1** and Table S3), $Mt9B^+$ produced a much higher amount of GlcOx[#]_1 & 2 from BC (up to ~89 μ g/mL) and AVI (up to ~19 μ g/mL) compared to the amount released by $Mt9B^-$ (~21 μ g/mL from BC and ~3 μ g/mL from AVI) in 24 h. For the RAC-digests, $Mt9B^-$ released higher amounts of oxidized products than $Mt9B^+$ in early time points (till 6 h) at 30 °C. After 6 h, $Mt9B^-$ did not release more GlcOx[#]_1 & 2 while $Mt9B^+$ continued generating GlcOx[#]_1 & 2. At 50 °C (Fig. S4

and Table S4), *Mt9B⁻* almost stopped releasing more GlcOx[#]_1 & 2 from all BC, AVI and RAC after 4 h, while *Mt9B⁺* still generated GlcOx[#]_1 & 2 (except AVI). Nevertheless, at 50 °C, the total amount of GlcOx[#]_1 & 2 in *Mt9B⁺* samples (~86 µg/mL from BC, ~6 µg/mL from AVI and ~56 µg/mL from RAC) was much higher compared to the total amount in *Mt9B⁻* samples (~18 µg/mL from BC, ~1 µg/mL from AVI and ~8 µg/mL from RAC), and after 24 h the difference was even larger than at 30 °C.

In summary, the oxidative cleavage by *MtLPMO9B* towards cellulose was modulated and influenced by the type of cellulose, CBM1 as well as temperature. Overall, the amount of GlcOx[#]_1 & 2 differed between the three types of cellulose, hinting at a different cellulose specific behavior. This result will be discussed further below. At an elevated temperature (50 °C), the cellulose specificity was still observed, however, based on the observation that oxidative cleavage stopped, the inactivation of *MtLPMO9B* was more pronounced compared to incubations at 30 °C. This was particularly striking for *Mt9B⁻* lacking CBM1. In addition, a larger difference in the amount of oxidized products obtained by *Mt9B⁺* and *Mt9B⁻* was observed after summation of the amounts in SUP and RES at 50 °C, compared to that at 30 °C. This further suggests that the CBM1 might stabilize or “help” the *MtLPMO9B* to act more pronounced and reduce the inactivation at an elevated temperature. A similar suggestion was reported for *BcLPMO10A*, a bacterial LPMO, for which it has been shown that removal of its CBM5 leads to fast enzyme inactivation, and, thus, decrease of oxidative cleavage (31). In addition to stabilization, it has been reported that CBM-driven substrate binding concentrates the LPMO on the substrate, which might further explain the observed higher amount of degradation products for *Mt9B⁺* compared to *Mt9B⁻* (26-30, 32).

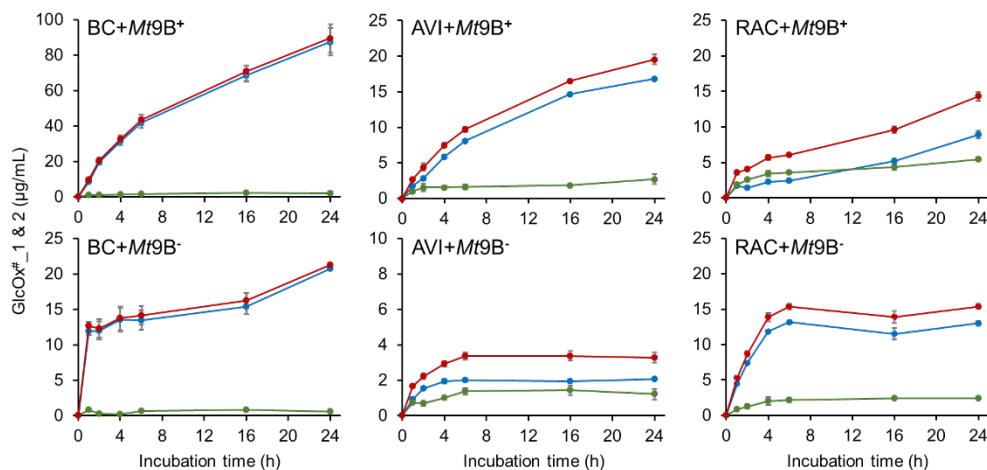


Fig. 1. Amount of gluconic acid (GlcOx[#]_1) and cellobionic acid (GlcOx[#]_2) generated from BC, AVI and RAC by *Mt9B⁺* and *Mt9B⁻* after subsequent hydrolysis, in supernatant (SUP, blue), residue (RES, green) and the sum of both (TOT, red) over time at 30 °C. Error bars (grey) indicate the standard deviations (\pm std) of duplicate incubations. Amount of GlcOx[#]_1 and GlcOx[#]_2 generated from BC, AVI and RAC by *Mt9B⁺* and *Mt9B⁻* after subsequent hydrolysis over time at 50 °C is shown in Figure S4.

Next, based on the results shown in **Fig. 1**, to further investigate the influence of cellulose types on the oxidized products released by LPMOs, the amounts of GlcOx[#]_1 & 2 present in SUP were calculated as percentages (%-Ox) of the total amount (TOT). This was done in order to mimic the parameter “percentage of soluble reducing sugar (%-Sugar)” used to describe the catalytic performance of cellulases in a filter paper assay (33-39). For both, *Mt9B*⁺ and *Mt9B*⁻, the %-Ox was higher than 95% in the BC digests, at 30 °C and 50 °C. However, the %-Ox decreased to approximately 85% and 62% in *Mt9B*⁺-AVI and -RAC digests (30 °C; 24 h), respectively, while at 50 °C it changed to approximately 75% and 80%, respectively. Likewise, in *Mt9B*⁻-AVI and -RAC digests, the %-Ox decreased to approximately 65% and 82% (30 °C; 24 h), respectively, while at 50°C *Mt9B*⁻-AVI and RAC digests were not representative due to the early inactivated *Mt9B*⁻. The difference of %-Ox in *Mt9B*⁺- and *Mt9B*⁻-AVI and RAC digests might relate to a more pronounced binding of CBM1 to crystalline cellulose than to RAC (40, 41). Courtade et al. observed a higher fraction of %-Ox in a full length ScLPMO10C-AVI digest compared to a CBM-truncated ScAA10-AVI digest (at comparable substrate concentration), which has been explained by the immobilizing effect of the CBM (28). This effect, as suggested by the authors, could keep the LPMO catalytic domain in a certain cellulose area, and, hereby, increases the chance for two (or more) cuts in the same cellulose chain (28). Indeed, the higher %-Ox from *Mt9B*⁺-AVI digest than *Mt9B*⁻-AVI digest might result from such immobilizing effect of CBM1. However, in contrast, product profiles obtained were similar for *Mt9B*⁺ and *Mt9B*⁻, and are further discussed in the later sections.

For cellulases, exo-cleavage and a processive catalytic action is considered if the parameter %-Sugar is higher than 90%, while 50%–70% reflects a more endo-cleavage and random-like action (35-37, 42). Due to the distinct structure and catalytic mechanisms compared to cellulases, LPMOs are not expected to act in a processive manner. However, the different %-Ox from BC, AVI and RAC still indicate that the catalytic performance of *MtLPMO9B* (both *Mt9B*⁺ and *Mt9B*⁻) is modulated by the type of cellulose. To gain more insight into the mode of cleavage of the three types of cellulose, corresponding product profiles were studied in detail.

3.3. Distinct product profiles of oxidized cello-oligosaccharide released from different cellulose types

Soluble oxidized cello-oligosaccharides formed at 24 h in BC, AVI and RAC (30 °C and 50 °C) were analyzed by HPAEC-PAD and the corresponding chromatograms are shown in **Fig. 2** and Fig. S5. Overall, non- and C1-oxidized cello-oligosaccharides (GlcOx[#]_n) were detected in the incubations of all three cellulose types with *Mt9B*⁺ and *Mt9B*⁻. However, the product profiles of C1-oxidized cello-oligosaccharides were different among BC, AVI and RAC samples. In both *Mt9B*⁺- and *Mt9B*⁻-BC samples, GlcOx[#]_2 was most pronounced followed by GlcOx[#]_3 and GlcOx[#]_4, at both 30°C and 50 °C (**Fig. 2** and Fig. S5). Only very low amounts of GlcOx[#]_5 to GlcOx[#]_8 were detected, indicating that both *Mt9B*⁺ and *Mt9B*⁻ formed mainly short oligosaccharides from BC. For the *Mt9B*⁺-AVI samples (**Fig. 2** and Fig. S5), again mainly GlcOx[#]_2–GlcOx[#]_4 were formed, but the ratio between GlcOx[#]_5–GlcOx[#]_8 and GlcOx[#]_2–GlcOx[#]_4

increased. At 30 °C and 50 °C, the amount of oxidized cello-oligosaccharides in *Mt9B*⁻-AVI sample was too low to see clear patterns. In comparison to the BC and AVI samples, a more even distribution pattern of GlcOx[#]_2–GlcOx[#]_8 products was observed in *Mt9B*⁺- and *Mt9B*⁻-RAC samples.

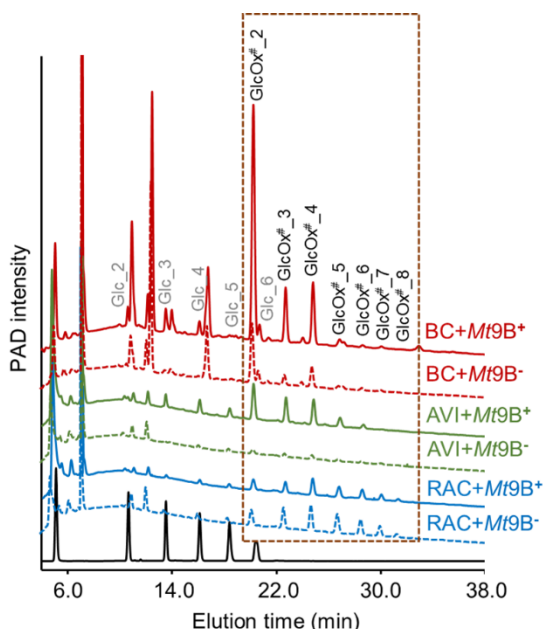


Fig. 2. HPAEC elution patterns of supernatants (SUP) from BC, AVI and RAC digests (24 h incubation) of *Mt9B*⁺ and *Mt9B*⁻ in the presence of Asc at 30 °C. HPAEC elution patterns of supernatants generated at 50 °C are shown in Fig. S5. HPAEC chromatogram from one of the duplicate samples is shown here, as they are identical. Annotation of non- (Glc_2 to Glc_6) and C1-oxidized cello-oligosaccharides (GlcOx[#]_2 to GlcOx[#]_8) is based on the previous study (18, 23). In this study, C1-oxidized cello-oligosaccharides are the most relevant and their elution range is located in a brown frame. Due to the presence of carboxylic acid end in the C1-oxidized cello-oligosaccharides, they bind stronger to the HPAEC column and thus are eluted later in HPAEC compared to non-oxidized cello-oligosaccharides. A standard containing a mixture of Glc_1 to Glc_6 (from left to right in chromatogram) is shown in black. SUP of control incubations is shown in Fig. S2.

To further investigate the product profiles of BC, AVI and RAC digests, we quantified each DP of oxidized cello-oligosaccharides formed over time. Due to the lack of GlcOx[#]_3–GlcOx[#]_8 standards, quantification was based on the peak area of each DP and expressed as the percentage of total peak area of oxidized cello-oligosaccharides from the corresponding 24 h sample (**Fig. 3** and Fig. S6). In line with the previously described patterns, both *Mt9B*⁺ and *Mt9B*⁻ predominantly released GlcOx[#]_2 (>60%) followed by GlcOx[#]_4 (~20%) and GlcOx[#]_3 (~10–15%) from BC over time at 30 °C (**Fig. 3**).

In the AVI samples, GlcOx[#]_2 and GlcOx[#]_4 reflected the main products, however, the proportion of larger GlcOx[#] products was higher than in the BC samples (**Fig. 3**). In the RAC samples, the percentages of GlcOx[#]_2 and GlcOx[#]_4 were the lowest at all-time points (**Fig. 3**). Additionally, compared to the BC and AVI samples, the proportion of higher DP oxidized cello-oligosaccharides (DP5–8) increased in the RAC samples, while the AVI-based product profiles represent an intermediate situation. As described in the previous section, the immobilizing effect of a CBM is also expected to result in smaller oxidized products, as shown by higher percentage of GlcOx[#]_2–GlcOx[#]_4 in the full-length ScLPMO-AVI digest compared to the CBM-truncated ScAA10-AVI digest (28). In that study the authors suggest that, when a CBM is present, the chance of multiple cleavages in the same cellulose chain is higher, and thus, shorter oxidized cello-oligosaccharides can be expected (28). However, our data

do not show such difference in product profiles for the full length and CBM-truncated LPMO used, and, hence, we cannot conclude that product profiles are CBM dependent.

These distinct profiles from BC, AVI and RAC samples at 30 °C were found to be similar in the BC, AVI and RAC digests with *Mt9B*⁺ and *Mt9B*⁻ at 50 °C (Fig. S6), though the amounts of each DP were different compared to the samples at 30 °C.

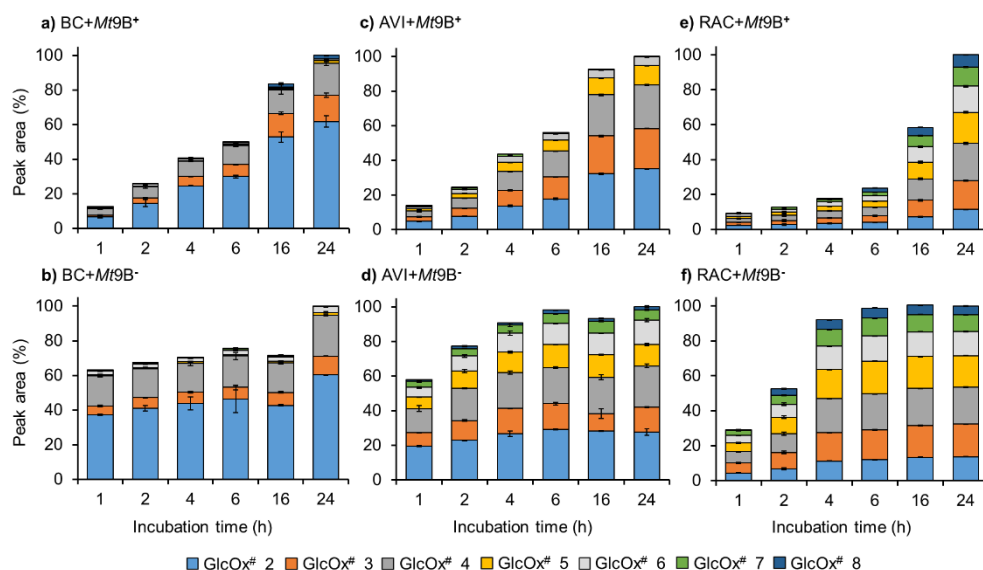


Fig. 3. Relative quantification of peak area of each DP of oxidized cello-oligosaccharides (GlcOx#_2 to GlcOx#_8) generated by *Mt9B*⁺ and *Mt9B*⁻ from the time-course incubation with BC, AVI and RAC at 30 °C. Total peak area of each 24 h sample was set to 100%. Relative quantification of peak area of released oxidized cello-oligosaccharides at 50 °C is shown in Fig. S6. Error bars for each DP of oxidized cello-oligosaccharides indicate the standard deviations (\pm std) of duplicate incubations.

To further substantiate that the mode of cleavage may relate to the type of cellulose, rather than to the type of LPMO, cellulose-digests of three other AA9 LPMOs were compared to those from *MtLPMO9B*. These three others (Table S1) were the previously characterized C1-oxidizing *MtLPMO9I* (no CBM) (18), C1-/C4-oxidizing *MtLPMO9H* (having a CBM1) (43), and C1-/C4-oxidizing *NcLPMO9M* (no CBM) (44), BC and RAC digests with these three LPMOs (24 h and 30 °C) were analyzed by HPAEC (**Fig. 4**). Similar to the product profiles in *MtLPMO9B*-BC digests, *MtLPMO9I* generated mainly short oxidized cello-oligosaccharides (DP2–4) from BC, while all DPs of oxidized cello-oligosaccharides were present in a more evenly distributed pattern in the RAC sample (**Fig. 4**). In *MtLPMO9H*- and *NcLPMO9M*-BC digests, short (C4-oxidized) cello-oligosaccharides (GlcOx*_n-GlcOx*_n+2) were predominant. Again, more even distribution profiles (of C4-oxidized cello-oligosaccharides) were seen in their RAC-digests. The concentration of C1-oxidized products released by *MtLPMO9H* and *NcLPMO9M* were too low to observe a clear cleavage pattern. These results indicate that not only *MtLPMO9B* (with and without a CBM1), but also other AA9 LPMOs generate distinct cellulose degradation product profiles ranging from mainly oxidized

cellobiose towards BC, to a more even distribution towards RAC. In addition, estimated from their peak areas, *MtLPMO9I* and *NcLPMO9M* released the highest quantities of oxidized products from RAC, whereas CBM1-containing *MtLPMO9H* released most oxidized products from BC.

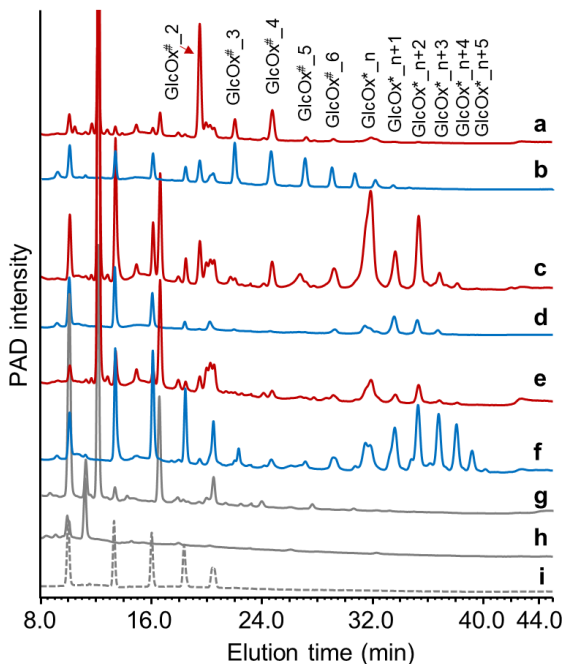


Fig. 4. HPAEC elution patterns of supernatants (SUP) from the incubation of BC (red) and RAC (blue) with *MtLPMO9I* (a and b), *MtLPMO9H* (c and d) and *NcLPMO9M* (e and f). SUP of control incubations (BC+Asc (g) and RAC+Asc (h)) and cello-oligosaccharide standard (i) are shown in gray solid lines and dashed line, respectively.

3.4. Proposed scenario of *MtLPMO9B* in degrading different cellulose types

As described above, the AA9 LPMO cellulose degradation profiles were mainly dependent on the type of cellulose used. BC, AVI and RAC have been reported to vary in surface area, crystallinity, DP and three-dimensional structure (13, 16, 45-53). In general, BC and AVI have a similarly high crystallinity (15, 16, 45); the crystallinity index (CrI (%)) of BC and AVI was determined as 95.7 ± 0.5 and 92.7 ± 1.1 , respectively (15). RAC has been shown to be completely amorphous due to the high concentration of phosphoric acid (86.2% wt/v) used to produce RAC (14). In a previous study, CrI (%) of RAC was determined in a dried state (67.4 ± 1.6) explaining the rather high value obtained (i.e., due to recrystallization during the drying process) (15). Compared to AVI, BC has a larger surface area and, therefore, has a higher accessibility (16, 45, 54, 55). Further, BC consists of long ribbon-like microfibrils with DPs ranging from 2000 to 6000, while AVI microfibrils are shorter and thicker with a much lower DP (100–300) (13, 16, 49-53). In addition, BC resembles a more well-arranged network compared to AVI (15, 16, 45). Gromovikh et al. suggested that the BC network forms three-dimensional layers of hollow cylinders, and each layer turns a small angle (56). For RAC, no typical DP lengths have been reported, but it can be expected that the DP of RAC-chains are shorter and more exposed compared to AVI

seen the process conditions (e.g., use of phosphoric acid) to produce RAC from AVI (14, 57).

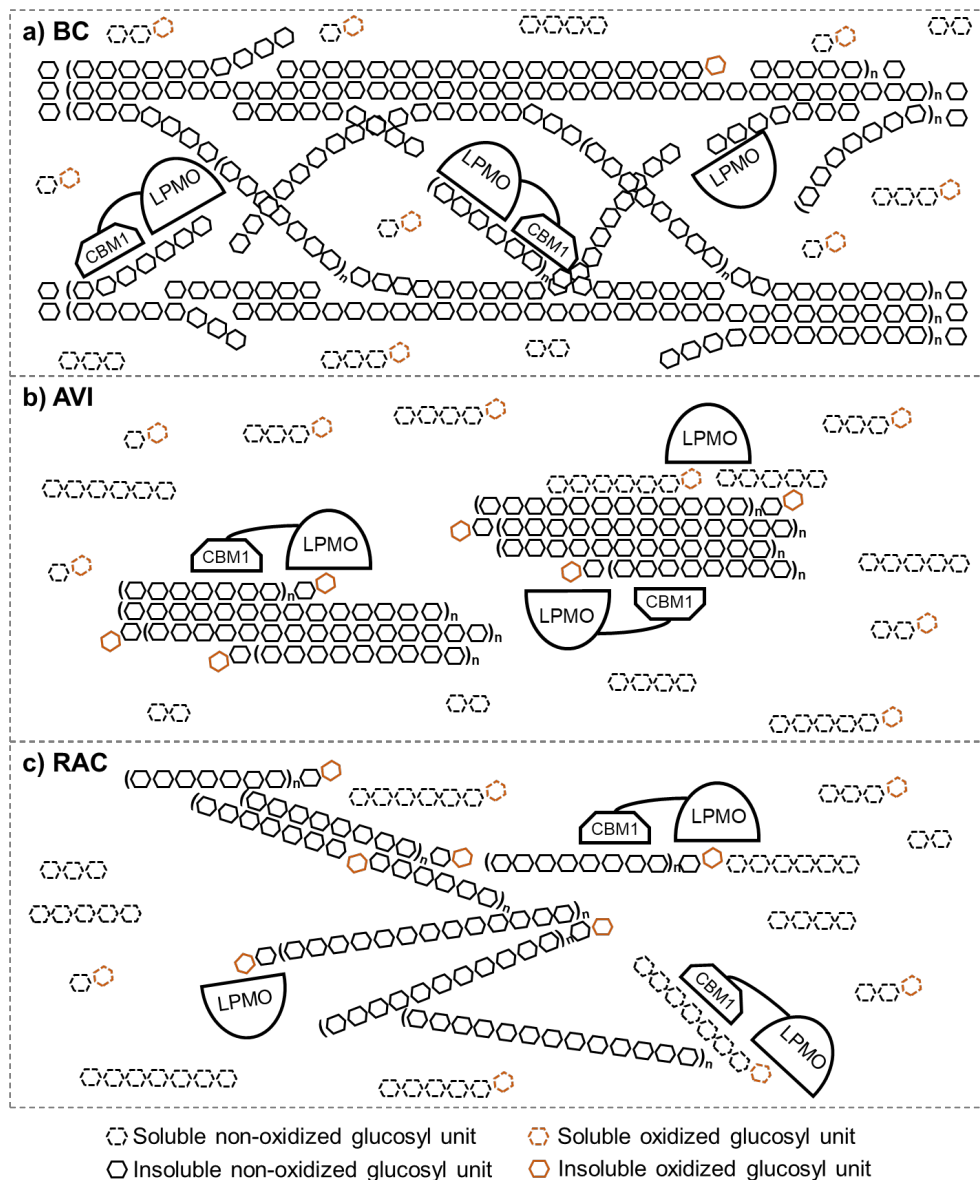


Fig. 5. Schematic representation of the proposed scenario of *MtlPMO9B* oxidative cleavage towards BC (A), AVI (B) and RAC (C). "n" indicates the number of repeating units. It should be noted that **Fig. 5** presents a generic and schematic representation in two-dimensions, and might oversimplify how LPMOs cleave within cellulose chains.

Based on the distinct characteristics of BC, AVI and RAC, together with our results, we propose different scenarios of how *MtlPMO9B* oxidatively cleaves various cellulose

types, as schematically depicted in **Fig. 5**. Based on the model suggested by Gromovykh et al (56), we propose that *MtLPMO9B* mainly cleaves the “connecting” region between layers. In between layers more chain ends can be expected, compared to the “layer” region, which might explain the pronounced formation of (oxidized) cellobiose (**Fig. 5a**). This scenario corresponds to the product profiles found in the BC digests, mainly reflecting formation of oxidized DP2–4 cello-oligosaccharides (oxidized cellobiose >60%) (**Fig. 2–4**) and an %-Ox of over 95% (**Fig. 1**). Next, we propose a scenario of the *MtLPMO9B* towards AVI (**Fig. 5b**). Due to the lower homogeneity and shorter chain lengths, compared to BC, the *MtLPMO9B* has less chance to stay active on AVI, resulting in formation of larger DPs of oxidized cello-oligosaccharides (**Fig. 2–4**), and an %-Ox of 60% to 80% (**Fig. 1**). Still, oxidized DP2–4 cello-oligosaccharides were most pronounced in these AVI-digests (**Fig. 2–4**). For the amorphous and homogenous RAC with more exposed glucan chains, we suggest that the LPMO has more chance to cleave in the middle of RAC-chains compared to the packed fibrous structure of BC and AVI (**Fig. 5c**). Due to the lower DP of RAC compared to the DP of AVI, the polymeric cellulose chain more easily becomes soluble. This would explain why a more evenly distributed oxidized cello-oligosaccharide profile (**Fig. 2–4**) is observed from the RAC digests compared to the BC and AVI digests.

The proposed scenario is mainly based on quantification data of oxidized ends, oxidized cello-oligosaccharide profiles, and morphological properties of the used cellulose types. Other factors, for example, how long LPMOs (i.e., with or without CBM) stay on the substrate are also important to consider, but not included in our model.

4. Conclusions

In this study, we compared several AA9 LPMOs for their reactivity with different types of cellulose and found that the substrate specificity and regioselectivity of the cleavage site were not altered by the presence of a CBM1. We also found that the CBM1 increased the release of oxidized cello-oligosaccharides by *MtLPMO9B*, especially at the elevated temperature. This increased release corresponds an increased binding affinity towards the substrates due to the presence of a CBM1. Intriguingly, the length of the released cello-oligosaccharide was dependent on the characteristics of the cellulose type. From BC mainly oxidized cellobiose was released, regardless the presence of CBM1, while from RAC and AVI a more evenly distributed mixture of oxidized cello-oligosaccharides (DP2–8) was obtained. Our study highlights the importance of considering biopolymeric substrate characteristics when cleavage profiles and kinetics of AA9 LPMOs are studied.

5. References

1. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, et al. The path forward for biofuels and biomaterials. *Science*. 2006;311:484-9.
2. Kim J-Y, Lee HW, Lee SM, Jae J, Park Y-K. Overview of the recent advances in lignocellulose liquefaction for producing biofuels, bio-based materials and chemicals. *Bioresour Technol*. 2019;279:373-84.
3. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.

4. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014;42:D490-D495.
5. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol.* 2018;9:1080.
6. Chylenski P, Bissaro B, Sorlie M, Rohr AK, Varnai A, Horn SJ, et al. Lytic polysaccharide monooxygenases in enzymatic processing of lignocellulosic biomass. *ACS Catal.* 2019;9:4970-91.
7. Vaaje-Kolstad G, Forsberg Z, Loose JS, Bissaro B, Eijsink VG. Structural diversity of lytic polysaccharide monooxygenases. *Curr Opin Struct Biol.* 2017;44:67-76.
8. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry.* 2010;49:3305-16.
9. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci.* 2019;20:6219.
10. Sun P, Laurent CVFP, Scheiblbrandner S, Frommhagen M, Kouzounis D, Sanders MG, et al. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnol Biofuels.* 2020;13:95.
11. Klemm D, Heublein B, Fink HP, Bohn A. Cellulose: Fascinating biopolymer and sustainable raw material. *Angew Chem Int Edit.* 2005;44:3358-93.
12. Gupta PK, Raghunath SS, Prasanna DV, Venkat P, Shree V, Chithananthan C, et al. An update on overview of cellulose, its structure and applications. *Cellulose.* 2019:846-1297.
13. Kambli ND, Mageshwaran V, Patil PG, Saxena S, Deshmukh RR. Synthesis and characterization of microcrystalline cellulose powder from corn husk fibres using biochemical route. *Cellulose.* 2017;24:5355-69.
14. Zhang YHP, Cui JB, Lynd LR, Kuang LR. A transition from cellulose swelling to cellulose dissolution by O-phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules.* 2006;7:644-8.
15. Valenzuela SV, Valls C, Schink V, Sanchez D, Roncero MB, Diaz P, et al. Differential activity of lytic polysaccharide monooxygenases on celluloses of different crystallinity. Effectiveness in the sustainable production of cellulose nanofibrils. *Carbohydr Polym.* 2019;207:59-67.
16. Choi SM, Shin EJ. The nanofication and functionalization of bacterial cellulose and its applications. *Nanomaterials-Basel.* 2020;10:406.
17. Chalak A, Villares A, Moreau C, Haon M, Grisel S, d'Orlando A, et al. Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol Biofuels.* 2019;12:206.
18. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym.* 2020;234:115917.
19. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels.* 2015;8:101.
20. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Industrial Biotechnology.* 2011;7:214-23.
21. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al., *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
22. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol.* 2017;102:1281-95.
23. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels.* 2016;9:186.
24. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem.* 2015;290:22955-69.

25. Danneels B, Tanghe M, Desmet T. Structural features on the substrate-binding surface of fungal lytic polysaccharide monooxygenases determine their oxidative regioselectivity. *Biotechnol J*. 2019;14:1800211.
26. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels*. 2015;8:90.
27. Forsberg Z, Nelson CE, Dalhus B, Mekasha S, Loose JS, Crouch LI, et al. Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J Biol Chem*. 2016;291:7300-12.
28. Courtade G, Forsberg Z, Heggset EB, Eijsink VGH, Aachmann FL. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J Biol Chem*. 2018;293:13006-15.
29. Mekasha S, Tuveng TR, Askarian F, Choudhary S, Schmidt-Dannert C, Niebisch A, et al. A trimodular bacterial enzyme combining hydrolytic activity with oxidative glycosidic bond cleavage efficiently degrades chitin. *J Biol Chem*. 2020;295:9134-46.
30. Kruer-Zerhusen N, Alahuhta M, Lunin VV, Himmel ME, Bomble YJ, Wilson DB. Structure of a *Thermobifida fusca* lytic polysaccharide monooxygenase and mutagenesis of key residues. *Biotechnol Biofuels*. 2017;10:243.
31. Mutahir Z, Mekasha S, Loose JS, Abbas F, Vaaje-Kolstad G, Eijsink VG, et al. Characterization and synergistic action of a tetra-modular lytic polysaccharide monooxygenase from *Bacillus cereus*. *FEBS Lett*. 2018;592:2562-71.
32. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, et al. High-resolution structure of a lytic polysaccharide monooxygenase from *Hypocrea jecorina* reveals a predicted linker as an integral part of the catalytic domain. *J Biol Chem*. 2017;292:19099-109.
33. Zhou W, Irwin DC, Escovar-Kousen J, Wilson DB. Kinetic studies of *Thermobifida fusca* Cel9A active site mutant enzymes. *Biochemistry*. 2004;43:9655-63.
34. Horn SJ, Sorlie M, Varum KM, Valjamae P, Eijsink VGH. Measuring processivity. *Method Enzymol*. 2012;510:69-95.
35. Jeon SD, Yu KO, Kim SW, Han SO. The processive endoglucanase EngZ is active in crystalline cellulose degradation as a cellulosomal subunit of *Clostridium cellulovorans*. *New Biotechnol*. 2012;29:365-71.
36. Li YC, Irwin DC, Wilson DB. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. *Appl Environ Microbiol*. 2007;73:3165-72.
37. Vuong TV, Wilson DB. Processivity, synergism, and substrate specificity of *Thermobifida fusca* Cel6B. *Appl Environ Microbiol*. 2009;75:6655-61.
38. Wilson DB, Kostylev M. Cellulase processivity. *Biomass Conversion: Springer*; 2012. p. 93-9.
39. Wu SS, Wu SF. Processivity and the mechanisms of processive endoglucanases. *Appl Biochem Biotechnol*. 2020;190:448-63.
40. Blake AW, McCartney L, Flint JE, Bolam DN, Boraston AB, Gilbert HJ, et al. Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes. *J Biol Chem*. 2006;281:29321-9.
41. Lehtö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc Natl Acad Sci U S A*. 2003;100:484-9.
42. Irwin DC, Spezio M, Walker LP, Wilson DB. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. *Biotechnol Bioeng*. 1993;42:1002-13.
43. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
44. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136:562-5.
45. Zhang YH, Lynd LR. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol Bioeng*. 2004;88:797-824.
46. Saxena IM, Brown RM. Cellulose biosynthesis: Current views and evolving concepts. *Ann Bot*. 2005;96:9-21.
47. Okiyama A, Motoki M, Yamanaka S. Bacterial cellulose. 3. Development of a new form of cellulose. *Food Hydrocoll*. 1993;6:493-501.

48. Watanabe K, Tabuchi M, Morinaga Y, Yoshinaga F. Structural features and properties of bacterial cellulose produced in agitated culture. *Cellulose*. 1998;5:187-200.
49. Shibazaki H, Kuga S, Okano T. Mercerization and acid hydrolysis of bacterial cellulose. *Cellulose*. 1997;4:75-87.
50. Mormann W, Demeter J, Wagner T. Partial silylation of cellulose with predictable degree of silylation–stoichiometric silylation with hexamethyldisilazane in ammonia. *Macromol Chem Phys*. 1999;200:693-7.
51. Kumar V, Kothari S, Banker GS. Effect of the agitation rate on the generation of low-crystallinity cellulose from phosphoric acid. *J Appl Polym Sci*. 2001;82:2624-8.
52. Hubbell CA, Ragauskas AJ. Effect of acid-chlorite delignification on cellulose degree of polymerization. *Bioresour Technol*. 2010;101:7410-5.
53. Matsuoka S, Kawamoto H, Saka S. What is active cellulose in pyrolysis? An approach based on reactivity of cellulose reducing end. *J Anal Appl Pyrolysis*. 2014;106:138-46.
54. Gayathry G, Gopalaswamy G. Production and characterisation of microbial cellulosic fibre from *Acetobacter xylinum*. *Indian J Fibre Text*. 2014;39:93-6.
55. Boisset C, Fraschini C, Schulein M, Henrissat B, Chanzy H. Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from *Humicola insolens* and its mode of synergy with cellobiohydrolase Cel7A. *Appl Environ Microbiol*. 2000;66:1444-52.
56. Gromovkykh TI, Pigaleva MA, Gallyamov MO, Ivanenko IP, Ozerova KE, Kharitonova EP, et al. Structural organization of bacterial cellulose: The origin of anisotropy and layered structures. *Carbohydr Polym*. 2020;237:116140.
57. Hong J, Ye XH, Zhang YHP. Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications. *Langmuir*. 2007;23:12535-40.

6. Supplementary information

6.1. Supplementary materials and methods

6.1.1. Purification of *Mt9B*⁺ and *Mt9B*⁻

Mt9B⁺ and *Mt9B*⁻ were purified in three subsequent chromatographic steps. The initial anion exchange chromatography (AEC) and cation exchange chromatography (CEC) steps were the same as described previously (1). The final CEC was performed with the following modifications. *Mt9B*⁺ and *Mt9B*⁻ obtained after the first CEC step were diluted two times with water before subjecting them to a Resource S column (30 × 16 mm internal diameter, GE Healthcare, Uppsala, Sweden) using an ÄKTA-Micro chromatography system (GE Healthcare). The column was pre-equilibrated with 10 mM sodium acetate buffer (pH 5.0). By using a flow rate of 1 mL/min, the unbound fraction was washed (three column volumes) and elution was performed by using a linear gradient from 0 to 0.5 M NaCl in 10 mM sodium acetate buffer pH 5.0 over one column volume. All unbound and eluted fractions (size 1 mL) were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to a concentration of 2 mg/mL (determined by using Pierce™ BCA Protein Assay Kit (Sigma-Aldrich) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) as described previously (2). The fractions containing either pure *Mt9B*⁺ or pure *Mt9B*⁻ were used for enzyme incubations.

6.1.2. Production and purification of *MtLPMO9I*, *MtLPMO9H* and *NcLPMO9M*

The genes encoding *MtLPMO9I* (MTCTH_2299721, UniProt ID: G2Q774) and *MtLPMO9H* (MYCTH_46583, UniProt ID: G2Q9T3) were homologously expressed in a low protease/low (hemi-) cellulose producing *Myceliophthora thermophila* C1 strain (IFF Nutrition & Biosciences, Leiden, The Netherlands), as described elsewhere (3, 4). The purification of *MtLPMO9I* has been described previously (5).

MtLPMO9H was purified by three subsequent chromatographic steps. Crude *MtLPMO9H*-rich fermentation broth was filtrated and dialyzed against 10 mM potassium phosphate buffer pH 7.6 before chromatographic purification. The dialyzed *MtLPMO9H* was purified by AEC followed by size exclusion chromatography (SEC). Purification settings and elution program of AEC and SEC have been described previously (5). The SEC-purified *MtLPMO9H*-containing fractions were further purified by CEC on an ÄKTA-Micro preparative chromatography system (GE Healthcare). *MtLPMO9H*-containing fractions were loaded on a Resource Q column (30 × 16 mm internal diameter, GE Healthcare) pre-equilibrated with 10 mM sodium acetate buffer pH 4.5 (eluent A). The unbound fraction was firstly washed (one column volume). Eluent B was 10 mM sodium acetate buffer (pH 4.5) containing 500 mM NaCl (eluent B). Elution (flow rate of 1 mL/min) was performed as follows: from 0% to 25% B in two column volumes; 25% B for one column volume; next 25% to 100% B over two column volumes and finally 100% B for four column volumes. All fractions were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to an approximate

concentration of 2 mg/mL determined by BCA assay and analyzed by SDS-PAGE, as described previously (2) to determine the *MtLPMO9H* fractions. CEC-purified *MtLPMO9H*-containing fractions were combined and used as final enzyme stock solution. Production and purification of *NcLPMO9M* have been described elsewhere (6, 7).

6.1.3. Fractionation of Celluclast® 1.5 L and Accellerase® BG

Commercial enzyme cocktails of Celluclast® 1.5 L and Accellerase® BG were fractionated by using SEC in an ÄKTA-Micro chromatography system (GE Healthcare). Prior to the fractionation, Celluclast® 1.5 L and Accellerase® BG were washed three times with a 20 mM potassium phosphate buffer (pH 7.0) containing 150 mM NaCl using Amicon® ultra centrifugal filters (3 kDa cut-off, Sigma-Aldrich). Enzyme samples were loaded on a Superdex™ 75 Increase 10/300 GL column (100 × 3 mm diameter, GE Healthcare) and eluted (1 mL/min) with a 20 mM potassium phosphate buffer (pH 7.0) containing 150 mM NaCl. Peak fractions (based on UV 280 nm) of each enzyme cocktail were collected, combined and concentrated via ultrafiltration. The fractionated Celluclast® 1.5 L and Accellerase® BG were dialyzed with 50 mM sodium acetate buffer (pH 5.0) and subsequently stored at 4 °C.

6.1.4. Screening of substrate specificities of *Mt9B*⁺ and *Mt9B*⁻

All carbohydrate substrates and their tested combinations are listed in **Table 1**. Substrates (final concentration of 2 mg/mL each) were dissolved or suspended in 50 mM ammonium acetate buffer (pH 5.0) in the absence or presence of 1 mM ascorbic acid (Asc, final concentration). Subsequently, *Mt9B*⁺ and *Mt9B*⁻ were added to the corresponding mixtures (0.5 mL total volume) at a final concentration of 0.75 µM. Control samples either contained substrates and enzymes without Asc, or the different cellulose types with 1 mM Asc. The reaction was incubated in an Eppendorf Thermomixer® Comfort at 30 °C under shaking at 800 rpm, placed in an almost vertical direction, for 24 h. The incubation was stopped by separation of supernatant (SUP) from the residue (RES) directly after centrifugation at 22000 × *g* for 10 min at 4 °C. SUP was collected and stored at -20 °C. SUP was diluted five times prior to high performance anion exchange chromatography (HPAEC) analysis.

6.1.5. Binding of *Mt9B*⁺ and *Mt9B*⁻ to different cellulose types

Binding experiments were performed according to Petrović et al. (8) and Courtade et al. (9). 5 µM *Mt9B*⁺ or *Mt9B*⁻ were added to 2 mg/mL BC, AVI or RAC in 50 mM ammonium acetate buffer pH 5.0. The protein-cellulose mixture was incubated at 30 or 50 °C in an Eppendorf ThermoMixer® Comfort (Eppendorf, Hamburg, Germany) at 800 rpm (in an almost vertical orientation) for 240 min. Subsequently, supernatant with free protein was separated and collected from residual cellulose with bound protein through centrifugation 22,000 × *g*, 2 min at 4 °C in a table centrifuge. The amount of bound protein was calculated based on the following equation (eq.1):

$$\text{Eq. 1} \quad \text{Bound protein } (\mu\text{M}) = \text{Total protein } (\mu\text{M}) - \text{Total protein } (\mu\text{M}) \times \frac{\text{Abs free (AU 280)}}{\text{Abs total (AU 280)}}$$

where *Total protein* = 5 μM ; *Abs free* = absorbance of free protein in the supernatant at 280 nm; *Abs total* = absorbance of total protein solution without substrate at 280 nm.

The absorbance of free protein and total protein was measured spectrophotometrically (SpectraMax M2e/iD3, Molecular Devices, USA) at 280 nm in a 96-well plate. Both the incubations and the measurements were performed in duplicate.

6.2. Supplementary data

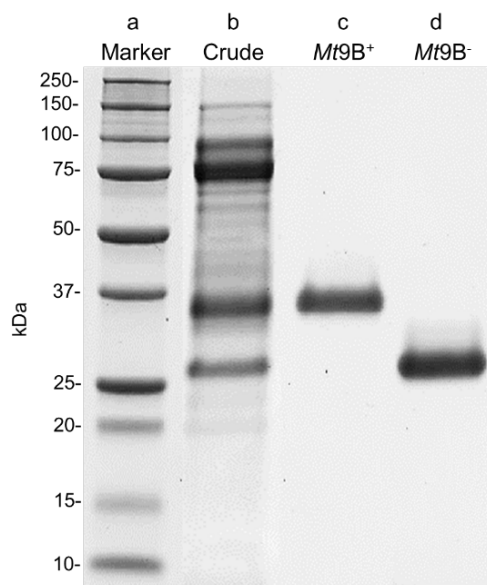


Fig. S1. SDS-PAGE analysis of crude *Mt9B*⁺ and *Mt9B*⁻-containing fraction (lane b) and purified *Mt9B*⁺ (lane c) and *Mt9B*⁻ (lane d). The molecular masses (kDa) of marker proteins (a) are indicated.

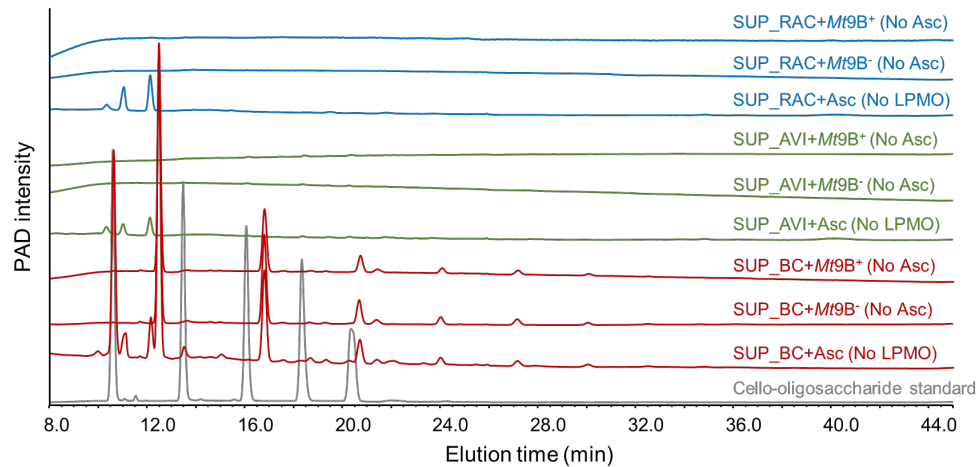


Fig. S2. HPAEC elution patterns of control samples. In the absence of Asc, the supernatants from the incubation (24 h and 30 °C) of *Mt9B*⁺ and *Mt9B*⁻ with BC (SUP_BC+*Mt9B*⁺ (no Asc) and SUP_BC+*Mt9B*⁻ (no Asc)), AVI (SUP_AVI+*Mt9B*⁺ (no Asc) and SUP_AVI+*Mt9B*⁻ (no Asc)) and BC (SUP_RAC+*Mt9B*⁺ (no Asc) and SUP_RAC+*Mt9B*⁻ (no Asc)) were analyzed. The supernatants from the incubation of BC, AVI and RAC without enzymes but with Asc (SUP_BC+Asc (no LPMO), SUP_AVI+Asc (no LPMO) and (SUP_BC+Asc (no LPMO))) were also analyzed. A standard containing a mixture of cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose (from left to right in chromatogram) is shown in gray. HPAEC elution patterns of control samples incubated at 30 °C and 50 °C were identical.

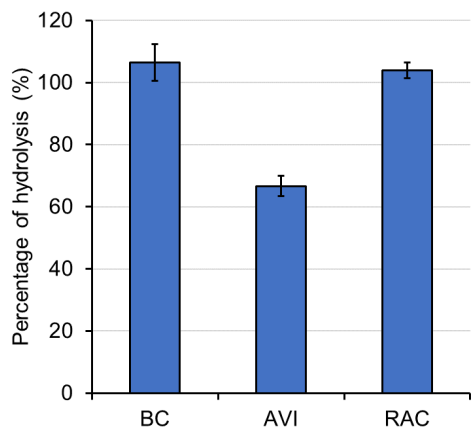


Fig. S3. Percentage of hydrolysis of BC, AVI and RAC by using the fractionated cellulase cocktail of Celluclast® 1.5 L and Accellerase® BG. The percentage of hydrolysis was calculated based on the amount of glucose released (corrected by the weight difference between glucose and its anhydrous form) divided by the initial substrate weight. The hydrolysis of each substrate was performed in triplicate.

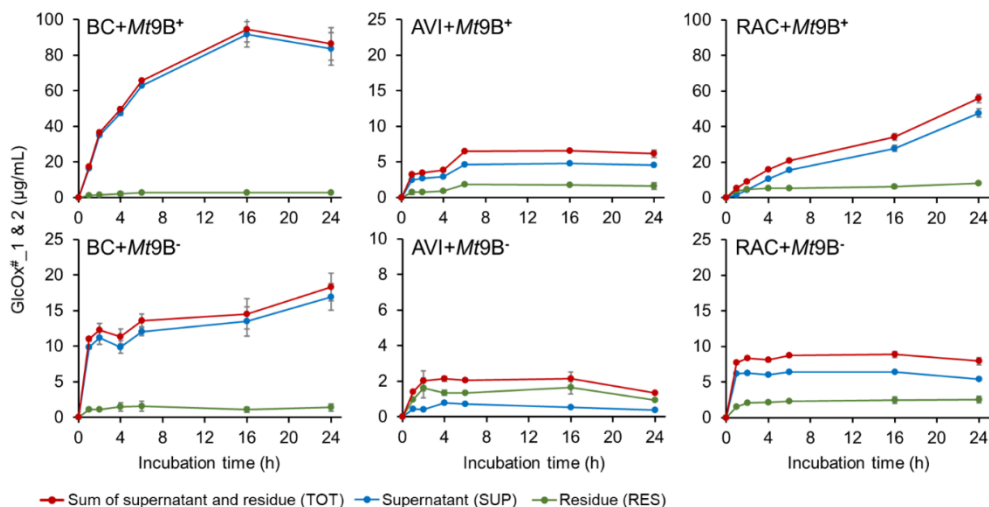


Fig. S4. Amount of gluconic acid (GlcOx#_1) and cellobionic acid (GlcOx#_2) generated from BC, AVI and RAC by $Mt9B^+$ and $Mt9B^-$ after subsequent hydrolysis, in supernatant (SUP, blue), residue (RES, green) and the sum of both (TOT, red) over time at 50 °C. Error bars (grey) indicate the standard deviations (\pm std) of duplicate measurements. Amount of GlcOx#_1 and GlcOx#_2 generated from BC, AVI and RAC by $Mt9B^+$ and $Mt9B^-$ after subsequent hydrolysis, in SUP, RES and TOT over time at 30 °C is shown in **Fig. 1**.

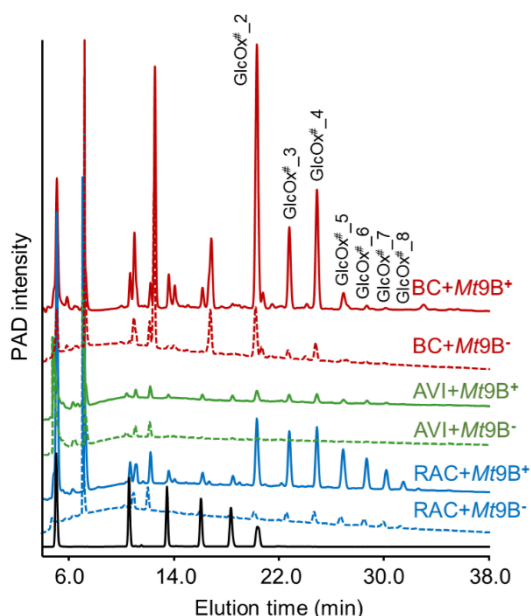


Fig. S5. HPAEC elution patterns of supernatants (SUP) from BC, AVI and RAC digests (24 h incubation) of $Mt9B^+$ and $Mt9B^-$ in the presence of Asc at 50 °C. HPAEC elution patterns of supernatants generated at 30 °C are shown in **Fig. 2**. Annotation of C1-oxidized cello-oligosaccharides (GlcOx#_2 to GlcOx#_8) is based on the previous study.(1, 5) A standard containing a mixture of glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose (from left to right in chromatogram) is shown in black. SUP of control incubations is shown in Fig. S2.

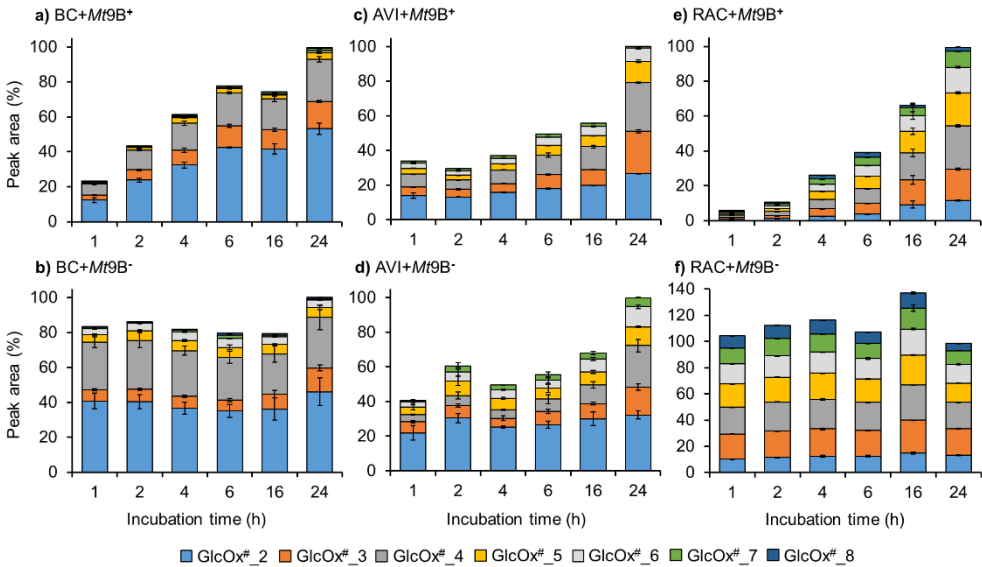


Fig. S6. Relative quantification of peak area of each DP of oxidized cello-oligosaccharides (GlcOx# 2 to GlcOx# 8) generated by Mt9B⁺ and Mt9B⁻ from the time-course incubation with BC, AVI and RAC at 50 °C. Total peak area of each 24 h sample was set to 100%. Relative quantification of the peak area of released oxidized cello-oligosaccharides at 30 °C is shown in **Fig. 3**.

Table S1. The presence of a CBM, regioselectivity and substrate specificity of four AA9 LPMOs used in this study. n.r. = not relevant.

Name	CBM	Regioselectivity	Substrate specificity	Reference
MtLPMO9B (Mt9B ⁺)	CBM1	C1	Cellulose	Frommhagen et al. (1) Grieco et al. (10)
Mt9B ⁻	No	Unknown	Unknown	n.r.
MtLPMO9I	No	C1	Cellulose	Sun et al. (5)
MtLPMO9H	CBM1	C1 and C4	Cellulose (only substrate tested)	Grieco et al. (10)
NcLPMO9M	No	C1 and C4	Cellulose, xyloglucan	Phillips et al. (11) Sun et al. (12)

Table S2. Amount of bound Mt9B⁺ and Mt9B⁻ on BC, AVI and RAC (total protein amount of 5 µM each).

Substrate	Bound (µM)			
	30 °C		50 °C	
	Mt9B ⁺	Mt9B ⁻	Mt9B ⁺	Mt9B ⁻
BC	1.08 ± 0.28	0.74 ± 0.19	1.42 ± 0.21	0.79 ± 0.07
AVI	1.34 ± 0.17	0.62 ± 0.21	1.08 ± 0.18	0.86 ± 0.22
RAC	2.85 ± 0.32	1.04 ± 0.11	2.90 ± 0.18	0.88 ± 0.21

Oxidized product profiles of AA9 LPMOs depend on the type of cellulose

Table S3. Amount of gluconic acid and cellobionic acid (GlcOx[#]_1 & 2, µg/mL) generated from BC, AVI and RAC by *Mt9B*⁺ and *Mt9B*⁻ after subsequent hydrolysis, in supernatant (SUP), residue (RES) and the sum of both (TOT) over time at 30 °C. Error indicate the standard deviations (± std) of duplicate measurements.

BC+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	9.6 ± 0.5	20.6 ± 1.5	32.4 ± 2.2	43.4 ± 3.0	70.8 ± 3.4	89.4 ± 7.9
SUP	0 ± 0	8.5 ± 0.4	19.6 ± 1.4	31.5 ± 2.2	41.8 ± 2.9	68.6 ± 3.4	87.5 ± 7.9
RES	0 ± 0	1.1 ± 0.3	1.0 ± 0.3	1.3 ± 0.3	1.6 ± 0.2	2.2 ± 0.4	2.1 ± 0.7
BC+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	12.7 ± 0.6	12.3 ± 1.3	13.7 ± 1.7	14.1 ± 1.3	16.2 ± 1.1	21.3 ± 0.0
SUP	0 ± 0	11.9 ± 0.5	12.0 ± 1.3	13.6 ± 1.7	13.5 ± 1.3	15.4 ± 1.1	20.7 ± 0.0
RES	0 ± 0	0.8 ± 0.2	0.3 ± 0.1	0.2 ± 0.1	0.7 ± 0.1	0.8 ± 0.1	0.6 ± 0.0
AVI+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	2.7 ± 0.2	4.4 ± 0.5	7.5 ± 0.3	9.7 ± 0.3	16.5 ± 0.0	19.5 ± 0.7
SUP	0 ± 0	1.7 ± 0.1	2.9 ± 0.1	5.9 ± 0.2	8.1 ± 0.2	14.7 ± 0.0	16.9 ± 0.0
RES	0 ± 0	1.0 ± 0.1	1.5 ± 0.5	1.6 ± 0.2	1.6 ± 0.3	1.8 ± 0.0	2.6 ± 0.7
AVI+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	1.6 ± 0.1	2.2 ± 0.2	2.9 ± 0.2	3.4 ± 0.2	3.4 ± 0.3	3.3 ± 0.3
SUP	0 ± 0	0.9 ± 0.1	1.5 ± 0.0	1.9 ± 0.1	2.0 ± 0.1	1.9 ± 0.1	2.1 ± 0.0
RES	0 ± 0	0.7 ± 0.0	0.7 ± 0.1	1.0 ± 0.1	1.4 ± 0.2	1.4 ± 0.2	1.2 ± 0.3
RAC+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	3.6 ± 0.3	4.1 ± 0.2	5.7 ± 0.4	6.0 ± 0.2	9.6 ± 0.5	14.3 ± 0.6
SUP	0 ± 0	1.7 ± 0.2	1.5 ± 0.2	2.3 ± 0.1	2.4 ± 0.1	5.2 ± 0.1	8.9 ± 0.5
RES	0 ± 0	1.9 ± 0.3	2.6 ± 0.1	3.4 ± 0.4	3.6 ± 0.2	4.4 ± 0.4	5.4 ± 0.3
RAC+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	5.3 ± 0.1	8.7 ± 0.4	13.9 ± 0.6	15.3 ± 0.4	13.9 ± 0.8	15.4 ± 0.4
SUP	0 ± 0	4.5 ± 0.1	7.4 ± 0.1	11.9 ± 0.1	13.2 ± 0.3	11.5 ± 0.8	13.0 ± 0.3
RES	0 ± 0	0.8 ± 0.0	1.3 ± 0.3	2.0 ± 0.6	2.2 ± 0.3	2.4 ± 0.2	2.4 ± 0.2

Chapter 2

Table S4. Amount of gluconic acid and cellobionic acid (GlcOx[#]_1 & 2, µg/mL) generated from BC, AVI and RAC by *Mt9B*⁺ and *Mt9B*⁻ after subsequent hydrolysis, in supernatant (SUP), residue (RES) and the sum of both (TOT) over time at 50 °C. Error indicate the standard deviations (± std) of duplicate measurements.

BC+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	17.4 ± 0.2	36.6 ± 0.6	49.6 ± 1.6	65.6 ± 0.8	94.6 ± 7.2	86.4 ± 9.2
SUP	0 ± 0	16.3 ± 0.2	35.0 ± 0.5	47.5 ± 1.4	63.0 ± 0.8	91.7 ± 7.2	83.5 ± 9.2
RES	0 ± 0	1.1 ± 0.0	1.6 ± 0.3	2.1 ± 0.9	2.6 ± 0.0	2.9 ± 0.1	2.8 ± 0.1
BC+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	11.0 ± 0.3	12.2 ± 0.9	11.3 ± 1.0	13.6 ± 0.9	14.5 ± 2.1	18.3 ± 1.9
SUP	0 ± 0	9.9 ± 0.2	11.1 ± 0.9	9.9 ± 0.9	12.1 ± 0.5	13.5 ± 2.1	16.9 ± 1.9
RES	0 ± 0	1.1 ± 0.2	1.1 ± 0.0	1.5 ± 0.6	1.5 ± 0.7	1.1 ± 0.4	1.4 ± 0.5
AVI+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	3.2 ± 0.3	3.4 ± 0.1	3.8 ± 0.2	6.5 ± 0.1	6.6 ± 0.1	6.1 ± 0.5
SUP	0 ± 0	2.5 ± 0.0	2.7 ± 0.0	2.9 ± 0.1	4.6 ± 0.1	4.8 ± 0.0	4.5 ± 0.0
RES	0 ± 0	0.7 ± 0.3	0.8 ± 0.1	0.9 ± 0.1	1.9 ± 0.1	1.8 ± 0.1	1.6 ± 0.5
AVI+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	1.4 ± 0.1	2.0 ± 0.6	2.1 ± 0.1	2.1 ± 0.1	2.2 ± 0.4	1.3 ± 0.1
SUP	0 ± 0	0.4 ± 0.0	0.4 ± 0.1	0.8 ± 0.0	0.7 ± 0.0	0.5 ± 0.0	0.4 ± 0.1
RES	0 ± 0	1.0 ± 0.1	1.6 ± 0.6	1.3 ± 0.1	1.3 ± 0.0	1.6 ± 0.4	0.9 ± 0.0
RAC+ <i>Mt9B</i> ⁺							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	5.5 ± 1.1	9.2 ± 0.8	15.9 ± 0.5	21.1 ± 0.7	34.2 ± 1.9	55.8 ± 2.3
SUP	0 ± 0	1.7 ± 1.1	4.4 ± 0.1	10.5 ± 0.3	15.6 ± 0.5	27.8 ± 1.7	47.7 ± 2.3
RES	0 ± 0	3.9 ± 0.4	4.8 ± 0.7	5.3 ± 0.4	5.4 ± 0.5	6.3 ± 1.0	8.1 ± 0.0
RAC+ <i>Mt9B</i> ⁻							
	0 h	1 h	2 h	4 h	6 h	16 h	24 h
TOT	0 ± 0	7.7 ± 0.2	8.3 ± 0.2	8.1 ± 0.2	8.7 ± 0.1	8.9 ± 0.5	7.9 ± 0.5
SUP	0 ± 0	6.2 ± 0.1	6.3 ± 0.2	6.0 ± 0.1	6.4 ± 0.1	6.4 ± 0.1	5.4 ± 0.2
RES	0 ± 0	1.5 ± 0.2	2.0 ± 0.1	2.1 ± 0.1	2.3 ± 0.1	2.5 ± 0.4	2.5 ± 0.5

6.3. Supplementary references

1. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
2. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
3. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al., *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
4. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnol*. 2011;7:214-23.
5. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
6. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
7. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
8. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sorlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci*. 2018;27:1636-50.
9. Courtade G, Forsberg Z, Heggset EB, Eijssink VGH, Aachmann FL. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J Biol Chem*. 2018;293:13006-15.
10. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
11. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6:1399-406.
12. Sun P, Laurent CVFP, Scheiblbrandner S, Frommhagen M, Kouzounis D, Sanders MG, et al. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnol Biofuels*. 2020;13:95.

Chapter 3

Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidized cello-oligosaccharides from their non-oxidized and reduced forms

Peicheng Sun, Matthias Frommhagen, Maloe Kleine Haar, Gijs van Erven, Edwin J. Bakx, Willem J.H. van Berkel and Mirjam A. Kabel

Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

Abstract

Lytic polysaccharide monooxygenases (LPMOs) are powerful enzymes that degrade recalcitrant polysaccharides, such as cellulose. However, the identification of LPMO-generated C1- and/or C4-oxidized oligosaccharides is far from straightforward. In particular, their fragmentation patterns have not been well established when using mass spectrometry. Hence, we studied the fragmentation behaviors of non-, C1- and C4-oxidized cello-oligosaccharides, including their sodium borodeuteride-reduced forms, by using hydrophilic interaction chromatography and negative ion mode collision induced dissociation - mass spectrometry. Non-oxidized cello-oligosaccharides showed predominantly C- and A-type cleavages. In comparison, C4-oxidized ones underwent B-/Y- and X-cleavage close to the oxidized non-reducing end, while closer to the reducing end C-/Z- and A-fragmentation predominated. C1-oxidized cello-oligosaccharides showed extensively A-cleavage. Reduced oligosaccharides showed predominant glycosidic bond cleavage, both B-/Y- and C-/Z-, close to the non-reducing end. Our findings provide signature mass spectrometric fragmentation patterns to unambiguously elucidate the catalytic behavior and classification of LPMOs.

Keywords

Lignocellulose; Biomass conversion; LPMOs; Cello-oligosaccharides; HILIC-ESI-CID-MS/MS²; Mass spectrometric fragmentation; Oxidation; Reduction

1. Introduction

An important step in the valorization of lignocellulose is the enzymatic degradation of the (hemi-)cellulosic fraction into fermentable monosaccharides (1). Hereto, hydrolytic polysaccharide degrading enzymes were considered unique for a long time. Therefore, they are abundantly present in commercial enzyme cocktails. Nevertheless, lytic polysaccharide monooxygenases (LPMOs) have been shown to boost hydrolases via oxidative cleavage of β -(1 \rightarrow 4)-linkages in recalcitrant polysaccharides, such as cellulose, in the last decade (2-5). This boosting effect has been explained by the fact that LPMOs generate new chain ends for hydrolases to act on and it has been suggested that LPMOs improve the substrate's accessibility for hydrolases as well (3, 6-8).

LPMOs have been shown to vary in their regioselectivity, which means that they are either specifically oxidizing the C1- or the C4-carbon position in polysaccharides. In addition, less regioselective LPMOs have been described able to oxidize both C1- and C4-carbon positions (9). Due to the oxidation of either the C1- or the C4-position, the corresponding glycosidic linkage becomes unstable leading to its cleavage, and, thus, eventually LPMO catalysis leads to formation of both oxidized and non-oxidized oligosaccharides (10-12). The action of C1-directed LPMOs results in release of unstable δ -lactones, which further undergo a ring cleavage in aqueous solution leading to the formation of aldonic acids. The action of C4-directed LPMOs generate 4-ketoaldoses, which might hydrate to their corresponding geminal diols in the presence of water (13, 14), but the latter is not always observed (15).

The identification of these C1- and C4-oxidized carbohydrate structures is far from straightforward, albeit various analytical approaches have been developed as thoroughly reviewed previously (16, 17). Among those analytical approaches, high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) and matrix assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS) are mostly used. Although HPAEC allows identification of LPMO-oxidized compounds based on their elution time, hereto, not commonly available standards of oxidized oligosaccharides products are required. The more complex the oxidized oligosaccharide are, which have been generated, for example LPMO-oxidized branched xyloglucan oligosaccharides (18-21), the more challenging the identification of C1- or C4- (or both) oxidation is based on their elution time. MALDI-TOF-MS can be used to detect oxidized oligosaccharides based on m/z values, but the determination of C1- and C4-oxidized oligosaccharides is far from easy. MALDI-TOF-MS cannot distinguish isomers and specifically cannot distinguish i.e., the geminal diol form of C4-oxidized from the same oligosaccharide but C1-oxidized having the same m/z (15, 22, 23). Therefore, mixed C4-, C1- or both C1-/C4-oxidized oligosaccharides cannot be distinguished by this method. To overcome these challenges, other methods, which combine chromatography directly with mass spectrometry have been studied. Thus obtained mass spectrometric fragmentation patterns are seen as distinct signatures, and differ for each oligosaccharide structure (24-28). Various attempts to separate and identify LPMO-generated oxidized products have been reported, for example by using reversed phase-ultra high performance

liquid chromatography (RP-UHPLC), porous graphitic carbon chromatography (PGC) and hydrophilic interaction chromatography (HILIC) techniques connected to electrospray ionization–collision induced dissociation–mass spectrometry (ESI-CID-MS/MS²) (13, 14, 29, 30). However, most intriguingly, a systematic study to generate MS² fragmentation patterns of the different chromatographically separated C1- and C4-oxidized oligosaccharides has not been performed yet.

In a few studies, CID-MS² of oxidized cello-oligosaccharides has been performed, albeit mainly in the positive ion mode. The positive ion mode is known to provide multiple MS² fragments with high intensity, but the presence of multiple sodium (and ammonium) adducts complicates the spectra obtained (31–34). Alternatively, CID-MS² of oligosaccharides can be performed in negative ion mode. This technique has been shown for non-oxidized oligosaccharides to give lower background noise, thus clearer spectra, and the fragmentation behaviors are suggested to be more specific and predictable (28, 35–38).

Furthermore, reduction of oligosaccharides has been shown to assist in their unambiguous separation and identification. Well known is the sodium borohydride (NaBH₄) initiated reduction of oligosaccharides to avoid α/β -anomers in various chromatographic approaches and co-occurring loss of the signal (11, 39–45). At the same time, sodium borodeuteride (NaBD₄) has been shown to label the reducing end of carbohydrates, which is helpful for the structural determination of many different oligosaccharides (46–48).

In this study, we investigated the CID-MS² fragmentation patterns of non-oxidized, LPMO-generated C1- and C4-oxidized cello-oligosaccharides in negative ion mode. Moreover, these cello-oligosaccharides were reduced by using NaBD₄ and fragmentation behaviors of the resulting oligosaccharides were studied. It is hypothesized that oxidized cello-oligosaccharides show distinct signature fragmentation patterns compared to their non-oxidized and reduced forms. Unambiguous signature fragmentation patterns of C1- and C4-oxidized cello-oligosaccharides will allow further understanding of LPMO mechanisms and their oxidized products.

2. Materials and methods

2.1. Carbohydrates, cellulose substrate and other chemicals

Galactose, NaBD₄ and ammonium acetate were purchased from Sigma-Aldrich (St. Louis, MO, USA). Cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose were purchased from Megazyme (Bray, Ireland). Regenerated amorphous cellulose (RAC) was prepared from Avicel® PH-101 (Sigma-Aldrich) as described previously (23). Ascorbic acid (Asc) and sorbitol were purchased from VWR International (Radnor, PA, USA). All water used was produced by a Milli-Q system (Millipore, Molsheim, France), unless mentioned otherwise.

2.2. Generation of non-, C4- and C1-oxidized cello-oligosaccharides by *MtLPMO9E* and *MtLPMO9I* from RAC

RAC was suspended in 50 mM ammonium acetate buffer (pH 5.0) to a concentration of 2 mg/mL. Subsequently, each LPMO from *Myceliophthora thermophila* C1 (*MtLPMO9E* (MTCTH_79765, UniProt ID: G2Q7A5) and *MtLPMO9I* (MTCTH_2299721, UniProt ID: G2Q774)) was added to the corresponding RAC suspension (in the presence of 1 mM Asc (final concentration)) at a concentration of 2 μ M. The expression, production and purification of *MtLPMO9E* and *MtLPMO9I* together with determination of their protein content and purity are described in Supplementary Information and Fig. S1. Control reactions were performed without the addition of Asc. *MtLPMO9E* incubations were carried out at 50 °C and those with *MtLPMO9I* at 30 °C by using an Eppendorf Thermomixer comfort, placed in a vertical direction, at 800 rpm (24 h reaction; 500 μ L total volume). All incubations were performed in duplicate. The reactions were stopped by immediately separating supernatants and pellets through centrifugation (22000 $\times g$, 15 min, 4 °C) in a table centrifuge. The resulting supernatants were collected and cleaned up directly or after reduction with solid phase extraction (SPE) as described in Section 2.3 prior to analysis.

2.3. Reduction of non- and oxidized-cello-oligosaccharides with NaBD₄ and clean-up with SPE

Reduction was performed by adding 200 μ L freshly prepared 0.5 M NaBD₄ to 200 μ L of the standard mixture (containing cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose, 50 μ g/mL each) and of supernatants obtained from the *MtLPMO9E*- and *MtLPMO9I*-RAC incubations at room temperature (20 °C) for 20 h. A clean-up procedure for reduced and non-reduced samples was carried out by using SPE with Supelclean™ ENVI-Carb™ columns (3 mL, Sigma-Aldrich). The SPE column was activated with 1.5 mL 80% (v/v) acetonitrile in water with 0.1% (v/v) trifluoroacetic acid (TFA), followed by adding three times 1.5 mL water. Samples were loaded on the column and washed four times with 1.5 mL water to remove unbound compounds. Bond oligosaccharides were eluted with two times 1.5 mL 60% (v/v) acetonitrile in water with 0.05% (v/v) TFA and the obtained samples were dried under a stream of nitrogen at room temperature (20 °C). The dried samples were dissolved in 50 μ L water prior to analysis.

2.4. Analytic methods

2.4.1. HPAEC-PAD analysis for profiling oligosaccharides

All samples, NaBD₄-reduced and non-reduced forms of (SPE cleaned) cello-oligosaccharides having a degree of polymerization (DP) of 2 to 6 and NaBD₄ reduced and non-reduced forms of (SPE cleaned) supernatants of RAC incubated with *MtLPMO9E* or *MtLPMO9I*, were analyzed by using HPAEC with an ICS-5000 system (Dionex, Sunnyvale, CA, USA) equipped with a CarboPac PA-1 column (2 mm ID \times 250 mm; Dionex) in combination with a CarboPac PA guard column (2 mm ID \times 50 mm; Dionex). The system was further equipped with PAD. Two mobile phases were

(A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH were kept under helium flushing and the column temperature was 20 °C. The elution profile applied has previously been described (15). Samples were diluted five times before analysis.

2.4.2. HILIC-ESI-CID-MS/MS² for elucidating fragmentation patterns

Oligosaccharides in non-reduced and NaBD₄-reduced forms of SPE-cleaned cello-oligosaccharides (DP2–6) and NaBD₄-reduced and non-reduced forms of SPE-cleaned supernatants of *MtLPMO9E*- or *MtLPMO9I*-incubated RAC samples were separated and analyzed by using HILIC-ESI-CID-MS/MS². A Vanquish UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an Acquity UPLC BEH Amide column (Waters, Millford, MA, USA; 1.7 µm, 2.1 mm ID × 150 mm) and a VanGuard pre-column (Waters; 1.7 µm, 2.1 mm ID × 150 mm) was used. The column temperature was set at 35 °C and the flow rate was 0.45 mL/min; injection volume was 2 µL. Water (A) and acetonitrile (B), both containing 0.1% (v/v) formic acid (FA) (all were UHPLC-grade; Biosolve, Valkenswaard, The Netherlands) were used as mobile phases. The separation was performed by using the following elution profile: 0–2 min at 82% B (isocratic), 2–42 min from 82% to 60% B (linear gradient), 42–42.5 min from 60% to 42% B (linear gradient), 42.5–49 min at 42% B (isocratic), 49–50 min from 42% to 82% B (linear gradient) and 50–60 min at 82% B (isocratic). The mass (*m/z*) of separated oligosaccharides was on-line detected by an LTQ Velos Pro mass spectrometer (Thermo Fisher Scientific) equipped with a heated ESI probe. MS data were obtained in negative ion mode with the following settings: source heater temperature 400°C, capillary temperature 250 °C, sheath gas flow 50 units, source voltage 2.5 kV and *m/z* range 300–1500. As MS² settings, CID with a normalized collision energy was set at 35%, with a minimum signal threshold of 5000 counts at an activation *Q* of 0.2 and activation time of 10 ms. Mass spectrometric data were processed by using Xcalibur 2.2 software (Thermo Fisher Scientific).

3. Results and discussion

3.1. Negative ion mode CID-MS² fragmentation of non-, C1- and C4-oxidized cello-oligosaccharides separated by HILIC

The aim of this research was to define specific signature fragmentation pathways for non-, C1- and C4-oxidized cello-oligosaccharides by using HILIC-ESI-CID-MS/MS² in negative ion mode. The specific signature fragmentation patterns of the NaBD₄-reduced forms of the non-, C1- and C4-oxidized oligosaccharides are discussed in Section 3.2.

3.1.1. Negative ion mode CID-MS² fragmentation of non-, C1- and C4-oxidized cello-oligosaccharides separated by HILIC

A mixture of commercially available non-oxidized cello-oligosaccharides with a DP of 2–6 (Glc₂–6) was used as the standard (Fig. S2 and S3). C4-oxidized cello-oligosaccharides (Glc_n*) were generated by incubating *MtLPMO9E* with RAC only when Asc was present (Fig. S2 and S3). We confirmed its C4-specific oxidative

cleavage with previously reported data of other C4-specific LPMOs, in particular by comparison of HPAEC elution behaviors (Fig. S2) and MALDI-TOF mass spectra (15, 29, 50). In addition to Glc_n* in the *MtLPMO9E*-RAC-digest, Glc₂-4 were also generated (Fig. S2f and Fig. S3b), which were confirmed by their elution time identical to the standards (Glc₂-4) by HPAEC (Fig. S2 and S3) and by their *m/z* values in HILIC-ESI-MS (Fig. S4). C1-oxidized cello-oligosaccharides (Glc_n[#]) were generated by incubating RAC with *MtLPMO9I* in the presence of Asc. Notably, we discovered a new C1-specific LPMO (Fig. S2 and S3) and confirmed its regioselectivity based on the comparison with previously reported data of RAC incubated with C1-specific *MtLPMO9B* and *MtLPMO9D* (15, 51). Apart from Glc_n[#], Glc₂-6 were present in the *MtLPMO9I*-RAC-digest (Fig. S2d and Fig. S3c) and their identities were similarly defined as just described for the Glc₂-4 present in the *MtLPMO9E*-RAC-digest. It should be noted that minor peaks next to each DP of Glc_n were also observed of which the identities remain so far unknown. Although these minor peaks might be of utmost importance to understand LPMO's mode-of-action in more detail, their exact nature and significance are out of the scope of this research and not further studied here.

Mass spectra of Glc_n and Glc_n* analyzed in full-scan mode with HILIC-ESI-MS indicated that these oligosaccharides were present mainly as their single charged deprotonated ([M-H]⁻) and formate adducted ([M+FA-H]⁻) products, while Glc_n[#] were present only in their [M-H]⁻ form (Fig. S4). In addition, the analyzed *m/z* of Glc_n* (*m/z* -2 compared to the *m/z* of Glc_n of the same DP) indicated that the C4-oxidized cello-oligosaccharides were in their 4-ketoaldose form. Previously reported geminal diol forms of Glc_n* (*m/z* +16 compared to the *m/z* of Glc_n of the same DP (13, 14)) were not observed under the analytical conditions used. The Glc_n[#] were observed as aldonic acids based on their *m/z* values (*m/z* +16 compared to the *m/z* of Glc_n of the same DP; Fig. S4). The separation of DP3, 4 and 5 within Glc_n, Glc_n* and Glc_n[#] (**Fig. 1**), with well distinguishable *m/z* values, allowed generation of their MS² fragmentation patterns. Hereto, the [M-H]⁻ products were preferred, because fragmentation of [M+FA-H]⁻ products was either not obtained or resulted in complex spectra with various formate adducted fragments.

Mass spectra of Glc_n and Glc_n* analyzed in full-scan mode with HILIC-ESI-MS indicated that these oligosaccharides were present mainly as their single charged deprotonated ([M-H]⁻) and formate adducted ([M+FA-H]⁻) products, while Glc_n[#] were present only in their [M-H]⁻ form (Fig. S4). In addition, the analyzed *m/z* of Glc_n* (*m/z* -2 compared to the *m/z* of Glc_n of the same DP) indicated that the C4-oxidized cello-oligosaccharides were in their 4-ketoaldose form. Previously reported geminal diol forms of Glc_n* (*m/z* +16 compared to the *m/z* of Glc_n of the same DP (13, 14)) were not observed under the analytical conditions used. The Glc_n[#] were observed as aldonic acids based on their *m/z* values (*m/z* +16 compared to the *m/z* of Glc_n of the same DP; Fig. S4). The separation of DP3, 4 and 5 within Glc_n, Glc_n* and Glc_n[#] (**Fig. 1**), with well distinguishable *m/z* values, allowed generation of their MS² fragmentation patterns. Hereto, the [M-H]⁻ products were preferred, because fragmentation of [M+FA-H]⁻ products was either not obtained or resulted in complex spectra with various formate adducted fragments.

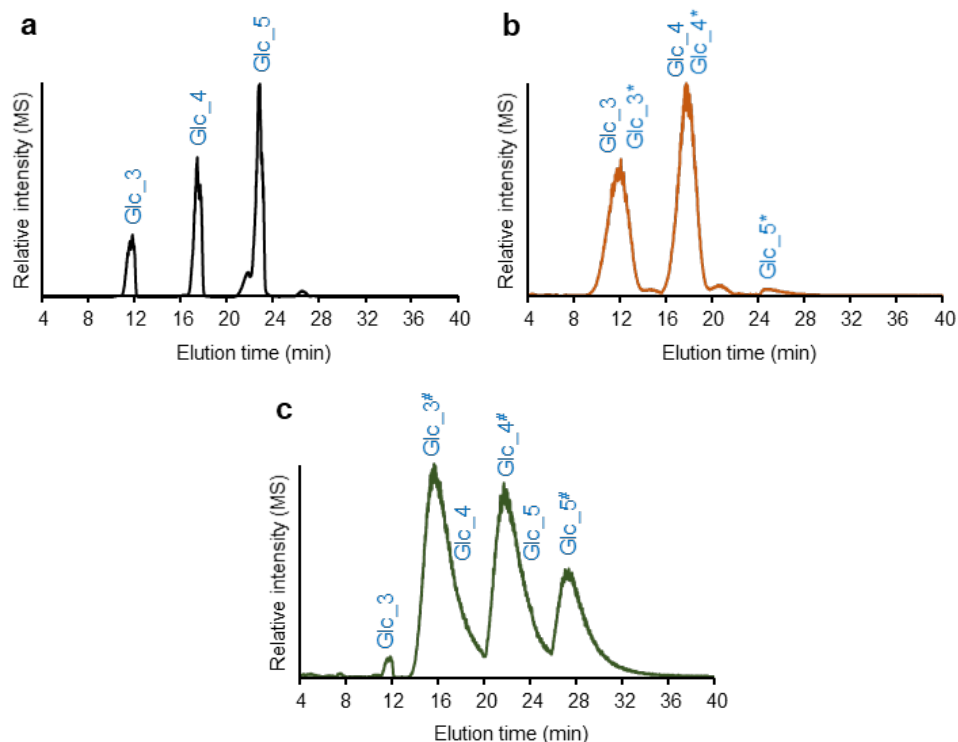


Fig. 1. HILIC extracted ion chromatograms of non-, C4- and C1-oxidized cello-oligosaccharides. (a) Non-oxidized cello-oligosaccharides: Glc_3, m/z 503.2; Glc_4, m/z 665.3; Glc_5, m/z 827.3; (b) Both non- and C4-oxidized cello-oligosaccharides: Glc_3-5 m/z see explanation panel A; Glc_3*, m/z 501.2; Glc_4*, m/z 663.3; Glc_5*, m/z 825.3; (c) Both non- and C1-oxidized cello-oligosaccharides: Glc_3-5, m/z see explanation panel A; Glc_3#, m/z 519.2; Glc_4#, m/z 681.3; Glc_5#, m/z 843.3. C4- and C1-oxidized cello-oligosaccharides were obtained from the incubation of RAC with either *MtLPMO9E* or *MtLPMO9I* in the presence of Asc (Fig. S2).

The MS² spectra of DP3, 4 and 5 of non-oxidized cello-oligosaccharides (Glc_3–5) are shown in Fig. S5 and annotated according to the nomenclature developed by Domon and Costello (52). The MS² spectrum of Glc_3 (Fig. S5) showed two most abundant fragments Z₁ (m/z 161) and C₂ (m/z 341). The C₁ fragment (m/z 179) was also found but in lower abundance. Cross-ring fragments were seen from the ions ^{0,2}A₂ (m/z 281) and ^{0,2}A₃ (m/z 443), and with consecutive water loss from the ions ^{0,2}A₂ (^{0,2}A₂–H₂O, m/z 263) and ^{0,2}A₃ (^{0,2}A₃–H₂O, m/z 425). This further loss of a water molecule of ^{0,2}A_n fragments into ^{0,2}A_n–H₂O ions (or annotated as ^{2,5}A_n ions in some studies) has previously been shown to occur (53–55). With increasing DP of non-oxidized cello-oligosaccharides (Glc_4 and Glc_5), C-type fragmentation was still predominant as shown by the pronounced relative intensity of C₂ (m/z 341), C₃ (m/z 503) and C₄ (m/z 665). In addition, the most abundant cross-ring fragments ^{0,2}A_n and ^{0,2}A_n–H₂O resulted from cleavages of the B ring (reducing end) and the B-neighboring ring. Similar CID-MS² results of neutral β-(1→4) linked gluco-oligosaccharides ([M–H][–]) with predominant C-type glycosidic cleavage and ^{0,2}A_n cross-ring fragmentation have been described (36, 56–58). Although the MS² fragmentation patterns of Glc_n in negative

ion mode can be found in literature, we still provided our MS² spectra of Glc_n here as the reference to better compare to the spectra of other types of cello-oligosaccharides.

The MS² fragmentation spectra of DP3 and DP4 of C4-oxidized cello-oligosaccharides (Glc_3* and Glc_4*) are shown in **Fig. 2**. The MS intensity of Glc_5* was too weak to obtain decent MS² spectra, mainly due to the very low amount of Glc_5* present in the *MtLPMO9E*-RAC-digest. Various attempts, for example increasing the dose of *MtLPMO9E*, were performed to increase the yield of Glc_5*. However, the yield of Glc_5* was not significantly improved (data not shown). It is thought that *MtLPMO9E* is able to cleave soluble cello-oligosaccharides having DP larger than 5, which has been reported previously (50). Nevertheless, to our opinion, the clear signature fragmentation behaviors and diagnostic fragments of DP3 and DP4 are representative for C4-oxidized cello-oligosaccharides in general. The MS² spectrum of Glc_3* indicated B₁ (*m/z* 159) and C₁ (*m/z* 177) ions, which were *m/z* -2 compared to Z₁ (*m/z* 161) and Y₁ (*m/z* 179), indicating, as expected, that an oxidized glucosyl unit was present at the non-reducing end (A ring). MS² ring-fragments of Glc_3*, ^{0,2}A₂ (*m/z* 279) and ^{0,2}A₃ (*m/z* 441; ^{0,2}A₃-H₂O, *m/z* 423) were also found, in decreasing abundance towards the A ring. Surprisingly, the fragment ^{2,4}X₂ (*m/z* 281) was detected, which has barely been shown to occur in negative ion mode CID-MS² of oligosaccharides so far. Fragmentation of the C4-oxidized cellotriose apparently resulted in loss of the oxidized end via ring cleavage leading to a ^{2,4}X₂ (*m/z* 281) fragment. In the MS² spectrum of Glc_4*, the three most abundant fragments were B₁ (*m/z* 159), Y₂ (*m/z* 341) and Y₃ (*m/z* 503). Hence, B- and Y-type fragmentation preferably took place at the glycosidic linkage at the right side of the A and A neighboring ring, most likely influenced by the C4-oxidized group. In contrast, pronounced ions of Z₁ (*m/z* 161), C₃ (*m/z* 501) and C₂ (*m/z* 339) suggested C- and Z-type fragmentation at the glycosidic linkage at the left side of the B ring, similar to the fragmentation of non-oxidized cello-oligosaccharides. This preference was further apparent from the very low intensity of C₁ (*m/z* 177), Y₁ (*m/z* 179), B₃ (*m/z* 483) and Z₃ (*m/z* 485) fragments. Likewise, for Glc_3*, A- and X-type of cross-ring fragmentation of Glc_4* was observed as ^{2,4}X₃ (*m/z* 443) and ^{2,4}X₃-H₂O (*m/z* 425). Hence, these ^{2,4}X_n fragments can be seen as diagnostic ions specifically for C4-oxidized cello-oligosaccharides. The intensity of the ^{2,4}X_n fragments decreased closer to the B ring, observed from the much lower relative intensity of ^{2,4}X₂ (*m/z* 281) and ^{2,4}X₂-H₂O (*m/z* 263). For the B ring and B neighboring rings ^{0,2}A_n (-H₂O) fragments were predominant and decreased in intensity when situated closer to the A ring.

A summary of the MS² fragmentation patterns of the (C4-oxidized) cello-oligosaccharides, including signature fragments, is given in **Table 1**.

The MS² fragmentation of C1-oxidized cello-oligosaccharides (aldonic acids, Glc_n[#], **Fig. 3** and Fig. S6) was very distinct from the fragmentation obtained for Glc_n and Glc_n*. In the Glc_3[#] MS² spectrum (**Fig. 3**), abundant B₁ (*m/z* 161) and C₁ (*m/z* 179) ions were observed, which represented the cleavage of the non-reducing end glucosyl unit (A ring). Apparently, glycosidic linkage cleavage next to, or in the neighborhood of the aldonic acid residue was less favored. In addition, A-type cross-ring fragments

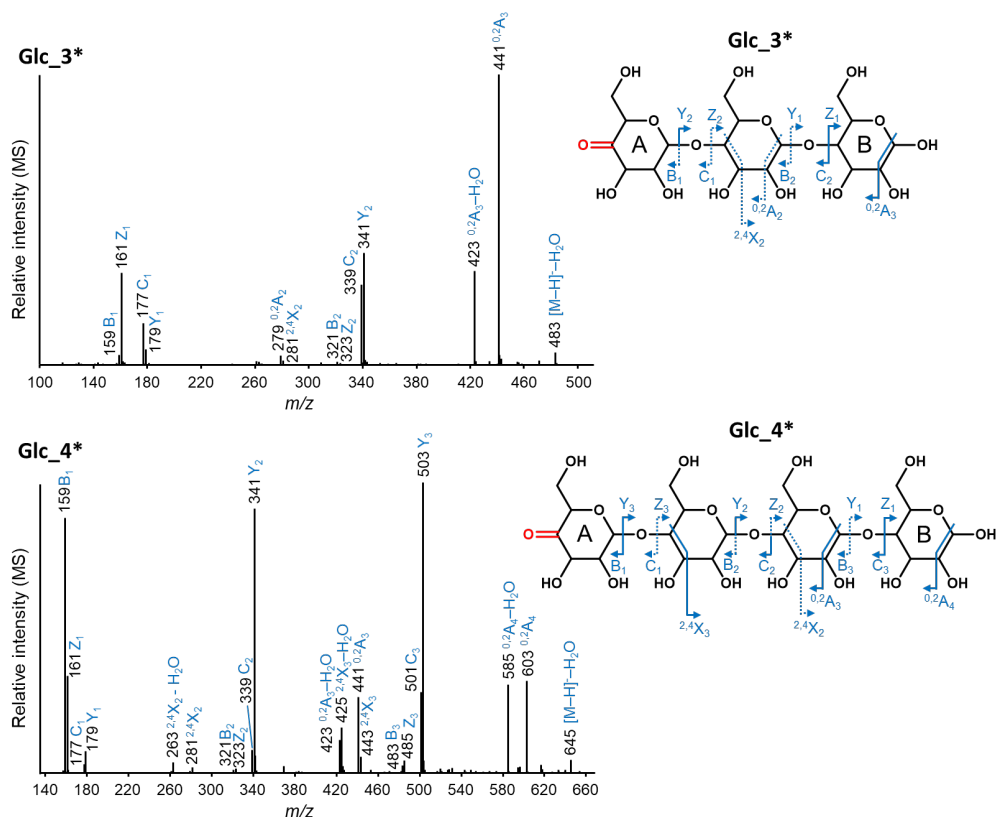


Fig. 2. Negative ion mode CID-MS² spectra of C4-oxidized cello-oligosaccharides DP3 and DP4 (Glc_3*, *m/z* 501.2 and Glc_4*, *m/z* 663.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (52). Blue solid arrows indicate abundant fragments, while dashed arrows indicate fragments with a lower relative intensity. C4-ketone groups are indicated in red. A and B ring represent the non-reducing and reducing end, respectively. C4-oxidized cello-oligosaccharides were generated by RAC incubated with *MtLPMO9E* in the presence of Asc (Fig. S2).

were detected, especially from cross-ring cleavage of the carboxyl end (B ring; ^{2,4}A₃ (*m/z* 383)). Likewise for Glc_3*, MS² of Glc_4* (**Fig. 3**) showed predominant cross-ring cleavage on the oxidized B ring and B neighboring ring (^{2,4}A₄ (*m/z* 545) and ^{2,4}A₃ (*m/z* 383), respectively; **Table 1**). Again ^{0,2}A_n (-H₂O) fragments were formed. Similarly, the MS² pattern of Glc_5* (Fig. S6) showed a series of ^{2,4}A_n and ^{0,2}A_n (-H₂O) fragments, which were predominant for the B ring and the B neighboring ring (e.g. ^{2,4}A₄, *m/z* 545 and ^{2,4}A₅, *m/z* 707). Glycosidic bond cleavage still occurred but was relatively less abundant (**Table 1**). Although similar data for the MS² spectrum of Glc_3* has been reported previously by Boulos and Nystrom (53), the systematic fragmentation behaviors of a series of C1-oxidized cello-oligosaccharides is here presented for the first time.

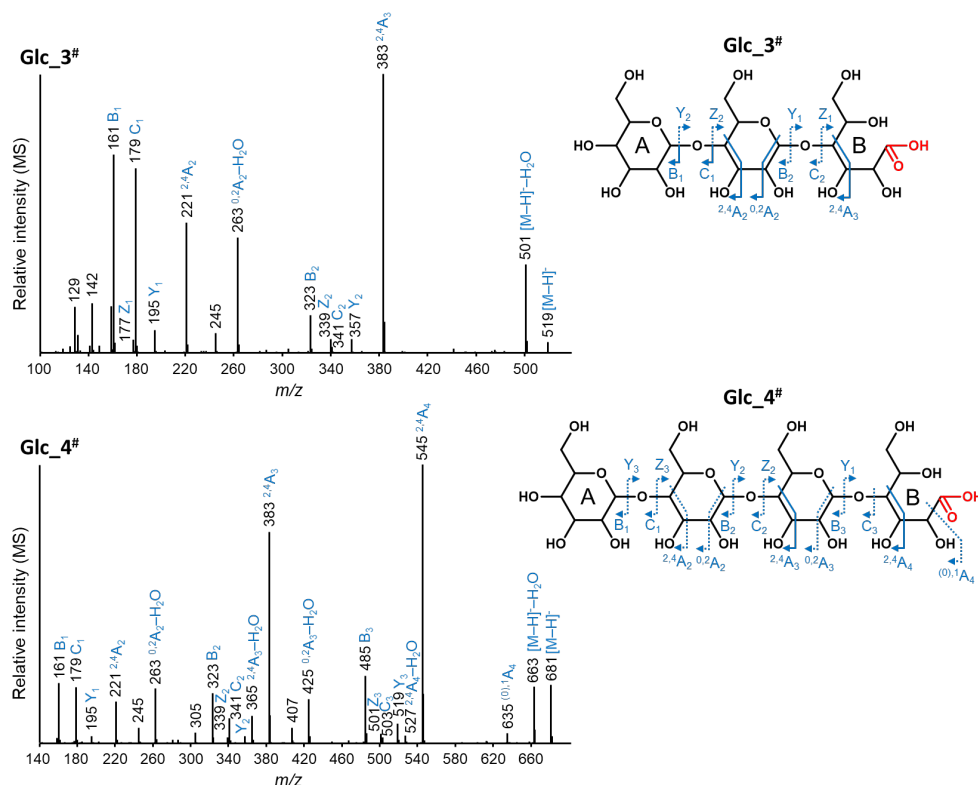


Fig. 3. Negative ion mode CID-MS² spectra of C1-oxidized cello-oligosaccharides DP3 (Glc_3#, *m/z* 519.2) and DP4 (Glc_4#, *m/z* 681.3). See Fig. S6 for DP5 (Glc_5#, *m/z* 843.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (52). Blue solid arrows indicate most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. Carboxyl groups are indicated in red. A and B ring represent the non-reducing and gluconic acid end, respectively. C1-oxidized cello-oligosaccharides were generated by RAC incubated with *MtLPMO9I* in the presence of Asc (Fig. S2).

3.2. Negative ion mode CID-MS² fragmentation of reduced non- and C4-oxidized cello-oligosaccharides separated by HILIC

Non- and C4-oxidized cello-oligosaccharides were successfully reduced by NaBD₄ to their alditol forms (RD-Glc_n and RD-Glc_n*) and analyzed by HPAEC (Fig. S7), while the obtained gluconic acid forms of the C1-oxidized cello-oligosaccharides cannot be reduced and remained in their aldonic acid form (Fig. S7). Through NaBD₄ reduction, the B ring of Glc_n was converted to the alditol form with one deuterium ion inserted, which led to *m/z* of 3 higher compared to the non-reduced Glc_n with the same DP (Fig. S8). Similarly to Glc_n, Glc_n* were reduced at their reducing end, however, the C4-oxidized A ring was also reduced with addition of one deuterium ion, which was evident from the *m/z* observed. RD-Glc_n* resulted in a *m/z* of 4 higher compared to the non-reduced Glc_n with the same DP (Fig. S8). Constituent monosaccharide analysis after TFA hydrolysis showed that reduction of the A ring led the formation of

glucosyl as well as galactosyl units (Fig. S9). HPAEC monosaccharide chromatograms clearly showed that galactose was generated in RD-Glc_n*, but not in RD-Glc_n after TFA hydrolysis. The generation of galactose after reduction of C4-oxidized cello-oligosaccharides and TFA hydrolysis was also described by Beeson et al. (11). It can be hypothesized that the reduction of the C4-oxidized glucosyl residue leads to the formation of a hydroxyl group at the C4-carbon atom, whose conformational orientation is in equilibrium between the equatorial or axial position. The formation of the two types of RD-Glc_n* was also seen from the HPAEC chromatogram (Fig. S7).

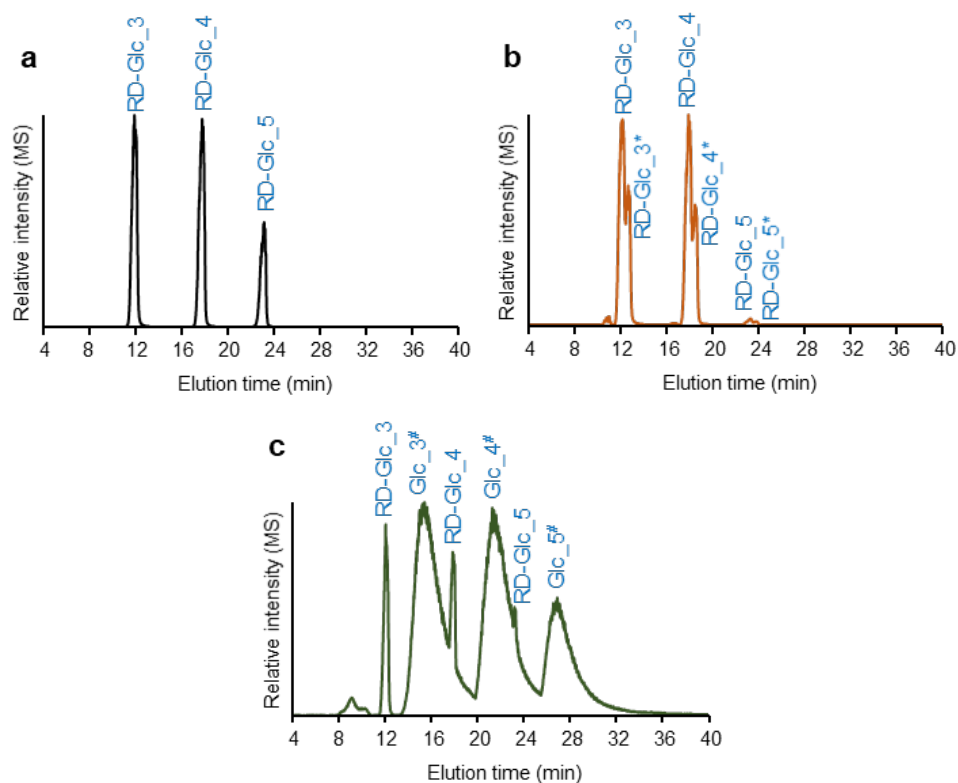


Fig. 4. HILIC extracted ion chromatograms of NaBD4-reduced non-, C4- and C1-oxidized cello-oligosaccharides. (a) Reduced non-oxidized cello-oligosaccharides: RD-Glc_3, m/z 506.3; RD-Glc_4, m/z 668.3; RD-Glc_5, m/z 830.3; (b) Both reduced non- and C4-oxidized cello-oligosaccharides: RD-Glc_3–5 m/z see explanation panel A; RD-Glc_3*, m/z 507.2; RD-Glc_4*, m/z 669.3; RD-Glc_5*, m/z 831.3; (c) Reduced non- and C1-oxidized cello-oligosaccharides: RD-Glc_3–5, m/z see explanation panel A; Glc_3#, m/z 519.2; Glc_4#, m/z 681.3; Glc_5#, m/z 843.3.

As RD-Glc_n* with glucosyl or galactosyl units have the same m/z values in HILIC-ESI-MS, we used RD-Glc_n* to indicate both (Fig. S8). RD-Glc_3–5 and RD-Glc_3–5* ($[M-H]^-$) were well separated by HILIC (Fig. 4), allowing accurate characterization of their MS² fragmentation behaviors. An overview of fragmentation behaviors of the reduced cello-oligosaccharides is shown in **Table 1**.

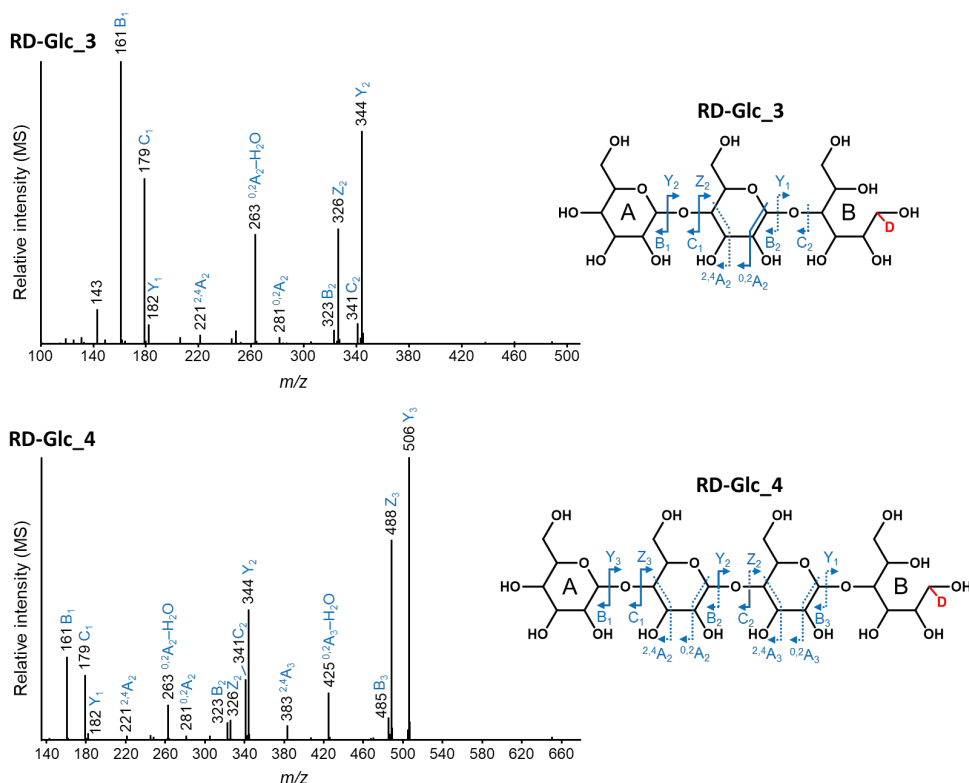


Fig. 5. Negative ion mode CID-MS² spectra of reduced non-oxidized cello-oligosaccharides DP3 (RD-Glc_3, *m/z* 506.3) and DP4 (RD-Glc_4, *m/z* 668.3). See Fig. S10 for DP5 (RD-Glc_5, *m/z* 830.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (52). Blue solid arrows indicate the most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. The deuterium ion is indicated in red. A and B ring represent the non-reducing and deuterated alditol end, respectively.

The MS² fragments of RD-Glc_n (**Fig. 5** and Fig. S10) showed again distinct cleavage patterns. In the MS² spectrum of RD-Glc_3 (**Fig. 5**), B₁ (*m/z* 161) and C₁ (*m/z* 179) were abundant ions and mainly generated from the A ring. Closer to the B ring, glycosidic bond fragmentation was less pronounced indicated by the low relative intensity of B₂ (*m/z* 323) and C₂ (*m/z* 341). The same fragmentation was observed from the high relative intensity of Y₂ (*m/z* 344) and Z₂ (*m/z* 326) and low abundance of Y₁ (*m/z* 182). Cross-ring fragmentation occurred, mainly seen from ^{2,4}A₂ (*m/z* 221) and more abundant ^{0,2}A₂-H₂O (*m/z* 263). RD-Glc_4 exhibited a similar trend of fragmentation as determined for RD-Glc_3 (**Fig. 5**). Again, glycosidic bond cleavage preferably occurred at the right side of the A ring and decreased closer to the B ring: Z₃ (*m/z* 488) and Y₃ (*m/z* 506) were the most abundant ions followed by Y₂ (*m/z* 344), B₁ (*m/z* 161) and C₁ (*m/z* 179). Cross-ring fragmentation (^{0,2}A_n and ^{2,4}A_n) was indicated, like for RD-Glc_3, but to a lesser extent. The MS² fragmentation patterns as described above were also obtained for RD-Glc_5 (Fig. S10), demonstrated by the abundant ions of Y₄ (*m/z* 668), Z₄ (*m/z* 650) and Y₃ (*m/z* 506) and relative low

intensities of B₄ (*m/z* 647), B₃ (*m/z* 485), C₃ (*m/z* 503), B₂ (*m/z* 323) and Z₂ (*m/z* 326).

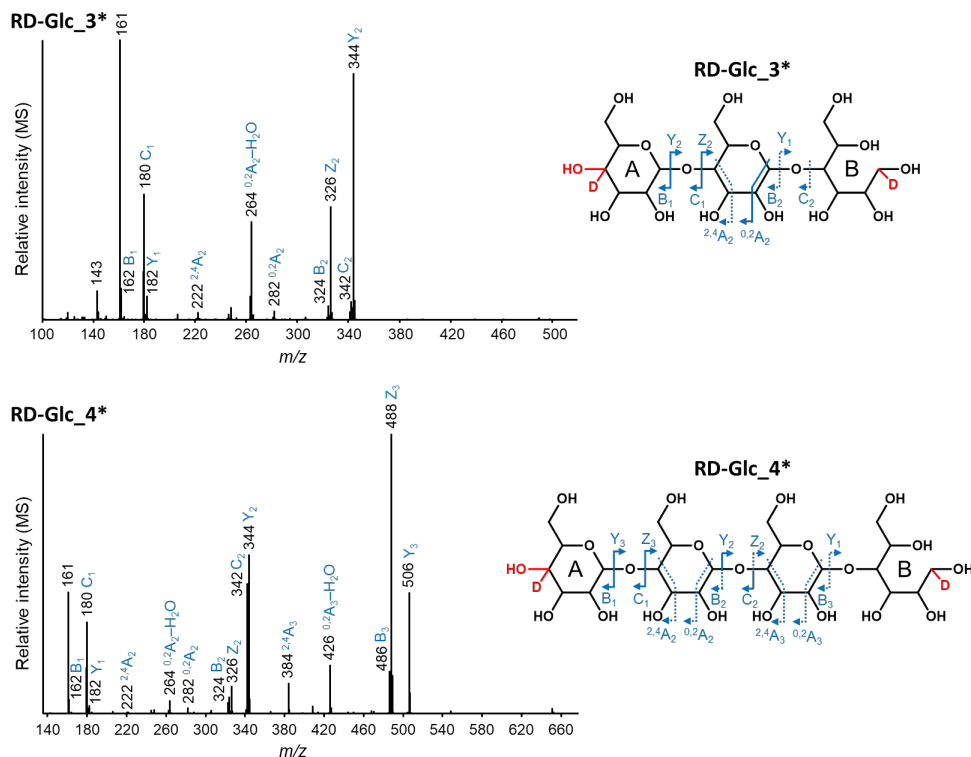


Fig. 6. Negative ion mode CID-MS² spectra of reduced C4-oxidized cello-oligosaccharides DP3 (RD-Glc_3*, *m/z* 507.2) and DP4 (RD-Glc_4*, *m/z* 669.3). See Fig. S11 for DP5 (RD-Glc_5*, *m/z* 831.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (52). Blue solid arrows indicate the most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. The formation of hydroxyl groups from the C4-ketone groups and deuterium ions are indicated in red. A and B ring represent the non-reducing and deuterated alditol end, respectively. Note that the A-ring can represent a glucosyl and galactosyl unit.

The MS² spectra of RD-Glc_n* (**Fig. 6** and Fig. S11) displayed overall the same patterns and fragmentation behavior as obtained for RD-Glc_n, which can be expected given their almost identical structures after reduction. The only difference is the deuterium ion at the C4-position of the glucosyl A ring of RD-Glc_n* compared to the hydrogen ion for the RD-Glc_n oligosaccharides, resulting in A ring fragments of *m/z* +1. This mass difference in the ions containing the A ring still allowed to distinguish RD-Glc_n from RD-Glc_n*. As indicated above, reduction also formed A ring galactosyl units, which could have potentially influenced the MS² fragmentation patterns of RD-Glc_n*. Yet, due to co-elution these effects cannot be further specified and the absence of cross-A-ring fragments furthermore does not allow to distinguish both structures.

Ions of *m/z* 162, 180, 324 and 342 in RD-Glc_3* (**Fig. 6**) were annotated as B₁, C₁, B₂ and C₂, respectively, as they obtained *m/z* +1 compared to the ions generated from

RD-Glc_3. Similarly, fragments of m/z 222, 264 and 282 were determined to be $^{2,4}A_2$, $^{0,2}A_2-H_2O$ and $^{0,2}A_2$, respectively. The similar MS² fragments having m/z +1 were also found in RD-Glc_4* (**Fig. 6**) and RD-Glc_5* (Fig. S11) compared to the corresponding annotate ions in RD-Glc_4 and RD-Glc_5, respectively. In the MS² spectra of RD-Glc_3* and RD-Glc_4*, m/z 161 was unexpectedly more abundant than the B₁ ions (m/z 162). It is hypothesized that m/z 161 came from the internal glucosyl ions resulting from a double cleavage on RD-Glc_n*. Such double cleavage generating internal glucosyl ions (m/z 161) can also occur for RD-Glc_n, but cannot be distinguished from its B₁ ion (m/z 161). This double cleavage has been previously shown on other types of gluco-oligosaccharides in negative ion mode, such as carbohydrate chains of glycoproteins having 3-linked N-acetylglucosamine units and endoglucanase-digested xyloglucan oligosaccharides (28, 55).

In summary, the five types of cello-oligosaccharides analyzed showed distinct fragmentation patterns in HILIC-separated negative ion mode CID-MS² (**Table 1**). In general, both glycosidic bond cleavage and cross-ring fragmentation occurred, though they were found to be rather different for each type of cello-oligosaccharide. Predominant C- and $^{0,2}A_n (-H_2O)$ -fragments were found for non-oxidized cello-oligosaccharides, as also widely described in literature. C4-oxidized cello-oligosaccharides obtained predominant B-/Y-type fragmentation on the oxidized A ring while C-/Z-fragments were pronounced on the reducing end B ring. Notably, diagnostic $^{2,4}X_n (-H_2O)$ -ions were annotated for the first time, which distinguish C4-oxidized cello-oligosaccharides from other ones. Extensive A-type cross-ring fragmentation over glycosidic bond cleavage was found in C1-oxidized cello-oligosaccharides, which is a unique feature of these particular cello-oligosaccharides. Though reduced non- and C4-oxidized cello-oligosaccharides showed identical fragmentation patterns, their structures can be distinguished by the different masses of ions containing the A ring. Compared to their non-reduced forms, reduced non- and C4-oxidized cello-oligosaccharides showed preferred glycosidic bond cleavage closer to the A ring.

Table 1. Summary of fragmentation patterns of five types of cello-oligosaccharides observed in HILIC-ESI-CID-MS².

Sugar type	Structure (β -(1 \rightarrow 4)-glucosyl backbone)	Fragmentation patterns	
		Glycosidic bond cleavage	Cross-ring cleavage
Non-oxidized (Glc _n)		C-ions are predominant	^{0,2} A _n ⁻ and ^{2,4} A _n ⁻ -fragments; ^{0,2} A _n ⁻ -fragments are predominant on B ring, but decreases closer to A ring; ^{2,4} A _n ⁻ -fragments on B ring and B neighboring ring
C4-oxidized (Glc _n [*])		Both B-/Y- and C-/Z-fragments; B-/Y-cleavage is predominant directly next to A ring, but decreases closer to B ring; C-/Z-cleavage is predominant directly next to B ring but decreases closer to A ring	^{2,4} X _n ⁻ and ^{0,2} A _n ⁻ -fragments; ^{2,4} X _n ⁻ -fragments are predominant on A ring, but decreases closer to B ring; ^{0,2} A _n ⁻ -fragments are predominant on B ring, but decreases closer to A ring
C1-oxidized (Glc _n [#])		Both B-/Y- and C-/Z-fragments; glycosidic bond cleavage is less pronounced compared to ring cleavage	^{0,2} A _n ⁻ , ^{2,4} A _n ⁻ and little ^{(0),1} A _n ⁻ fragments; ^{2,4} A _n ⁻ -fragments are predominant (even over glycosidic bond cleavage) on the ring closed to B ring, but decreases closer to A ring
Reduced- non-oxidized (RD-Glc _n)		Both B-/Y- and C-/Z-fragments; glycosidic bond cleavage is predominant closer to A ring, but decreases closer to B ring	^{0,2} A _n ⁻ and ^{2,4} A _n ⁻ -fragments; not on A and B rings, but on the neighboring rings of A and B
Reduced-C4- oxidized (RD-Glc _n [*])		Both B-/Y- and C-/Z-fragments; glycosidic bond cleavage is predominant closer to A ring, but decreases closer to B ring	^{0,2} A _n ⁻ and ^{2,4} A _n ⁻ -fragments; not on A and B rings, but on the neighboring rings of A and B

4. Conclusions

In this study, fragmentation patterns of non-, C1- and C4-oxidized cello-oligosaccharides released by LPMOs and the reduced forms of non- and C4-oxidized cello-oligosaccharides were analyzed by using HILIC-ESI-CID-MS/MS² in negative ion mode. All type of cello-oligosaccharides showed both glycosidic bond and cross-ring cleavage fragments, but the fragmentation pattern of each type is distinct. Based on the outcomes of this research, also other structures of different oligosaccharides containing a β -(1 \rightarrow 4)-glucosyl (cellulose-like) backbone, for example branched oxidized xyloglucan oligosaccharides, can be elucidated by their diagnostic ions and the specific fragmentation patterns. The structural elucidation of these complex (oxidized) oligosaccharides will further help understanding the mode-of-action of LPMOs regarding their ability to oxidatively degrade a range of plant cell wall polysaccharides, including cellulose and xyloglucan. Moreover, it can be envisaged that

beyond the LPMO-field, our study contributes to the characterization of (chemically) oxidized oligosaccharides in general.

5. References

1. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, et al. The path forward for biofuels and biomaterials. *Science*. 2006;311:484-9.
2. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33:747-61.
3. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.
4. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunes AC, Stenstrom Y, MacKenzie A, et al. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20:1479-83.
5. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørle M, et al. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219-22.
6. Martinez AT. How to break down crystalline cellulose. *Science*. 2016;352:1050-1.
7. Villares A, Moreau C, Bennati-Granier C, Garajova S, Foucat L, Falourd X, et al. Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. *Sci Rep*. 2017;7:40262.
8. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*. 2010;49:3305-16.
9. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol*. 2018;9:1080.
10. Kim S, Stahlberg J, Sandgren M, Paton RS, Beckham GT. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxygen rebound mechanism. *Proc Natl Acad Sci U S A*. 2014;111:149-54.
11. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134:890-2.
12. Beeson WT, Vu VV, Span EA, Phillips CM, Marletta MA. Cellulose degradation by polysaccharide monooxygenases. *Annu Rev Biochem*. 2015;84:923-46.
13. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, et al. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289:2632-42.
14. Westereng B, Arntzen MO, Aachmann FL, Varnai A, Eijsink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monooxygenases acting on cellulose. *J Chromatogr A*. 2016;1445:46-54.
15. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
16. Eijsink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnol Biofuels*. 2019;12:58.
17. Monclaro AV, Filho EXF. Fungal lytic polysaccharide monooxygenases from family AA9: Recent developments and application in lignocellulose breakdown. *Int J Biol Macromol*. 2017;102:771-8.
18. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Varnai A, et al. FgLPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Lett*. 2016;590:3346-56.
19. Kojima Y, Varnai A, Ishida T, Sunagawa N, Petrovic DM, Igarashi K, et al. A lytic polysaccharide monooxygenase with broad xyloglucan specificity from the brown-rot fungus *Gloeophyllum trabeum* and its action on cellulose-xyloglucan complexes. *Appl Environ Microbiol*. 2016;82:6557-72.
20. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sorlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci*. 2018;27:1636-50.

21. Ladeveze S, Haon M, Villares A, Cathala B, Grisel S, Herpoel-Gimbert I, et al. The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2017;10:215.
22. Forsberg Z, Mackenzie AK, Sorlie M, Rohr AK, Helland R, Arvai AS, et al. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2014;111:8446-51.
23. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
24. Westphal Y, Kuhnelt S, Schols HA, Voragen AG, Gruppen H. LC/CE-MS tools for the analysis of complex arabino-oligosaccharides. *Carbohydr Res*. 2010;345:2239-51.
25. Westphal Y, Schols HA, Voragen AG, Gruppen H. Introducing porous graphitized carbon liquid chromatography with evaporative light scattering and mass spectrometry detection into cell wall oligosaccharide analysis. *J Chromatogr A*. 2010;1217:689-95.
26. Kool MM, Gruppen H, Sworn G, Schols HA. Comparison of xanthans by the relative abundance of its six constituent repeating units. *Carbohydr Polym*. 2013;98:914-21.
27. Pfenninger A, Karas M, Finke B, Stahl B. Structural analysis of underivatized neutral human milk oligosaccharides in the negative ion mode by nano-electrospray MS⁽ⁿ⁾ (part 2: application to isomeric mixtures). *J Am Soc Mass Spectrom*. 2002;13:1341-8.
28. Chai W, Piskarev V, Lawson AM. Negative-ion electrospray mass spectrometry of neutral underivatized oligosaccharides. *Anal Chem*. 2001;73:651-7.
29. Frommhagen M, van Erven G, Sanders M, van Berkel WJH, Kabel MA, Gruppen H. RP-UHPLC-UV-ESI-MS/MS analysis of LPMO generated C4-oxidized gluco-oligosaccharides after non-reductive labeling with 2-aminobenzamide. *Carbohydr Res*. 2017;448:191-9.
30. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels*. 2015;8:90.
31. Harvey DJ. N-(2-diethylamino)ethyl-4-aminobenzamide derivative for high sensitivity mass spectrometric detection and structure determination of N-linked carbohydrates. *Rapid Commun Mass Spectrom*. 2000;14:862-71.
32. Deery MJ, Stimson E, Chappell CG. Size exclusion chromatography/mass spectrometry applied to the analysis of polysaccharides. *Rapid Commun Mass Spectrom*. 2001;15:2273-83.
33. Reis A, Domingues MRM, Domingues P, Ferrer-Correia AJ, Coimbra MA. Positive and negative electrospray ionisation tandem mass spectrometry as a tool for structural characterisation of acid released oligosaccharides from olive pulp glucuronoxylans. *Carbohydr Res*. 2003;338:1497-505.
34. Kailemia MJ, Ruhaak LR, Lebrilla CB, Amster IJ. Oligosaccharide analysis by mass spectrometry: A review of recent developments. *Anal Chem*. 2014;86:196-212.
35. Harvey DJ. Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans. *J Am Soc Mass Spectrom*. 2005;16:631-46.
36. Chai W, Luo J, Lim CK, Lawson AM. Characterization of heparin oligosaccharide mixtures as ammonium salts using electrospray mass spectrometry. *Anal Chem*. 1998;70:2060-6.
37. Lawson AM, Chai W, Cashmore GC, Stoll MS, Hounsell EF, Feizi T. High-sensitivity structural analyses of oligosaccharide probes (neoglycolipids) by liquid-secondary-ion mass spectrometry. *Carbohydr Res*. 1990;200:47-57.
38. Gao C, Zhang Y, Liu Y, Feizi T, Chai W. Negative-ion electrospray tandem mass spectrometry and microarray analyses of developmentally regulated antigens based on Type 1 and Type 2 backbone sequences. *Anal Chem*. 2015;87:11871-8.
39. York WS, Kolli VSK, Orlando R, Albersheim P, Darvill AG. The structures of arabinoxyloglucans produced by solanaceous plants. *Carbohydr Res*. 1996;285:99-128.
40. Abdel-Akher M, Hamilton JK, Smith F. The reduction of sugars with sodium borohydride. *J Am Chem Soc*. 1951;73:4691-2.
41. Barr BK, Hsieh YL, Ganem B, Wilson DB. Identification of two functionally different classes of exocellulases. *Biochemistry*. 1996;35:586-92.
42. Hantus S, Pauly M, Darvill AG, Albersheim P, York WS. Structural characterization of novel L-galactose-containing oligosaccharide subunits of jojoba seed xyloglucans. *Carbohydr Res*. 1997;304:11-20.
43. Kawasaki N, Ohta M, Hyuga S, Hyuga M, Hayakawa T. Application of liquid chromatography/mass spectrometry and liquid chromatography with tandem mass

- spectrometry to the analysis of the site-specific carbohydrate heterogeneity in erythropoietin. *Anal Biochem.* 2000;285:82-91.
44. Vinogradov E, Perry MB, Conlan JW. Structural analysis of *Francisella tularensis* lipopolysaccharide. *Eur J Biochem.* 2002;269:6112-8.
45. Mazumder K, York WS. Structural analysis of arabinoxylans isolated from ball-milled switchgrass biomass. *Carbohydr Res.* 2010;345:2183-93.
46. Xie Y, Liu J, Zhang J, Hedrick JL, Lebrilla CB. Method for the comparative glycomic analyses of O-linked, mucin-type oligosaccharides. *Anal Chem.* 2004;76:5186-97.
47. Ring SG, Selvendran RR. Purification and methylation analysis of cell wall material from *Solanum tuberosum*. *Phytochemistry.* 1978;17:745-52.
48. Selvendran RR, King SE. Structural features of the cell-wall polysaccharides of the parchment layers of the pods of mature runner beans. *Carbohydr Res.* 1989;195:87-99.
49. Irwin DC, Spezio M, Walker LP, Wilson DB. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. *Biotechnol Bioeng.* 1993;42:1002-13.
50. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. *PLOS ONE.* 2018;13:e0202148.
51. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol.* 2017;102:1281-95.
52. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 1988;5:397-409.
53. Boulos S, Nystrom L. UPLC-MS/MS investigation of beta-glucan oligosaccharide oxidation. *Analyst.* 2016;141:6533-48.
54. Mulrone B, Barrie Peel J, Traeger JC. Theoretical study of deprotonated glucopyranosyl disaccharide fragmentation. *J Mass Spectrom.* 1999;34:856-71.
55. Quémener B, Vigouroux J, Rathahao E, Tabet JC, Dimitrijevic A, Lahaye M. Negative electrospray ionization mass spectrometry: a method for sequencing and determining linkage position in oligosaccharides from branched hemicelluloses. *J Mass Spectrom.* 2015;50:247-64.
56. Palma AS, Liu Y, Zhang HT, Zhang YB, McCleary BV, Yu GL, et al. Unravelling glucan recognition systems by glycome microarrays using the designer approach and mass spectrometry. *Mol Cel Proteomics.* 2015;14:974-88.
57. Quémener B, Désiré C, Debrauwer L, Rathahao E. Structural characterization by both positive and negative electrospray ion trap mass spectrometry of oligogalacturonates purified by high-performance anion-exchange chromatography. *J Chromatogr.* 2003;984:185-94.
58. Pfenninger A, Karas M, Finke B, Stahl B. Structural analysis of underivatized neutral human milk oligosaccharides in the negative ion mode by nano-electrospray MS⁽ⁿ⁾ (part 1: methodology). *J Am Soc Mass Spectrom.* 2002;13:1331-40.
59. Van Soest Pv, Robertson J, Lewis B. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J Dairy Sci.* 1991;74:3583-97.

6. Supplementary information

6.1. Supplementary materials and methods

6.1.1. Enzyme expression, production and purification

The genes encoding *MtLPMO9I* (MYCTH_2299721, UniProt ID: G2Q774) and *MtLPMO9E* (MYCTH_79765, UniProt ID: G2Q7A5, previously reported as *MtLPMO9J* (1)), were homologously expressed in a low protease/low (hemi-) cellulose producing *Myceliophthora thermophila* C1 strain, essentially as described elsewhere (2, 3).

The crude *MtLPMO9I*-rich fermentation broth was filtrated and dialyzed against 10 mM potassium phosphate buffer pH 7.6 before chromatographic purification. First, anion exchange chromatography (AEC) was performed by using an ÄKTA-Explorer preparative chromatography system (GE Healthcare, Uppsala, Sweden). A Source 15Q column with a 100 × 70 mm internal diameter (GE Healthcare) was used and pre-equilibrated with two column volumes by 20 mM potassium phosphate buffer pH 7.6 (eluent A). Eluent B was 20 mM potassium phosphate buffer pH 7.6 containing 1 M NaCl. *MtLPMO9I*-containing samples were loaded onto the column and the unbound fraction was washed out with three column volumes of Eluent A. Elution was performed as follows: from 0% to 40% B in 15 column volumes; next from 40% to 100% over one column volume and finally two column volumes of 100% B. The flow rate was 20 mL/min and the elution was monitored by UV absorption at 220 and 280 nm. All fractions were collected (40 mL size) and immediately stored on ice. Peak fractions were adjusted to an approximate concentration of 2 mg/mL and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; see Section 6.1.3) to determine the *MtLPMO9I*-containing fractions. For a second purification step, size exclusion chromatography (SEC) was used. Hereto, combined *MtLPMO9I*-containing fractions were loaded on a self-packed Superdex™ 75 PG column (100 × 3 cm internal diameter, GE Healthcare) and eluted (5 mL/min) with a 10 mM potassium phosphate buffer (pH 7.0) containing 150 mM NaCl. Peak fractions (based on UV 280 nm) were collected (10 mL) and immediately stored on ice. These fractions again were adjusted to an approximate concentration of 2 mg/mL and analyzed by SDS-PAGE. SEC-purified *MtLPMO9I*-containing fractions were used directly for enzyme incubations.

For *MtLPMO9E*, three subsequent chromatographic purification steps were performed. Crude *MtLPMO9E*-rich fermentation broth was filtrated and dialyzed against 10 mM potassium phosphate buffer pH 7.0. First, the *MtLPMO9E*-containing enzyme preparation was loaded on a Source 15Q column (GE Healthcare) pre-equilibrated with 20 mM potassium phosphate buffer pH 7.0. After washing the unbound fraction (two column volumes), elution (flowrate 20 mL/min) was performed by using a linear gradient from 0 to 1 M NaCl in 20 mM potassium phosphate buffer pH 7.0 over one column volume. All fractions were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to an approximate concentration of 2 mg/mL and analyzed by SDS-PAGE to determine the *MtLPMO9E*-containing fractions. Subsequently, SEC purification was performed similarly as described for purification of *MtLPMO9I*. Combined SEC-purified *MtLPMO9E*-containing fractions were further

purified by using cation exchange chromatography (CEC) on an ÄKTA-Micro preparative chromatography system (GE Healthcare). *MtLPMO9E*-containing fractions were loaded on a Resource Q column (30 × 16 mm internal diameter, GE Healthcare) pre-equilibrated with 10 mM sodium acetate buffer (pH 5.0). By using a flow rate of 1 mL/min, the unbound fraction was washed (two column volumes) and elution was performed by using a linear gradient from 0 to 0.5 M NaCl in 10 mM sodium acetate buffer pH 5.0 over one column volume. All fractions were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to an approximate concentration of 2 mg/mL and analyzed by SDS-PAGE to determine the *MtLPMO9E* fractions. CEC-purified *MtLPMO9E*-containing fractions were used directly for enzyme incubations.

6.1.2. Protein content

The protein content of purified *MtLPMO9I* and *MtLPMO9E* was analyzed using the BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). Bovine serum albumin was used as calibration standard in a range of 20-2000 µg/mL.

6.1.3. Electrophoresis of enzyme samples to evaluate purity and molecular mass

SDS-PAGE was used to analyze the apparent molecular mass of *MtLPMO9I* and *MtLPMO9E* corresponding fractions and to determine their purity. Hereto, protein samples were reduced with dithiothreitol, heated for 15 min and loaded on 12% polyacrylamide gels (Mini-PROTEAN TGX Gels, Bio-Rad Laboratories, Hempel Hempstead, UK). In addition, a protein marker (Precision Plus ProteinTM Standards, Bio-Rad Laboratories) was loaded to calibrate the subunit mass. After the electrophoresis, gels were stained with InstantBlueTM (Expedeon, Heidelberg, Germany).

6.1.4. TFA hydrolysis of reduced cello-oligosaccharides to monosaccharides

Freeze-dried reduced cello-oligosaccharides were hydrolyzed with 1 mL of 2 M TFA at 121 °C for 1 h with closed caps. Subsequently, hydrolyzed samples were dried at 40 °C under a stream of nitrogen. To the dried samples 1 mL of water was added and dried again at 40 °C under a stream of nitrogen to remove the remaining TFA. The dried hydrolyzed samples were dissolved in 100 µL MilliQ-pure water prior to HPAEC analysis.

6.1.5. HPAEC analysis of monosaccharides after TFA hydrolysis

Monosaccharides were analyzed using the same HPAEC system (Thermo Fisher Scientific) with additional post-column (Thermo Fisher Scientific; addition of 0.1 M NaOH). The analysis was performed as published elsewhere (4). TFA hydrolyzed samples were diluted ten times prior to HPAEC post-column analysis. Glucose (25 µg/mL), galactose and sorbitol (10 µg/mL each) were used as standards.

6.2. Supplementary data

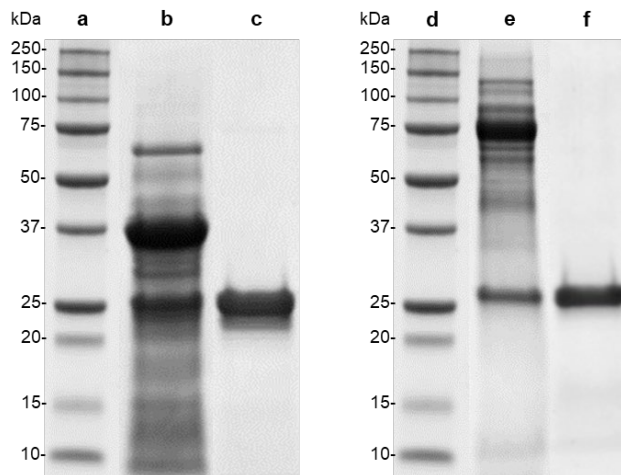


Fig. S1. SDS-PAGE analysis of *MtlPMO9E*- and *MtlPMO9I*-containing fractions before and after purification. *MtlPMO9E* (lane c) and *MtlPMO9I* (lane f) were step-wise purified from the crude fermentation broth rich of *MtlPMO9E* (lane b) and *MtlPMO9I* (lane e), respectively. In lanes a and d, the known molecular masses (kDa) of marker proteins are indicated.

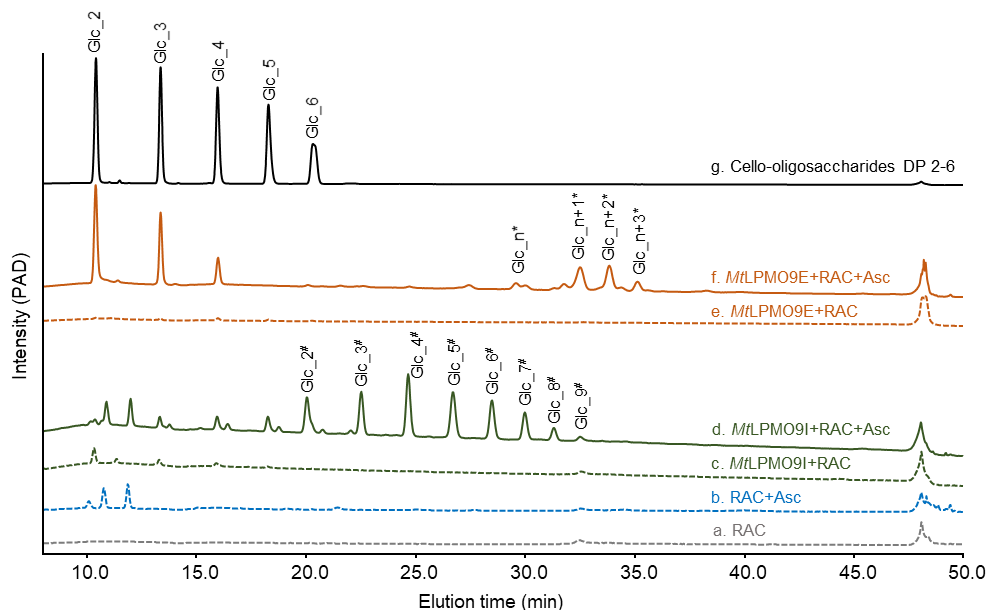


Fig. S2. HPAEC chromatograms of generated oligosaccharides from RAC incubations with *MtlPMO9E* and *MtlPMO9I* without SPE clean-up. Incubation of RAC with *MtlPMO9E* resulted in the release of C4-oxidized oligosaccharides (Glc_n^* ; f), while *MtlPMO9I* resulted in the generation of C1-oxidized oligosaccharides ($\text{Glc}_n^\#$, d), in the presence of ascorbic acid (Asc). Control reactions were (a) only RAC, (b) RAC with Asc, (c) *MtlPMO9I* incubated with RAC in the absence of Asc and (e) *MtlPMO9E* incubated with RAC in the absence of Asc. Cello-oligosaccharides DP2-6 were used as standards (g).

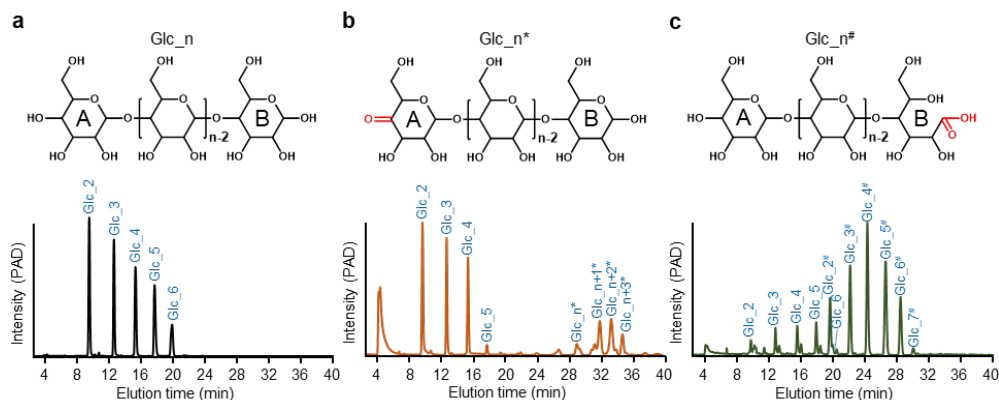


Fig. S3. Structures and HPAEC chromatograms (after SPE clean-up) of non-, C4- and C1-oxidized cello-oligosaccharides, indicated as Glc_n, Glc_{n*} and Glc_{n#}, respectively. (a) Non-oxidized cello-oligosaccharides of DP2–6 (Glc₂–6). (b) C4-oxidized cello-oligosaccharides (Glc_{n*}) having the C4-ketone group (indicated in red) on the A ring as present in the *MtLPMO9E*-RAC sample. In addition, non-oxidized cello-oligosaccharides are present as a result of the *MtLPMO9E* catalytic reaction towards RAC. (c) C1-oxidized cello-oligosaccharides (Glc₂–7[#]) having the carboxyl group on the B ring (gluconic acid group) as present in the *MtLPMO9I*-RAC sample. In addition, non-oxidized cello-oligosaccharides are present as a result of the *MtLPMO9I* catalytic reaction towards RAC. C4- and C1-oxidized cello-oligosaccharides were obtained from the incubation of RAC with either *MtLPMO9E* or *MtLPMO9I* in the presence of Asc (Fig. S2).

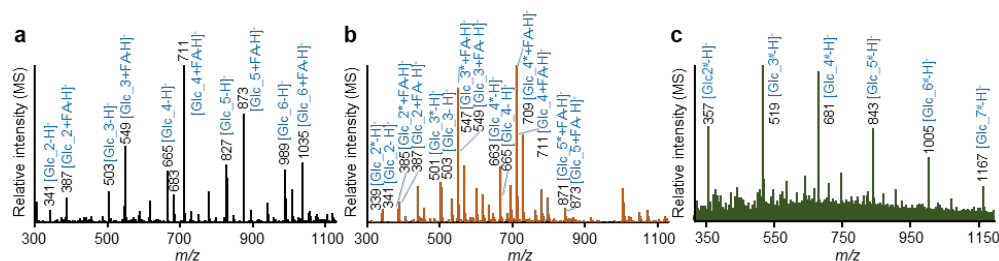


Fig. S4. HILIC-ESI full-scan mass spectra (elution time 4–40 minutes) of (a) standard cello-oligosaccharides DP2–6, (b) *MtLPMO9E*-RAC sample and (c) *MtLPMO9I*-RAC sample, all after SPE clean-up. (a) Non-oxidized cello-oligosaccharides (Glc_n) were present both in [Glc_n-H]⁺ and [Glc_n+FA-H]⁺. (b) C4-oxidized cello-oligosaccharides (Glc_{n*}) were present as 4-ketoaldoses indicated by their *m/z* of -2 compared to the Glc_n of the same DP. Glc_{n*} were present as [Glc_{n*}-H]⁺ and [Glc_{n*}+FA-H]⁺. (c) C1-oxidized cello-oligosaccharides (Glc_{n#}) were present as aldonic acids indicated by their *m/z* of +16 compared to the Glc_n of the same DP. Glc_{n#} were only present as [Glc_{n#}-H]⁺ ions. FA= formic acid. C4- and C1-oxidized cello-oligosaccharides were obtained from the incubation of RAC with either *MtLPMO9E* or *MtLPMO9I* in the presence of Asc (Fig. S2).

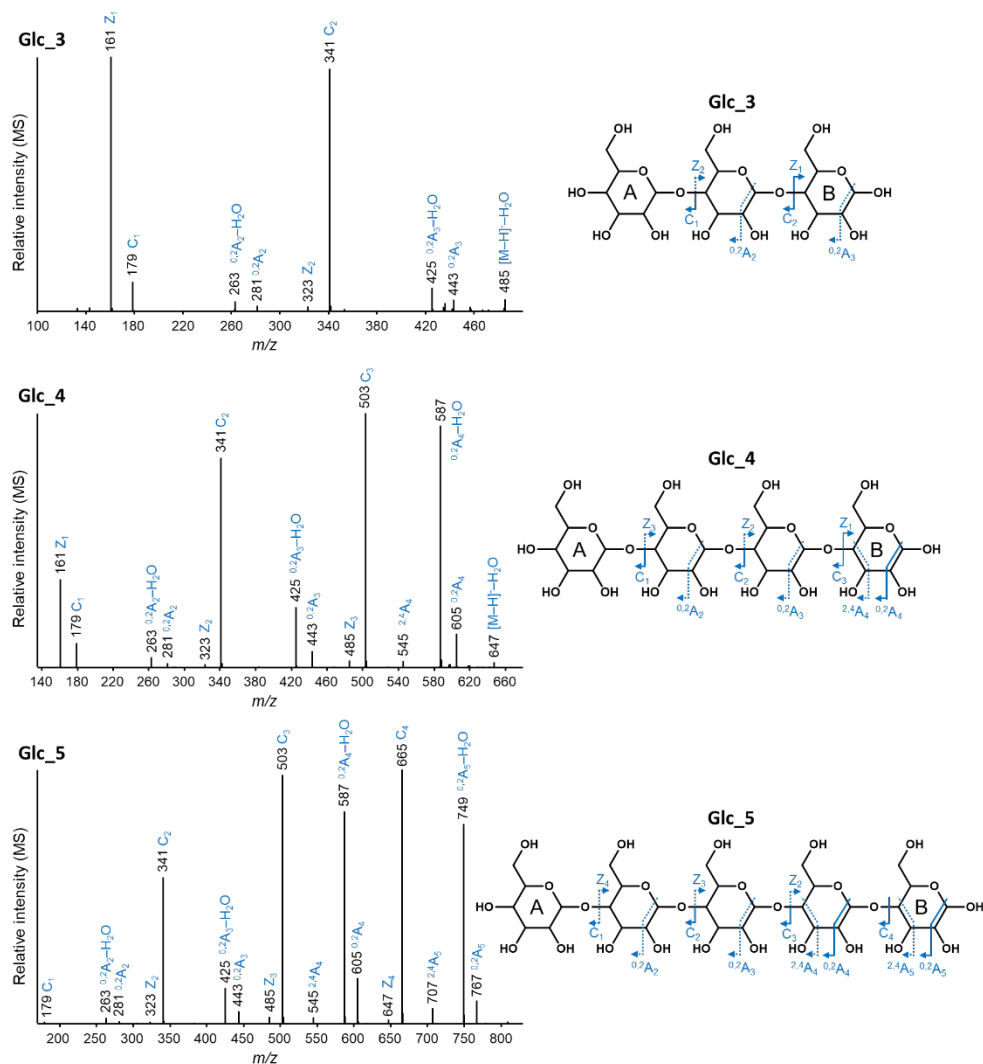


Fig. S5. Negative ion mode CID-MS² spectra of non-oxidized cello-oligosaccharides DP3-5 (Glc_3, m/z 503.2; Glc_4, m/z 665.3 and Glc_5, m/z 827.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (5). Blue solid arrows indicate the most abundant fragments, while dashed arrows represent fragments with relatively low intensity. A and B ring represent the non-reducing and reducing end, respectively.

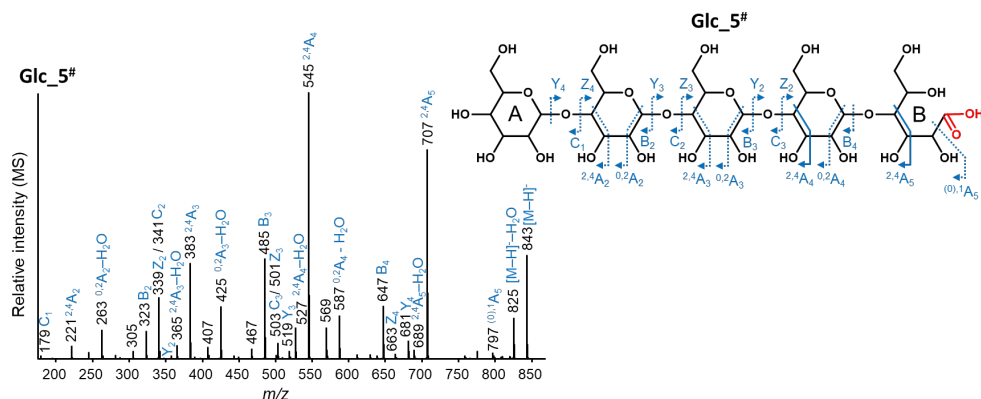


Fig. S6. Negative ion mode CID-MS² spectrum of the C1-oxidized cello-oligosaccharide DP5 (Glc_5#, *m/z* 843.3). See **Fig. 3** for DP3 (Glc_3#, *m/z* 519.2) and DP4 (Glc_4#, *m/z* 681.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (5). Blue solid arrows indicate most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. Carboxyl groups are indicated in red. A and B ring represent the non-reducing and gluconic acid end, respectively. C1-oxidized cello-oligosaccharides were generated by RAC incubated with *MtLPMO9I* in the presence of Asc (Fig. S2).

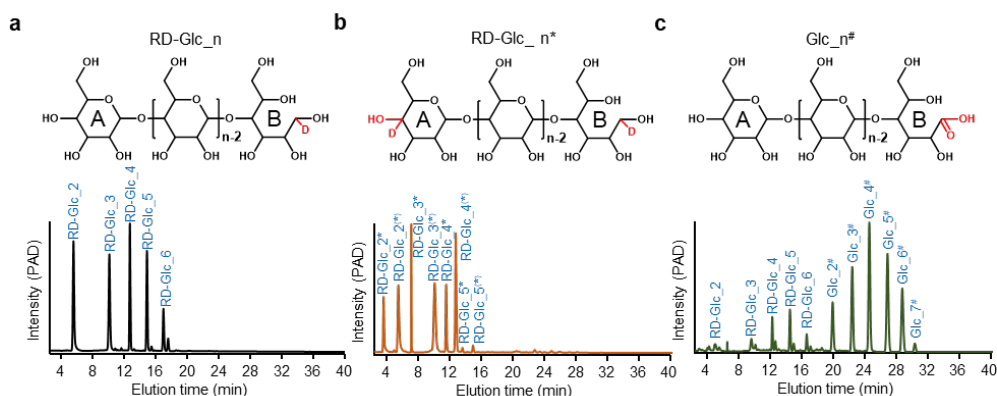


Fig. S7. Structures and HPAEC chromatograms (after SPE clean-up) of NaBD₄-reduced non-, C4- and C1-oxidized cello-oligosaccharides, indicated as RD-Glc_n, RD-Glc_n* and Glc_n#, respectively. (a) Reduced non-oxidized cello-oligosaccharides DP2-6 (RD-Glc₂-6). (b) Reduced C4-oxidized cello-oligosaccharides with resulting deuterium ions (indicated in red) on both the A and B ring. Reduction of the oxidized C4-group resulted both in glucosyl and galactosyl units, hence two series of oligosaccharides eluted in HPAEC (RD-Glc_n* and RD-Glc_n''). (c) C1-oxidized cello-oligosaccharides (Glc_n*) were non-reducible and remained in their aldonic acid form. C4- and C1-oxidized cello-oligosaccharides were obtained from the incubation of RAC with either *MtLPMO9E* or *MtLPMO9I* in the presence of Asc (Fig. S2).

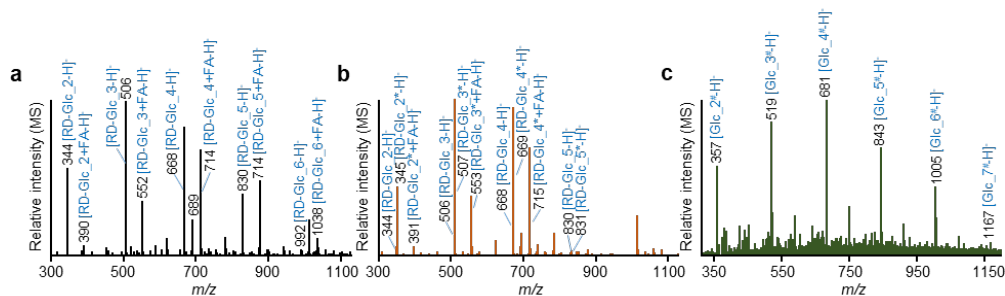


Fig. S8. HILIC-ESI full-scan mass spectra (elution time 4–40 minutes) of NaBD₄-reduced (a) standard cello-oligosaccharides DP2–6, (b) *MtlPMO9E*-RAC sample and (c) *MtlPMO9I*-RAC sample, all after SPE clean-up. (A) Reduced non-oxidized cello-oligosaccharides (RD-Glc_n) were present both in [RD-Glc_n-H]⁻ and [RD-Glc_n+FA-H]⁻. (B) Reduced C4-oxidized cello-oligosaccharides (RD-Glc_n^{*}) shown $m/z + 1$ compared to the RD-Glc_n of the same DP were also present as [RD-Glc_n^{*}-H]⁻ and [RD-Glc_n^{*}+FA-H]⁻. (C) C1-oxidized cello-oligosaccharides (Glc_n[#]) were non-reducible and remained in their deprotonated aldonic acid form [Glc_n[#]-H]⁻.

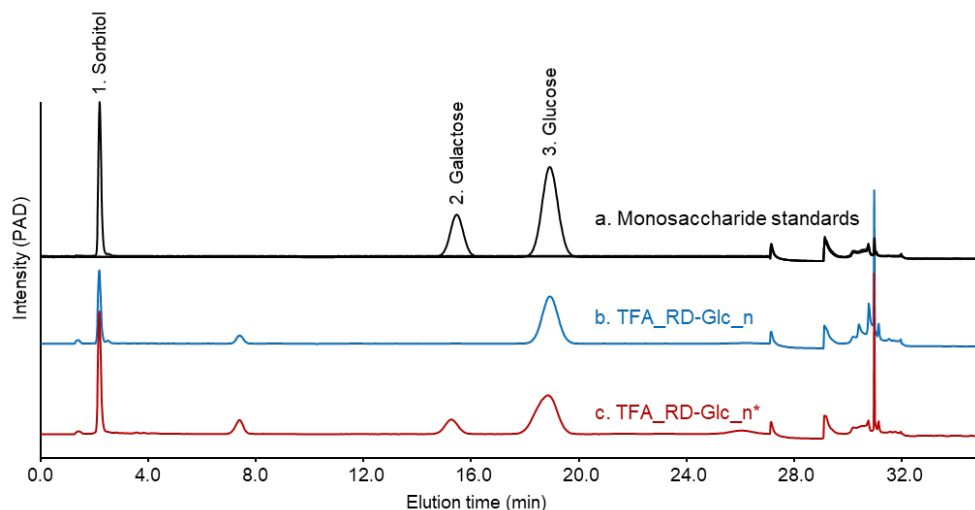


Fig. S9. HPAEC monosaccharide chromatograms of (a) monosaccharide standard (sorbitol, galactose and glucose; black line), (b) TFA hydrolyzed reduced native cello-oligosaccharides (TFA_RD-Glc_n, blue line) and (c) TFA hydrolyzed reduced C4-oxidized cello-oligosaccharides (TFA_RD-Glc_n^{*}, red line). Galactose was present in TFA_RD-Glc₂–5^{*} indicating that the galactosyl unit was formed during reduction of C4-oxidized cello-oligosaccharides.

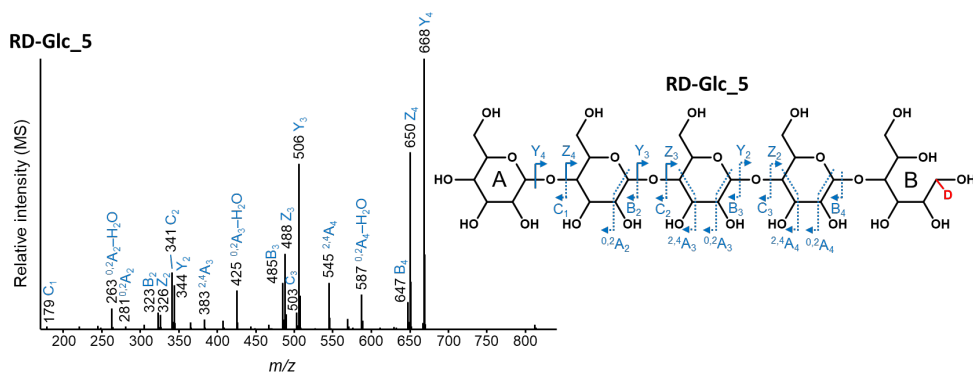


Fig. S10. Negative ion mode CID-MS² spectrum of the reduced non-oxidized cello-oligosaccharide DP5 (RD-Glc_5, m/z 830.3). See **Fig. 5** for DP3 (RD-Glc_3, m/z 506.3) and DP4 (RD-Glc_4, m/z 668.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (5). Blue solid arrows indicate the most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. The deuterium ion is indicated in red. A and B ring represent the non-reducing and deuterated alditol end, respectively.

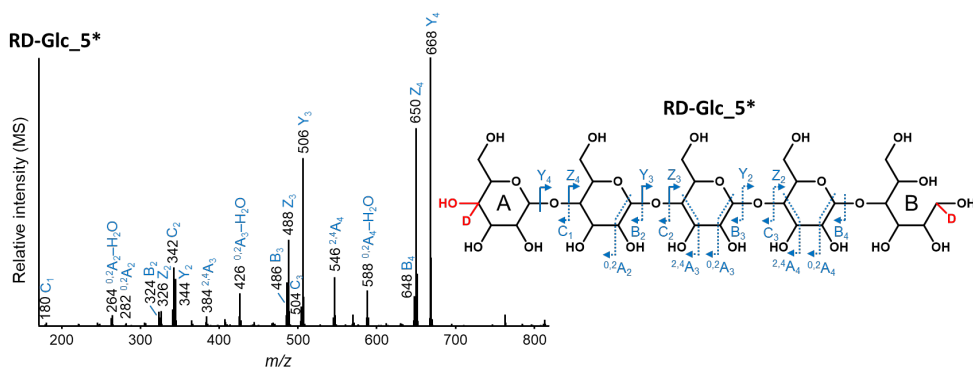


Fig. S11. Negative ion mode CID-MS² spectrum of the reduced C4-oxidized cello-oligosaccharide DP5 (RD-Glc_5*, m/z 831.3). See **Fig. 6** for DP3 (RD-Glc_3*, m/z 507.2) and DP4 (RD-Glc_4*, m/z 669.3); average spectra across chromatographic peaks. The fragments are annotated according to Domon and Costello (5). Blue solid arrows indicate the most abundant fragments, while dashed arrows indicate fragments with relatively lower intensity. The formation of hydroxyl groups from the 4-keto groups and deuterium ions are indicated in red. A and B ring represent the non-reducing and deuterated alditol end, respectively. Note that the A-ring can represent a glucosyl and galactosyl unit.

6.3. Supplementary references

1. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. PLOS ONE. 2018;13:e0202148.
2. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al., *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
3. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. Ind Biotechnol. 2011;7:214-23.
4. Sengkhamparn N, Verhoef R, Schols HA, Sajjaanantakul T, Voragen AG. Characterisation of cell wall polysaccharides from okra (*Abelmoschus esculentus* (L.) Moench). Carbohydr Res. 2009;344:1824-32.
5. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate J. 1988;5:397-409.

Chapter 4

Regioselective C4 and C6 double oxidation of cellulose by lytic polysaccharide monooxygenases

Peicheng Sun¹, Christophe V.F.P. Laurent^{2,3}, Vincent J.P. Boerkamp¹, Gijs van Erven¹, Roland Ludwig², Willem J.H. van Berkel¹ and Mirjam A. Kabel¹

¹Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

²Biocatalysis and Biosensing Laboratory, Department of Food Science and Technology, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

³Institute of Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

Abstract

Lytic polysaccharide monooxygenases (LPMOs) play a key role in enzymatic degradation of hard-to-convert polysaccharides, such as chitin and cellulose. It is widely accepted that LPMOs catalyze a single regioselective oxidation of the C1 or C4 carbon of a glycosidic linkage, after which the destabilized linkage breaks. We discovered a series of novel C4/C6 double oxidized cello-oligosaccharides. Products were characterized, aided by sodium borodeuteride reduction, and hydrophilic interaction chromatography coupled to mass spectrometric analysis. The C4/C6 double oxidized products were generated by C4 and C1/C4 oxidizing LPMOs, but not by C1 oxidizing ones. By performing incubation and reduction in H_2^{18}O , we confirmed that the C6 gem-diol structure resulted from oxygenation, though oxidation to a C6 aldehyde, followed by hydration to the C6 gem-diol, could not be excluded. Our findings can be extended to how the reactive LPMO-cosubstrate complex is positioned towards the substrate.

Keywords

Isotopic labeling; Lytic polysaccharide monooxygenase; Mass spectrometry; Oligosaccharides; Regioselectivity

1. Introduction

In context of fungal physiology, biocatalysis and biorefinery processes, lytic polysaccharide monooxygenases (LPMOs) are beneficial to the enzymatic degradation of hard-to-convert polysaccharides, such as cellulose and chitin, and, simultaneously, introduce oxidized functional groups (1-6).

LPMOs are classified as Auxiliary Activities (AA) families in the Carbohydrate-Active enZyme (CAZy) database (7, 8). Based on amino acid sequence similarity, LPMOs are currently categorized in AA9–AA11 and AA13–AA17 (9, 10). So far, AA9 is the largest LPMO family, in which all candidates are from fungal origin, and all characterized LPMOs are oxidatively cleaving cellulose (11, 12).

LPMOs exploit a mono-copper dependent active site, where the copper is coordinated by a unique histidine brace (4, 13-15). Although part of the LPMO's mono(per)oxygenase mechanisms remain under debate, it is generally accepted that an external reductant and a cosubstrate (O_2 or H_2O_2) is required for catalysis (11, 12, 16, 17). LPMOs catalyze a regioselective single oxidation reaction, of either the C1 or the C4 position in the glycosidic linkage, which results in cleavage of this destabilized linkage (11). In addition to non-oxidized products, C1 oxidation generates δ -lactones, while after C4 oxidation 4-ketoaldoses are formed. In aqueous solutions, however, δ -lactones convert to aldonic acids, while the 4-ketoaldoses are in pH-dependent equilibrium with the corresponding gem-diols (11, 12). The LPMO regioselectivity attracts a lot of scientific attention, as it reflects how LPMOs interact with their substrate, and how the copper-bound oxygen (or peroxide) species and the glycosidic linkage are positioned during the catalytic reaction (18-21).

So far, singly C1 and/or C4 oxidized cello-oligosaccharides, resulting from oxidative cleavage of cellulose, have been identified via a combination of chemically oxidized standards, high performance anion exchange chromatographic with pulsed amperometric detection analysis (HPAEC-PAD) and mass spectrometric (m/z) data (21-24). Following the same strategy, however, without including the corresponding standards, double, C1 and C4, oxidized cello-oligosaccharides have also been suggested as LPMO products (21-24). It can however be questioned whether the observed double oxidized products are indeed C1/C4 oxidized, or C1/C6 oxidized as suggested elsewhere (25), since both have the same m/z values. In this work, we analyzed different LPMO cellulose digests aided by sodium borodeuteride ($NaBD_4$) reduction and hydrophilic interaction chromatography–electrospray ionization–negative ion mode collision induced dissociation–mass spectrometry (HILIC-ESI-CID-MS/MS²). We found that C4 and C1/C4 oxidizing LPMOs generated C4/C6 double oxidized cello-oligosaccharides having a gem-diol structure at C6 carbon. Our findings provide new insights into the LPMO-mediated regioselective oxidation of cellulose, which is of high relevance for biorefinery applications.

2. Materials and methods

2.1. Enzymatic incubation of AA9 LPMOs with RAC

Materials, production and purification of AA9 LPMOs used in this study are described in Supplementary information. LPMO-RAC incubation was based on Sun et al. (26). 2 mg/mL of RAC was mixed with 50 mM (final concentration) ammonium acetate buffer (pH 5.0). Subsequently, 1 μ M of LPMO was added to RAC mixture with and without 1 mM Asc (final concentration). Control reactions were performed in the absence of LPMO with and without Asc. All reactions (500 μ L total volume) were incubated at 30 °C by using an Eppendorf Thermomixer® comfort (Hamburg, Germany), placed in an almost vertical direction, at 800 rpm in duplicate. At 6 h and 16 h, another 1 μ M of *Mt*LPMO9H or *Mt*LPMO9E was added to their corresponding incubation tubes (in total 3 μ M), while the same volume of water was added to other LPMO-RAC incubation tubes. The reactions continued till 24 h and were stopped by immediately separating supernatants after centrifugation at 22000 $\times g$ for 10 mins at 4 °C in a benchtop centrifuge (Z 233 MK-2, HERMLE Labortechnik GmbH, Wehingen, Germany). The resulting supernatants were collected and part of supernatants were reduced by NaBD₄ followed by solid phase extraction (SPE) as described below (27). The non-reduced digests were diluted five times before HPAEC-PAD analysis. A mixture of cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose (1 μ g/mL each) were analyzed as standards.

2.2. Reduction of generated (oxidized) cello-oligosaccharides with NaBD₄ and clean-up with SPE

Reduction with NaBD₄ was performed according to Sun et al. (26). A clean-up procedure for reduced digests was carried out by using SPE with Supelclean™ ENVI-Carb™ columns (3 mL, Sigma-Aldrich) as described previously (26). The dried reduced samples were dissolved in 60% (v/v) acetonitrile in water. The reduced RAC digests were directly used for HILIC-ESI-CID-MS/MS² analysis or diluted twenty times for HPAEC-PAD analysis.

2.3. Enzymatic incubation and reduction in H₂¹⁸O

Prior to enzymatic incubation, RAC, buffer and enzyme stock solution were prepared in H₂¹⁸O. 0.7 mL of RAC suspension was thoroughly mixed with 0.7 mL of H₂¹⁸O followed by centrifugation 22000 $\times g$ for 10 min at 4 °C in a table centrifuge. Subsequently, the supernatant was discarded and H₂¹⁸O was added to bring back to the same volume. These steps were repeated another four times to achieve that H₂¹⁸O in RAC was >99.9%. 50 mM (final concentration in the incubation) ammonium acetate buffer (pH 5.0) containing 1 mM Asc (final concentration in the incubation) was freshly prepared in H₂¹⁸O prior to incubation. Stock solutions of *Mt*LPMO9E, *Mt*LPMO9H, *Nc*LPMO9C and *Nc*LPMO9M were concentrated ten times by using Amicon® Ultra-0.5 centrifugal filter units (Merck Millipore, U.S.A.) with a 3 kDa cutoff. Afterwards, 50 mM ammonium acetate buffer (pH 5.0) in H₂¹⁸O was added to bring back to the same volume. These cycles were repeated another four times to achieve that H₂¹⁸O in enzyme stock solution was >99.9%. The enzymatic incubation with RAC was

performed as described above. The reduction was carried out by using NaBD₄ in H₂¹⁸O. The SPE procedures were performed as described above. The dried reduced samples were dissolved in 60% (v/v) acetonitrile in H₂¹⁸O just prior to analysis.

2.4. HPAEC-PAD analysis for profiling oligosaccharides

Samples and standards were analyzed by a HPAEC-PAD with an ICS-5000 system (Dionex, Sunnyvale, California, U.S.A.) equipped with a CarboPac PA-1 column (2 mm ID × 250 mm; Dionex) in combination with a CarboPac PA guard column (2 mm ID × 50 mm; Dionex). Two mobile phases were (A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH and other setting have been described previously (26, 28). Non-reduced samples and standards were analyzed with a 92.5 min elution program: 0–45 min, linear gradient 0–25% B; 45–67.5 min, linear gradient 25–50% B; 67.5–77.5 min isocratic gradient 100% B and 77.5–92.5 isocratic gradient 0% B. The elution program for analysis of reduced samples and standards has been described previously (26, 28). Chromatographic data were processed by using Chromeleon 7.2.10 software (Thermo Fisher Scientific, Waltham, Massachusetts, U.S.A.).

2.5. HILIC-ESI-CID-MS/MS² for structural elucidation of reduced cello-oligosaccharides

The reduced RAC digests were analyzed by HILIC-ESI-CID-MS/MS² on a Vanquish UHPLC system (Thermo Fisher Scientific) coupled to an LTQ Velos Pro mass spectrometer (Thermo Fisher Scientific). The UHPLC and MS (negative ion mode) settings, column and mobile phases have been described previously (26). The elution program was also described previously (29). MS² was performed under dependent scan mode as described previously with the only modification of Activation Q set to 0.25 (26). Mass spectrometric data were processed by using Xcalibur 2.2 software (Thermo Fisher Scientific). All chemical structures used were created by using ChemDraw 18.0.0.231 (PerkinElmer, Waltham, Massachusetts, U.S.A.).

3. Results and discussion

We first analyzed, by using HPAEC-PAD (Fig. S1), regenerated amorphous cellulose (RAC) digested with published C1/C4 oxidizing LPMOs from two different fungi: LPMO9H from *Myceliophthora thermophila* C1 (*MtLPMO9H*) (23) and LPMO9M from *Neurospora crassa* (*NcLPMO9M*) (11). The oligosaccharides eluting after 45 min correspond to the previously suggested double, C1 and C4, oxidized cello-oligosaccharides (23). The same double oxidized oligomers, eluting after 45 min, were detected in the RAC digests of the distinct C4 oxidizing *MtLPMO9E* (26) and *NcLPMO9C* (19, 30, 31), though in (relatively) lower amounts, and not in C1 oxidizing *MtLPMO9B* (28, 32)- and *NcLPMO9F* (19, 31)-RAC digests (Fig. S1).

Next, to further understand the position of the oxidations in the double oxidized oligomers, these digests were reduced by NaBD₄ and subsequently analyzed by HPAEC-PAD (Fig. S2) and HILIC-ESI-CID-MS/MS² (**Fig. 1** and Fig. S3) (26, 29). Reduction with NaBD₄ reduces the reducing end of the oligosaccharides under addition

of a deuterium atom (D) (see structures in **Fig. 1**), which distinguishes this reduced end in MS² fragmentation spectra from the non-reducing end otherwise having the same m/z values. In addition, the C4 oxidized end becomes reduced leading to a “regular” non-reduced end glycosyl unit, but with addition of D ($m/z + 1$ compared to hydrogen atom (H)). Both HPAEC-PAD and HILIC-MS showed that, besides reduced non-oxidized (RD-G2–6) and reduced C4 oxidized cello-oligosaccharides (RD-GALC4oxG2–7 and RD-GLC4oxG2–7 due to the formation of galactosyl or glucosyl ends after reduction) (26), an additional series of reduced oligomers was observed (Fig. S3). Reduced oligomers of this additional series were collected after HILIC analysis, and a re-analysis by HPAEC-PAD confirmed that these products corresponded to the additional series defined by HPAEC-PAD, of not yet defined, possibly double, (reduced) oxidized products (Fig. S4).

Further, based on m/z values obtained from the base peak chromatogram (Fig. S3), and the extracted ion chromatogram (**Fig. 1a**) of the HILIC-MS/MS² analysis, the additional series indicated to be a series of reduced C4/C6 double oxidized cello-oligosaccharides (DP2–7; RD-C4C6oxG2–7). To explain in brief, for example, for RD-C4C6oxG4 (**Fig. 1b**, others in Fig. S5), m/z 685.5 ($[M-H]^-$) was detected. In our previous study, we showed that i) NaBD₄ reduction of non-oxidized cello-oligosaccharides (e.g., RD-G4, m/z 668.5) increases their m/z values ($m/z + 3$) due to the reduction and insertion of one D to the reduced alditol end; ii) NaBD₄ reduction of C4-oxidized cello-oligosaccharides (e.g., RD-C4oxG4, m/z 669.5) leads to an m/z increase of +6 compared to the non-reduced ones (e.g., C4oxG4, m/z 663.5), since the C4 ketone end is reduced to –OH with one D inserted at the C4 carbon ($m/z + 3$), and reduction and insertion of one D to the reduced alditol end ($m/z + 3$); iii) (C4 oxidized) gem-diol forms and C1 oxidized cello-oligosaccharides (aldonic acid form) cannot be reduced by NaBD₄ (26). Based on these findings, the in literature proposed double C1/C4 oxidized cello-oligosaccharides (e.g., C1C4oxG4, m/z 679.5 (C4 ketone)/ m/z 697.5 (C4 gem-diol); Fig. S6) should, after NaBD₄ reduction lead to an m/z increase of +3 in case of a C4 ketone end, or remain the same m/z after NaBD₄ reduction (Fig. S6). The detected m/z 685.5 ($[M-H]^-$, **Fig. 1b**, others in Fig. S5) did not match with the m/z of RD-C1C4oxG4, however, matched with a reduced form of a C4 ketone end having a C6 gem-diol group (**Fig. 1**). Corresponding m/z values of reduced double, C1/C4, oxidized cello-oligosaccharides were not detected. Further, a detailed MS² analysis of m/z 685.5 confirmed this product as a reduced form of a C4 oxidized cellotetraose having an additional (oxidized) C6 gem-diol group (**Fig. 1** and Fig. S7). In particular, m/z values of fragments Y₃ and S₄ showed the presence of an additional gem-diol group on the (NaBD₄ reduced) C4 ketone end of the oligosaccharides.

The combined results of the NaBD₄ reduction and the HILIC-MS/MS² analysis of the double oxidized series in the LPMO-RAC digests, can be translated into the structure of these products prior to reduction: a C4 ketone and a C6 gem-diol group in the same product (**Scheme 1**).

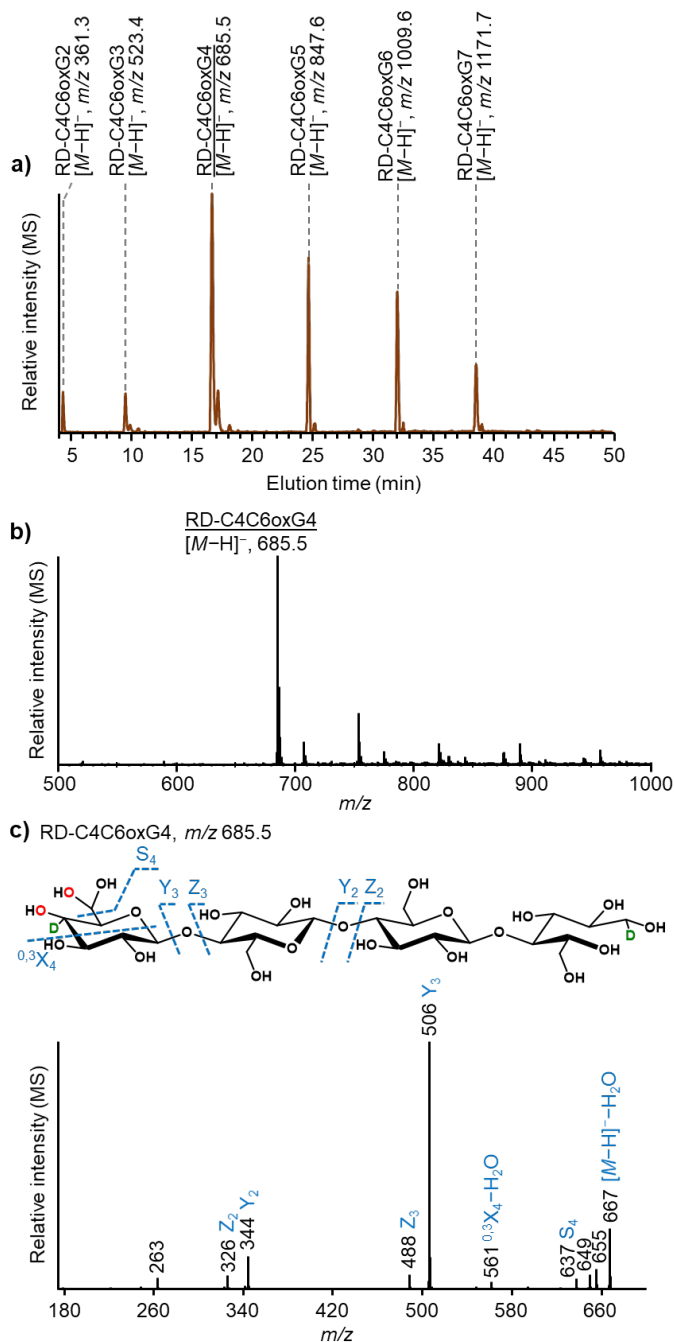


Fig. 1. LPMO-generated C4/C6-double oxidized cello-oligosaccharides after NaBD₄ reduction in H₂O. (a) HILIC-ESI-MS (negative mode; [M-H]⁻) extracted ion chromatogram of the reduced C4/C6-double oxidized cello-oligosaccharides from DP2 to DP7 (RD-C4C6-oxG2–7) from the *Nc*LPMO9M-RAC digest; similar data were obtained for *Mt*LPMO9H⁻, *Nc*LPMO9C⁻ and *Mt*LPMO9E⁻

RAC digests. (b) Representative example of an MS spectrum of the reduced double oxidized cellotetraose (RD-C4C6-oxG4, m/z 685.5 ($[M-H]^-$)). (c) Representative example of MS² spectrum of RD-C4C6-oxG4. The oxygen atom from molecular oxygen is indicated in red and the deuterium atom is indicated in green. MS² spectra of RD-C4C6-OxG2-3 and RD-C4C6-OxG5-6 are shown in Fig. S7. Annotation of fragments is according to the nomenclature developed by Domon and Costello (33).

LPMOs have been shown to incorporate molecular oxygen at the C4 position, resulting in a C4 ketone, when oxygen is used as cosubstrate. The LPMO route to a C6 gem-diol formation has yet to be defined. We hypothesize two routes (**Scheme 1**): **I** the addition of a second hydroxyl group at the C6 carbon is catalyzed by the LPMO, via molecular oxygen, similar as suggested for C4 oxidation; **II** the LPMO oxidizes the C6-OH into a C6 aldehyde, which hydrates further into the C6 gem-diol (34). The major difference is that in route **I** the second oxygen atom in the C6 gem-diol group stems also from LPMO, while in route **II** one oxygen atom is from water.

To test the origin of the inserted oxygen atoms in the C4 ketone and C6 gem-diol group, we performed the LPMO-RAC incubations, and the NaBD₄ reduction, in H₂¹⁸O. In these digests, reduced single C4 oxidized cello-oligosaccharides showed a m/z of +2 compared to the m/z of corresponding RD-C4ox generated in H₂O (Fig. S8). Corresponding MS² spectra suggest the presence of an ¹⁸O atom (from H₂¹⁸O) at the C4 carbon of the non-reducing end. To confirm that ¹⁸O exchange occurred during reduction, H₂O and H₂¹⁸O digests without reduction were subjected to HILIC-MS/MS². Although the HILIC separation was poor, and the MS signal intensity of non-reduced samples was less pronounced compared to the reduced products, we could still collect proper MS and MS² spectra. From these spectra, it can be concluded that both the H₂O and H₂¹⁸O digests comprised m/z values consistent with ¹⁶O incorporation at the non-reducing end of the C4 oxidized products. For example, for C4oxG4 m/z 663.5 ($[M-H]^-$) was found in both H₂O and H₂¹⁸O (Fig. S9 and S10).

Prior to reduction, in H₂¹⁸O the m/z +2 values were also abundant, for example m/z 665.5 ($[M-H]^-$). This m/z 665.5 can represent non-oxidized G4 or C4oxG4 having ¹⁸O at the C4 of the non-reducing end (¹⁸O-C4oxG4), which coelutes in HILIC. They obtain the same m/z values that cannot be distinguished by MS and MS². Nevertheless, since the m/z +2 values (e.g., 665.5; Fig. S9 and S10) were more abundant in H₂¹⁸O than in H₂O, we suggest that the ¹⁶O-C4 ketone is hydrated in H₂¹⁸O to a C4 gem-diol (e.g., with both -¹⁶OH and -¹⁸OH), which subsequently converts back to the ¹⁶O- and ¹⁸O-C4 ketones (35). After the reduction in H₂¹⁸O, mostly ¹⁸O-C4 ketones end-products are observed due to the favored side of the C4 ketone over the gem-diol in this equilibrium (Fig. S11).

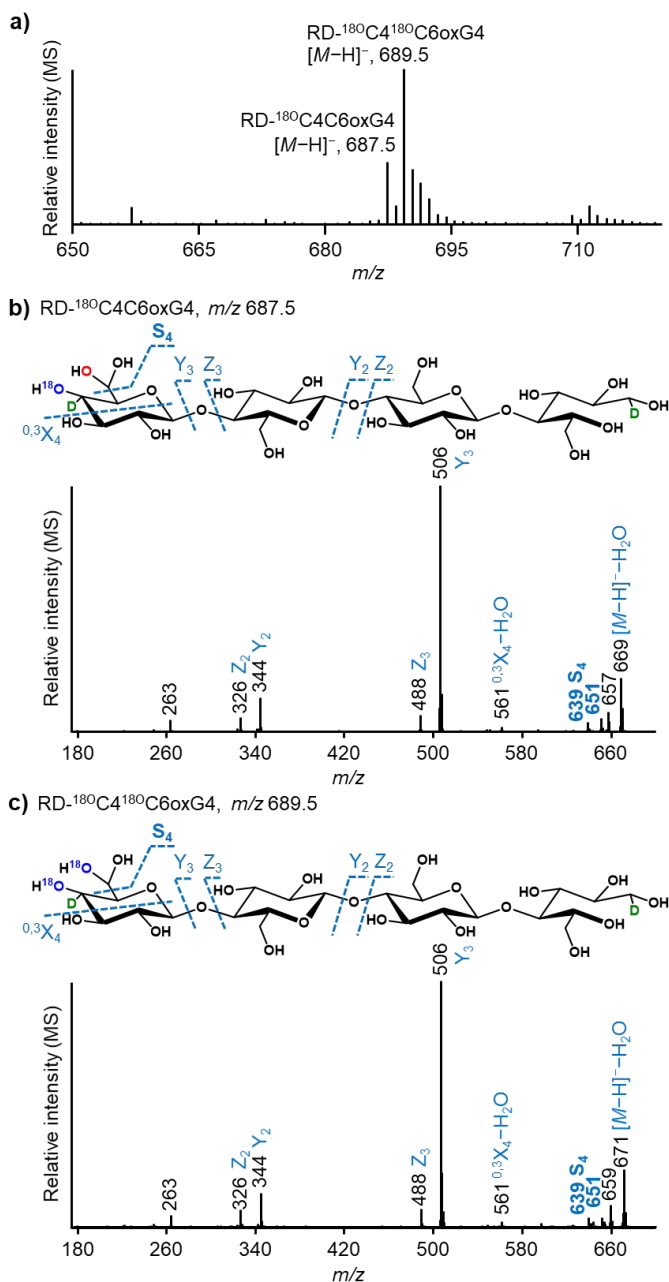
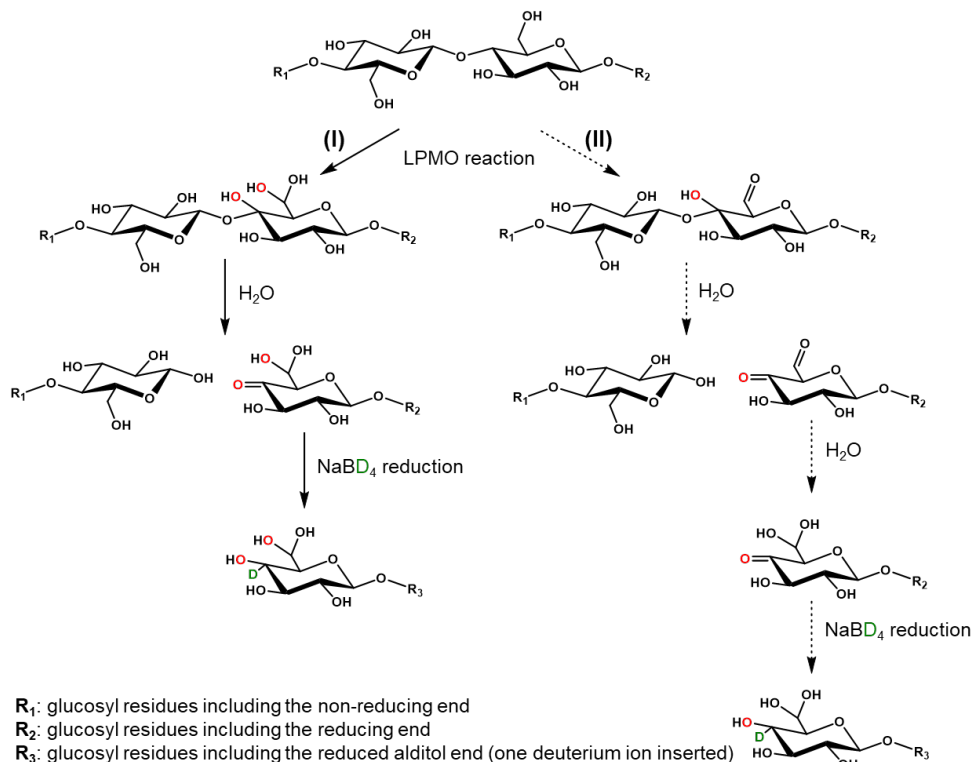


Fig. 2. LPMO-generated C4/C6-double oxidized cello-oligosaccharides after NaBD₄ reduction in H₂¹⁸O. a) Representative example of an MS spectrum indicating the m/z of reduced double oxidized cellotetraose (digest and reduction in H₂¹⁸O; RD- $^{180}\text{C}_4\text{C}_6\text{-oxG}_4$, m/z 687.5 [$M-H$] $^-$ and RD- $^{180}\text{C}_4^{180}\text{C}_6\text{-oxG}_4$, m/z 689.5 [$M-H$] $^-$). MS² spectra of (b) RD- $^{180}\text{C}_4\text{C}_6\text{-oxG}_4$ and (c) RD- $^{180}\text{C}_4^{180}\text{C}_6\text{-oxG}_4$. The oxygen atom from molecular oxygen or from H₂¹⁸O is indicated in red and blue, respectively. The deuterium atom is indicated in green. Annotation of fragments is according to the nomenclature developed by Domon and Costello (33).

With respect to the reduced double oxidized products in the H₂¹⁸O digests (**Fig. 2a**), compared to those in H₂O (**Fig. 1b**), the MS spectra showed a series of m/z +2 and of +4 values (e.g., m/z 687.5 [$M-H$]⁻; RD-¹⁸O C₄C₆oxG₄, and m/z 689.5 [$M-H$]⁻; RD-¹⁸O C₄¹⁸O C₆oxG₄). The corresponding MS² fragmentation patterns showed that the one double oxidized series (e.g., m/z +2) contained ¹⁸O at the non-reducing end only (**Fig. 2b**). Notably, the m/z loss resulting in the MS² S₄ fragment (e.g., loss of C₆ gem-diol from the parental molecule) was the same in RD-C₄OxG₄ and RD-¹⁸O C₄OxG₄ (both 47), again indicating that ¹⁶O was found in the C₆ gem-diol group, even in H₂¹⁸O (**Fig. 1c and 2b**). This diagnostic observation confirmed route **I**: a second hydroxyl group in the C₆ gem-diol is formed via LPMO (per)oxygation (**Scheme 1** and Fig. S12). The other double oxidized series (e.g., m/z +4) was identified, via corresponding MS² spectra, to have ¹⁸O both at the non-reducing end and at the C₆ carbon (**Fig. 2c**; RD-¹⁸O C₄¹⁸O C₆oxG₄), confirmed by the mass loss of 49 of the S₄ fragment (**Fig. 2c**). Again, as described above for ¹⁸O exchange at the C₄ position (e.g., via gem-diol formation), formation of RD-¹⁸O C₄¹⁸O C₆oxG₄ possibly results from RD-¹⁸O C₄C₆oxG₄, by exchange of ¹⁶O to ¹⁸O in H₂¹⁸O (Fig. S12). However, route **II** (**Scheme 1**) cannot be fully excluded as RD-¹⁸O C₄¹⁸O C₆oxG₄ may also form from hydration of the C₆ aldehyde in H₂¹⁸O (Fig. S12). Prior to reduction, m/z values of C₆ aldehyde structures were recognized (only DP3 and 4 were seemingly present; Fig. S13).

However, due to poor separation of non-reduced samples and too low MS intensities of these products, the data were not conclusive, and route **II** was not confirmed nor disapproved. Oxidation to C₆ aldehydes followed by hydration to C₆ gem-diols has been described as products from galactose oxidases and raffinose oxidases (36-38). These enzymes oxidize the C₆ carbon of a galactosyl residue to a galactoaldehyde, followed by hydration to the C₆ gem-diol, hemiacetal formation and other minor modifications (36-38). C₆ aldehyde and the two latter products were not observed to be significantly present in the LPMO-RAC digests, and products with C₆ gem-diol structure were predominant. These results suggested that most likely C₆ gem-diol is formed directly via oxygenation catalyzed by LPMOs, as demonstrated in route **I**.

We described two routes how C₄/C₆ double oxidized cello-oligosaccharides formed, however, it remains to be answered whether C₄/C₆ double oxidized cello-oligosaccharides were generated simultaneously with C₄ oxidized ones, or that C₆ oxidation was generated after C₄ oxidation took place. However, the chance that C₆ oxidation occurred at (early on) formed oxidized C₄ ends, and *vice versa* is considered to be very low. More likely would be that in the same LPMO-substrate complex both C₄ and C₆ oxidation took place. The latter is strengthened by our observation of (only) formation of C₄/C₆ double oxidized cello-oligosaccharide, and C₄ oxidized cello-oligosaccharides in the LPMO reactions studied, also at earlier incubation time points (data not shown). We attempted to pick up if single C₆ oxidized products were formed. Since without further cleavage of the glycosidic linkages, only C₆ oxidized cellulose will not become soluble, we have investigated cellulases hydrolyzed residues of RAC and of an LPMO-RAC digest with proton nuclear magnetic resonance (¹H NMR) spectroscopy (Fig. S14). For comparison, the soluble fraction of the LPMO-RAC digest was also analyzed by NMR.



Scheme 1. Routes for the LPMO catalyzed generation of C4/C6-double oxidized cello-oligosaccharides. In route I, LPMO catalyzes the insertion of oxygen atoms at both C4- and C6-carbon atoms in the cellulose. The reaction at the C4-carbon will destabilize the glycosidic bond, leading to the bond cleavage and formation of a C4-ketone in water. After NaBD₄ reduction, the C4-ketone is reduced to glucosyl and galactosyl residues (only one structure is used to illustrate the scheme), and the C6-gem-diol remains unaltered. Route **I** is confirmed by performing the reaction and reduction in H₂¹⁸O. In route **II**, the LPMO catalyzes the insertion of one oxygen atom at the C4-carbon, but oxidizes the C6-carbon into a C6-aldehyde group. In water, a C4-ketone is formed, and the C6-aldehyde prefers to convert to a C6-gem-diol. The subsequent reduced product is the same as in route **I**. Route **II** could not be confirmed nor disproved. The routes for generation of C4/C6-double oxidized cello-oligosaccharides in H₂¹⁸O are presented in Fig. S11. The scheme does not distinguish whether the C4- and C6- oxidation occurs simultaneously or that C4-oxidation occurs prior to C6-oxidation.

Based on these NMR results, we observed basically no differences in the expected gem-diol region (based on chemical shift predictions) in both RAC residues, while in the same region signals were observed for the soluble part of the LPMO digest (Fig. S14). Thus, we concluded that there is no evidence that single C6 oxidation occurred. Nevertheless, further research is recommended to study the order of events of C4 and C6 double oxidation, which will shed more light on how the cosubstrate, the carbohydrate and the enzyme are oriented.

4. Conclusion

In this research, we characterized, for the first time, a series of novel LPMO-catalyzed C4/C6 double oxidized cello-oligosaccharides by NaBD₄ reduction and HILIC-ESI-CID-

MS/MS². We found that C4/C6 double oxidized cello-oligosaccharides were generated by C4 and C1/C4 oxidizing LPMOs, but not by C1 oxidizing ones. By performing LPMO-RAC incubation and NaBD₄ reduction in H₂¹⁸O, we confirmed that the C6 gem-diol structure resulted from oxygenation, though oxidation to a C6 aldehyde, followed by hydration to the C6 gem-diol, could not be excluded.

5. References

- Corrêa TL, dos Santos LV, Pereira GA. AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol*. 2016;100:9-16.
- Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33:747-61.
- Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.
- Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørli M, et al. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219-22.
- Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, et al. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20:1479-83.
- Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*. 2010;49:3305-16.
- Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6:41.
- Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490-D5.
- Ipsen JO, Hallas-Møller M, Brander S, Lo Leggio L, Johansen KS. Lytic polysaccharide monooxygenases and other histidine-brace copper proteins: structure, oxygen activation and biotechnological applications. *Biochem Soc Trans*. 2021;49:531-40.
- Sabbadin F, Urresti S, Henrissat B, Avrova AO, Welsh LRJ, Lindley PJ, et al. Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes. *Science*. 2021;373:774.
- Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6:1399-406.
- Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134:890-2.
- Quinlan RJ, Sweeney MD, Leggio LL, Otten H, Poulsen J-CN, Johansen KS, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A*. 2011;108:15079-84.
- Tandrup T, Frandsen KEH, Johansen KS, Berrin J-G, Lo Leggio L. Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochem Soc Trans*. 2018;46:1431.
- Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J Mol Biol*. 2008;383:144-54.
- Bissaro B, Rohr AK, Müller G, Chylenski P, Skaugen M, Forsberg Z, et al. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat Chem Biol*. 2017;13:1123-8.
- Wang BJ, Wang ZF, Davies GJ, Walton PH, Rovira C. Activation of O₂ and H₂O₂ by lytic polysaccharide monooxygenases. *ACS Catal*. 2020;10:12760-9.
- Kim S, Stahlberg J, Sandgren M, Paton RS, Beckham GT. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc Natl Acad Sci U S A*. 2014;111:149-54.
- Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
- Danneels B, Tanghe M, Desmet T. Structural features on the substrate-binding surface of fungal Lytic polysaccharide monooxygenases determine their oxidative regioselectivity. *Biotechnol J*. 2019;14:1800211.

21. Forsberg Z, Mackenzie AK, Sorlie M, Rohr AK, Helland R, Arvai AS, et al. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2014;111:8446-51.
22. Fanuel M, Garajova S, Ropartz D, McGregor N, Brumer H, Rogniaux H, et al. The *Podospira anserina* lytic polysaccharide monooxygenase PaLPMO9H catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnol Biofuels*. 2017;10:63.
23. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
24. Huttner S, Varnai A, Petrovic DM, Bach CX, Anh DTK, Thanh VN, et al. Specific xylan activity revealed for AA9 lytic polysaccharide monooxygenases of the thermophilic fungus *Malbranchea cinnamomea* by functional characterization. *Appl Environ Microbiol*. 2019;85:e01408-19.
25. Bey M, Zhou S, Poidevin L, Henrissat B, Coutinho PM, Berrin JG, et al. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospira anserina*. *Appl Environ Microbiol*. 2013;79:488-96.
26. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx EJ, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
27. Irwin DC, Spezio M, Walker LP, Wilson DB. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. *Biotechnol Bioeng*. 1993;42:1002-13.
28. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SWA, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
29. Sun P, Laurent CVFP, Scheiblbrandner S, Frommhagen M, Kouzounis D, Sanders MG, et al. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnol Biofuels*. 2020;13:95.
30. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.
31. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
32. Sun P, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng*. 2021;9:14124-33.
33. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J*. 1988;5:397-409.
34. García-Alles LF, Zahn A, Erni B. Sugar recognition by the glucose and mannose permeases of *Escherichia coli*. Steady-state kinetics and inhibition studies. *Biochemistry*. 2002;41:10077-86.
35. Reusch W. Virtual Textbook of Organic Chemistry. Michigan: Michigan State University; 1999.
36. Parikka K, Master E, Tenkanen M. Oxidation with galactose oxidase: multifunctional enzymatic catalysis. *J Mol Catal B: Enzym*. 2015;120:47-59.
37. Parikka K, Tenkanen M. Oxidation of methyl α -D-galactopyranoside by galactose oxidase: products formed and optimization of reaction conditions for production of aldehyde. *Carbohydr Res*. 2009;344:14-20.
38. Andberg M, Møllerup F, Parikka K, Koutaniemi S, Boer H, Juvonen M, et al. A novel *Colletotrichum graminicola* raffinose oxidase in the AA5 family. *Appl Environ Microbiol*. 2017;83:e01383-17.

6. Supplementary information

6.1. Materials and methods

6.1.1. Materials

Cellobiose, Cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose were purchased from Megazyme (Bray, Ireland). Regenerated amorphous cellulose (RAC) was prepared from AVI (Avicel® PH-101, Sigma-Aldrich, St. Louis, Missouri, USA) as described previously (1). Ascorbic acid (Asc) was purchased from VWR International (Radnor, Pennsylvania, USA) and sodium borodeuteride (NaBD₄) was purchased from Sigma-Aldrich. Water (H₂¹⁶O) used was produced via a Milli-Q system (Millipore, Molsheim, France). H₂¹⁸O (GMP grade, ¹⁸O >97%) was purchased from ABX advanced biochemical compounds-Biomedizinische Forschungsreagenzien GmbH (Radeberg, Germany). Deuterated water (D₂O, D >99%) was purchased from Eurisotop (SaintAubin, France). Celluclast® 1.5L was purchased from Novozymes (Bagsværd, Denmark) and fractionated before use as described previously (2). Other chemicals were purchased from either Sigma-Aldrich or VWR International.

6.1.2. Production and purification of AA9 LPMOs

The genes encoding *MtLPMO9B* (MYCTH_80312; UniProt ID, G2QCJ3), *MtLPMO9E* (MTCTH_79765, UniProt ID: G2Q7A5) and *MtLPMO9H* (MYCTH_46583, UniProt ID: G2Q9T3) were homologously expressed in a low protease/low (hemi-) cellulose producing *Myceliophthora thermophila* C1 strain (IFF Nutrition & Biosciences, Leiden, The Netherlands), as described elsewhere (2, 3). Purification of the resulting *MtLPMO* enzymes has been reported previously (4, 5). Production and purification of *NcLPMO9C*, *NcLPMO9F* and *NcLPMO9M* have been described elsewhere (6, 7).

6.1.3. Proton nuclear magnetic resonance spectroscopy analysis of LPMO-RAC residue and supernatant of AA9 LPMOs

Prior to proton nuclear magnetic resonance (¹H NMR) spectroscopy analysis, the insoluble residue from the *MtLPMO9H*-RAC digest was treated with the cellulase cocktail Celluclast® 1.5L (48.58 mg protein/mL) in order to solubilize the residue. Hereto, 5 µL of ten times diluted (with D₂O) Celluclast® 1.5L was added to repeatedly D₂O washed untreated RAC and to *MtLPMO9H*-digested RAC residue suspension (400 µL; ~1 mg dry matter content; washed five times with 1 mL D₂O). A control sample was prepared with only the 5 µL of ten times diluted (with D₂O) cellulase cocktail Celluclast® 1.5L in 400 µL D₂O. These three samples were incubated at 50 °C in an Eppendorf Thermomixer® comfort (Hamburg, Germany) rotating at 800 rpm. After every hour of incubation, another 5 µL of ten times diluted Celluclast® 1.5L was added to the corresponding samples, with 7 additions in total. Afterwards, samples were immediately transferred to 5 mm NMR tubes and stored at 4 °C before NMR analysis. After the extensive cellulase treatment, still some insoluble material was observed for both RAC samples. For comparison, supernatants from the *MtLPMO9H*-RAC digest and from the incubation of RAC with only Asc were prepared for NMR analysis. These

supernatants (1 mL each) were dried under nitrogen flow, redissolved in 450 μL D_2O , and transferred to 5 mm NMR tubes. 1 mg of the standard cellopentaose was dissolved in 450 μL D_2O . For NMR measurement, the D_2O contained a trace amount of methanol which was used as internal reference (δ_{H} 3.34 ppm). Spectra were recorded on a Bruker Avance HD 700 MHz NMR spectrometer (Bruker BioSpin, Switzerland) equipped with a 5 mm BBI-probe. The internal temperature of the probe was set to 298 K. In total 8192 scans were collected using an FID size of 64k, a spectral width of 20 ppm, and offset of 4.7 ppm (δ_{H} 14.7 to -5.3 ppm). The number of scans was set this high to be able to detect minor signals as well. The relaxation time was set to 2 s, the acquisition time to 2.3 s, the 90° pulse length was automatically determined ($\sim 8 \mu\text{s}$), and the receiver gain was maximized. Water suppression was achieved using the standard Bruker pulse program "noesygppr1d", where the water signal is pre-saturated during the relaxation delay, mixing time, and spoil gradient. The same acquisition parameters as above were used for the standard cellopentaose, except for the number of scans which was reduced to 128. The data was processed in MestReNova v14.1 (Mestrelab Research, S.L., Santiago de Compostela, Spain). For the Fourier transform, an exponential window function was applied using a line broadening factor of 0.3 Hz.

6.2. Supplementary data

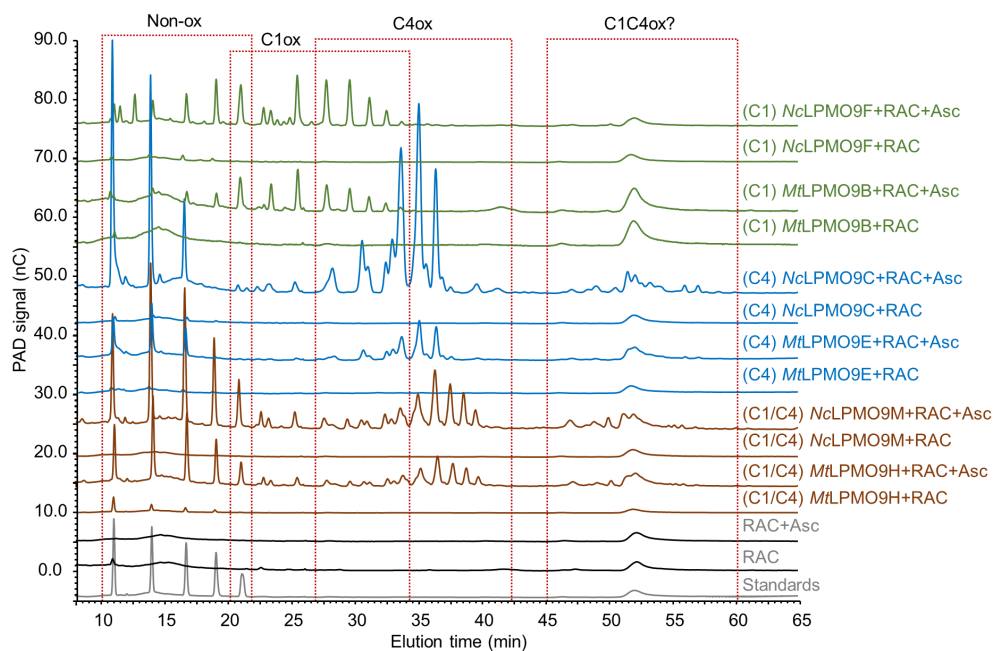


Fig. S1. HPAEC chromatograms of LPMO-RAC digests and control samples. *MtLPMO9H* and *NcLPMO9M* were previously characterized as C1/C4 oxidizing LPMOs (8, 9). *MtLPMO9E* and *NcLPMO9C* were previously characterized as C4 oxidizing LPMOs (5, 10). *MtLPMO9B* and *NcLPMO9F* were previously characterized as C1 oxidizing LPMOs (6, 11). Elution time regions of non-, C1, C4 and C1/C4 double oxidized cello-oligosaccharides are indicated as non-ox, C1ox, C4ox and C1C4ox, respectively. A standard mixture of cellobiose, cellotriose, cellotetraose,

cellopentaose and cellohexasaccharide (standards) is also shown. The assignment of C1C4ox has been suggested previously (8, 12). Ascorbic acid (Asc) was the reducing agent for the LPMO reactions.

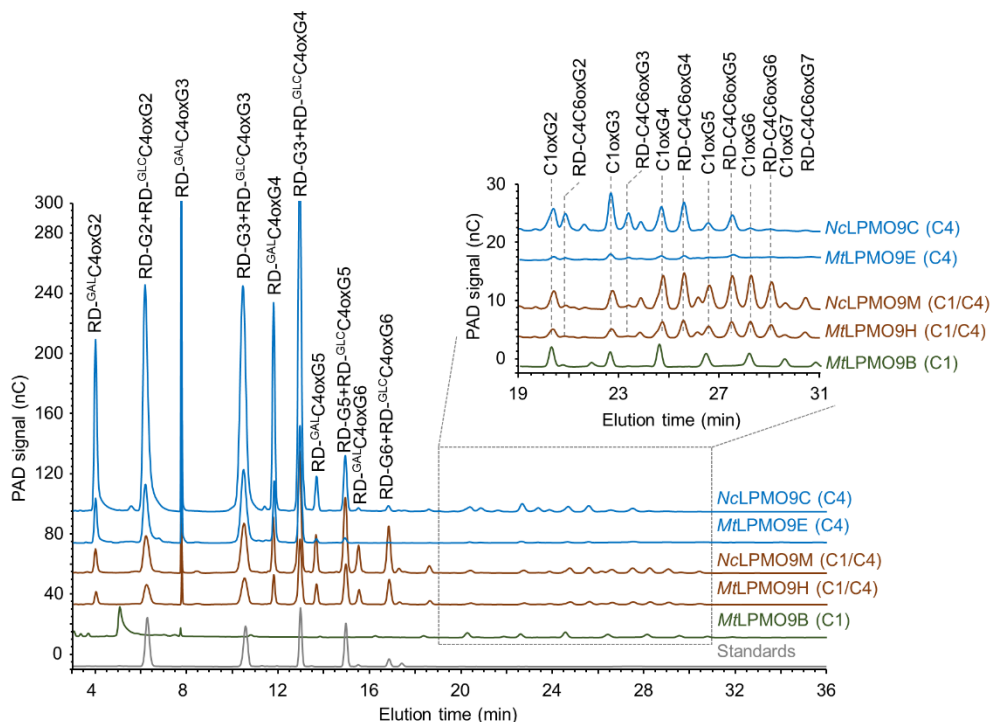


Fig. S2. HPAEC chromatograms of reduced *MtLPMO9H*-, *NcLPMO9M*-, *MtLPMO9E*- and *NcLPMO9C*-RAC digests. After reduction, non-oxidized cello-oligosaccharides are converted to their alditol forms (RD-Gn), C4 oxidized ones are converted to their alditol forms with either glucosyl or galactosyl residues at the non-reducing end (RD-^{GLCC}C4oxGn and RD-^{GALC}C4oxGn) (5), and C4C6 double oxidized ones are converted to their corresponding alditol forms, which are described in the main text (RD-C4C6oxGn). RD-^{GLCC}C4oxGn and RD-^{GALC}C4oxGn were annotated based on NaBD₄ reduction of i) cello-oligosaccharide standard mixture (cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexasaccharide) (RD-G2–RD-G6, in grey) and ii) β-glucosidase hydrolyzed C4 oxidizing LPMO-RAC digest, as described previously (5). *MtLPMO9B*-RAC digest is also included for the annotation of C1 oxidized cello-oligosaccharides (C1oxGn). A standard mixture of cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexasaccharide after reduction (standards) is also shown. All digests shown were prepared with ascorbic acid as the reducing agent for the LPMO reactions.

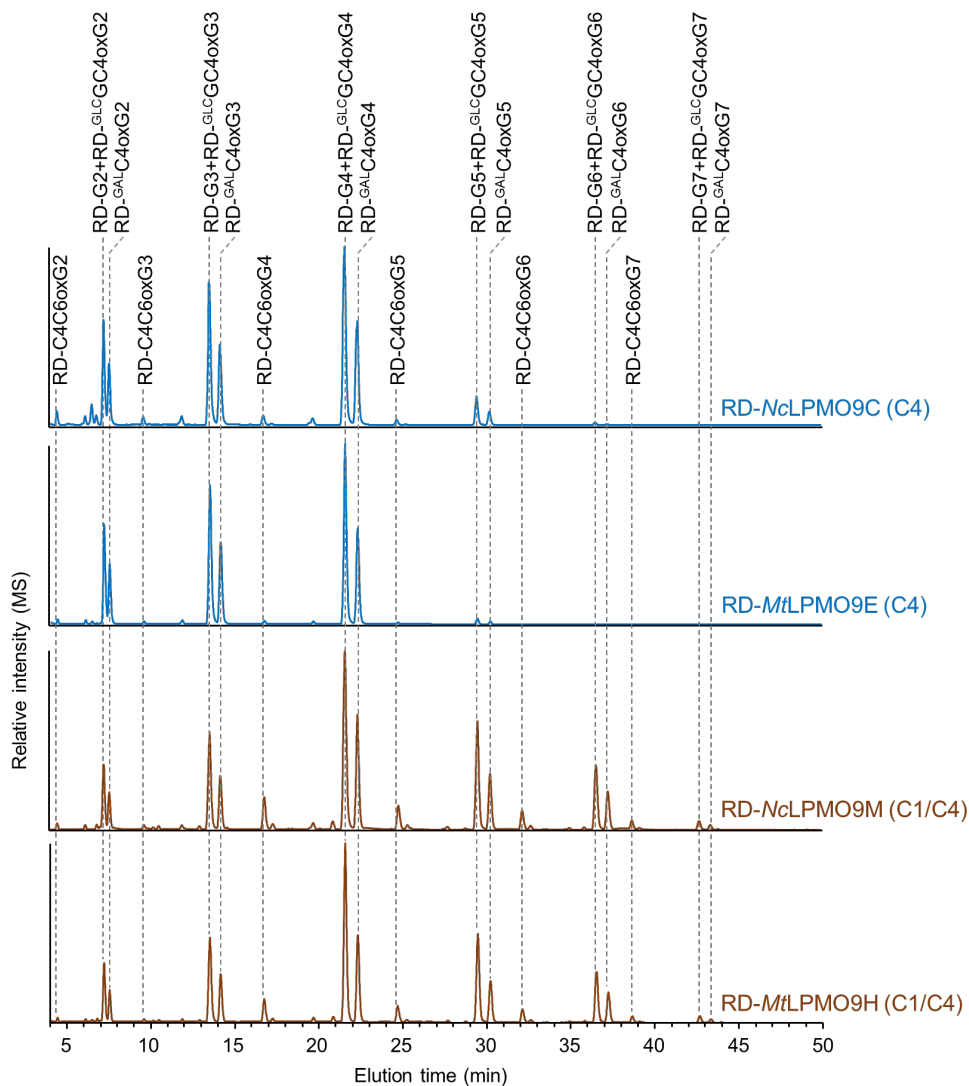


Fig. S3. HILIC-ESI-MS base peak chromatograms of reduced *MtLPMO9H*-, *NcLPMO9M*-, *MtLPMO9E*- and *NcLPMO9C*-RAC digests (RD-*MtLPMO9H*, RD-*NcLPMO9M*, RD-*MtLPMO9E* and RD-*NcLPMO9C*). All digests shown were prepared with ascorbic acid as the reducing agent for the LPMO reactions. After reduction, C4/C6 double oxidized cello-oligosaccharides DP2–7 (RD-C4C6oxG2 to RD-C4C6oxG7) were identified in addition to non- (RD-G2 to RD-G7) and C4 oxidized ones (RD-^{GLC}C6oxG2 to RD-^{GLC}C4oxG7; RD-^{GAL}C6oxG2 to RD-^{GAL}C4oxG7).

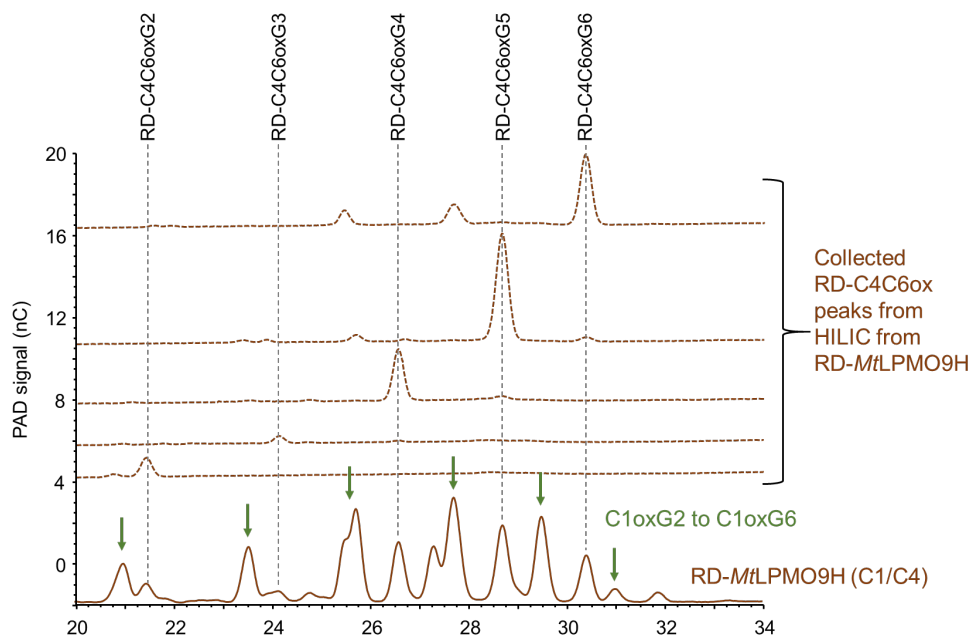


Fig. S4. HPAEC chromatograms of RD-C4C6-oxG2 to RD-C4C6-oxG6 collected from HILIC separation prior to MS analysis. Green arrows indicate the C1oxG2 to C1oxG6.

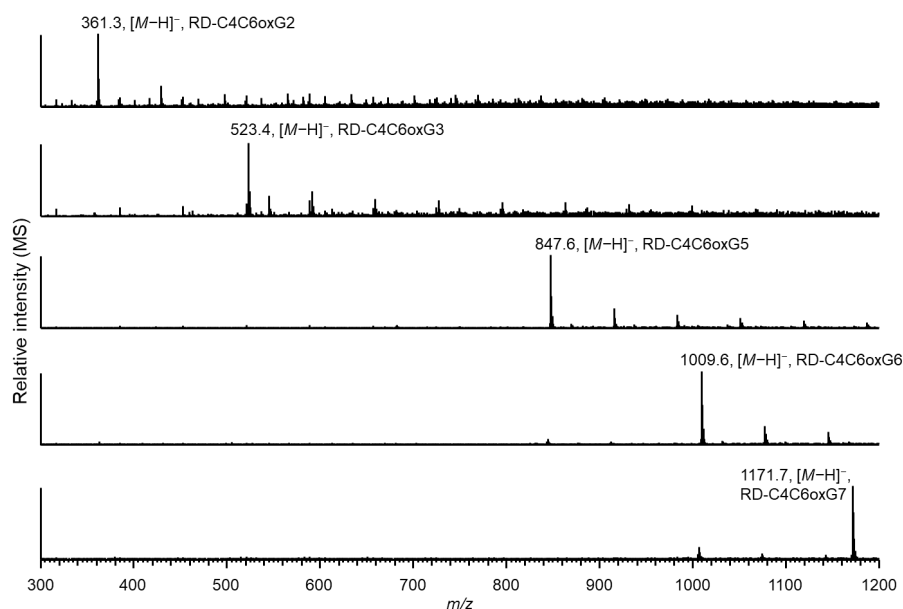


Fig. S5. HILIC-ESI-MS (negative mode; $[M-H]^-$) of RD-C4C6-oxG2 (m/z 361.3), RD-C4C6-oxG3 (m/z 524.3), RD-C4C6-oxG5 (m/z 847.6), RD-C4C6-oxG6 (m/z 1009.6) and RD-C4C6-oxG7 (m/z 1171.7). HILIC-ESI-MS of RD-C4C6-oxG4 is shown in **Fig. 1**.

C4/C6-double oxidized cello-oligosaccharides

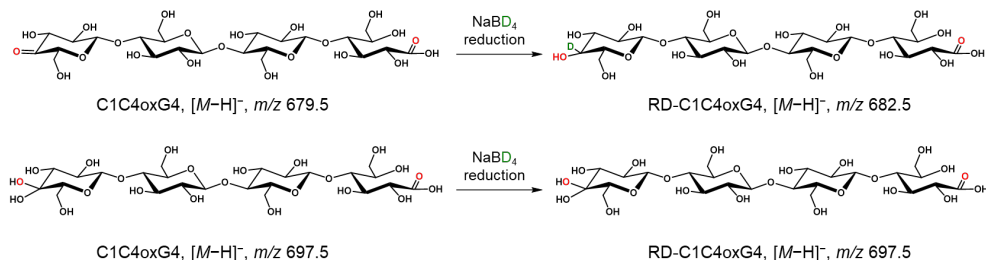
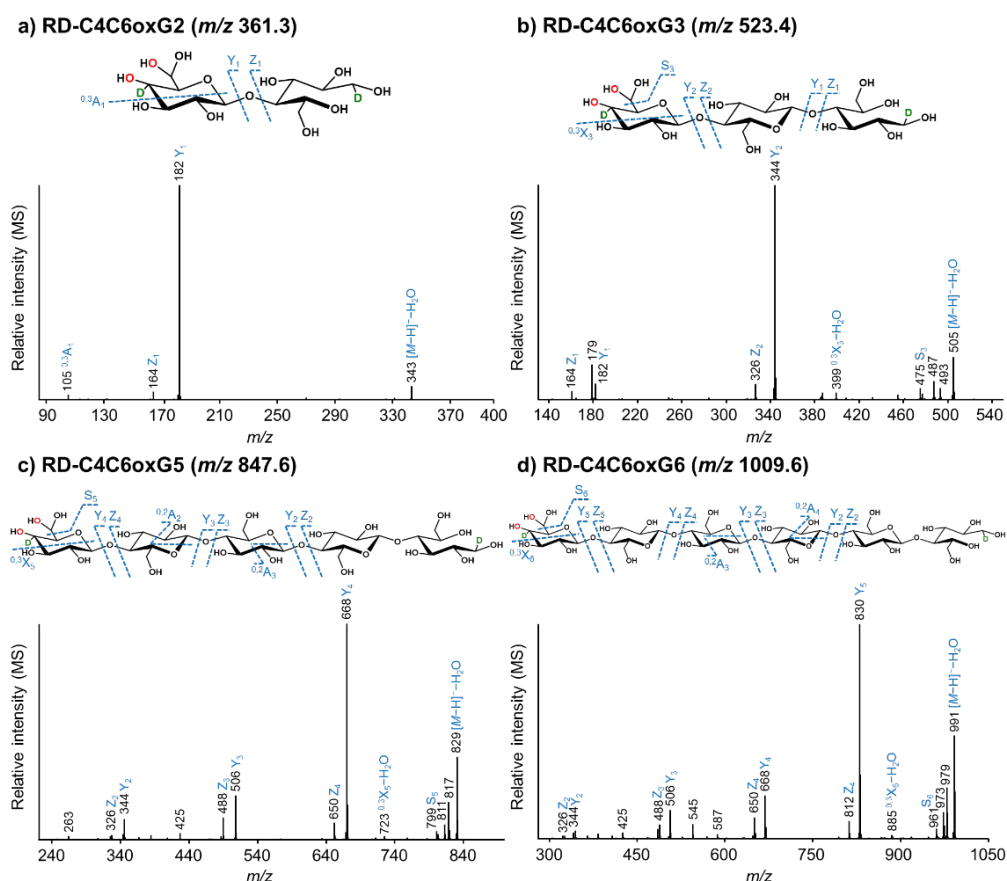


Fig. S6. Previously suggested two forms of C1/C4-double oxidized cello-oligosaccharides (e.g., DP4, C1C4oxG4). m/z values here are calculated and based on negative ion mode ($[M-H]^-$). The C4 ketone form of C1C4ox has a m/z difference of 3 before and after NaBD₄ reduction (e.g., C1C4oxG4, m/z 679.5 and RD-C1C4oxG4, m/z 682.5). If the C1C4ox is in the C4-gem-diol form, the m/z remains the same after reduction, because the C1-aldehydic acid and C4-gem-diol structure cannot be reduced. The oxygen atom from molecular oxygen is shown in red.



and *NcLPMO9C-RAC* digests. The oxygen atom from molecular oxygen is shown in red and the deuterium atom is shown in green. Annotation of fragments is according to the nomenclature developed by Domon and Costello (13).

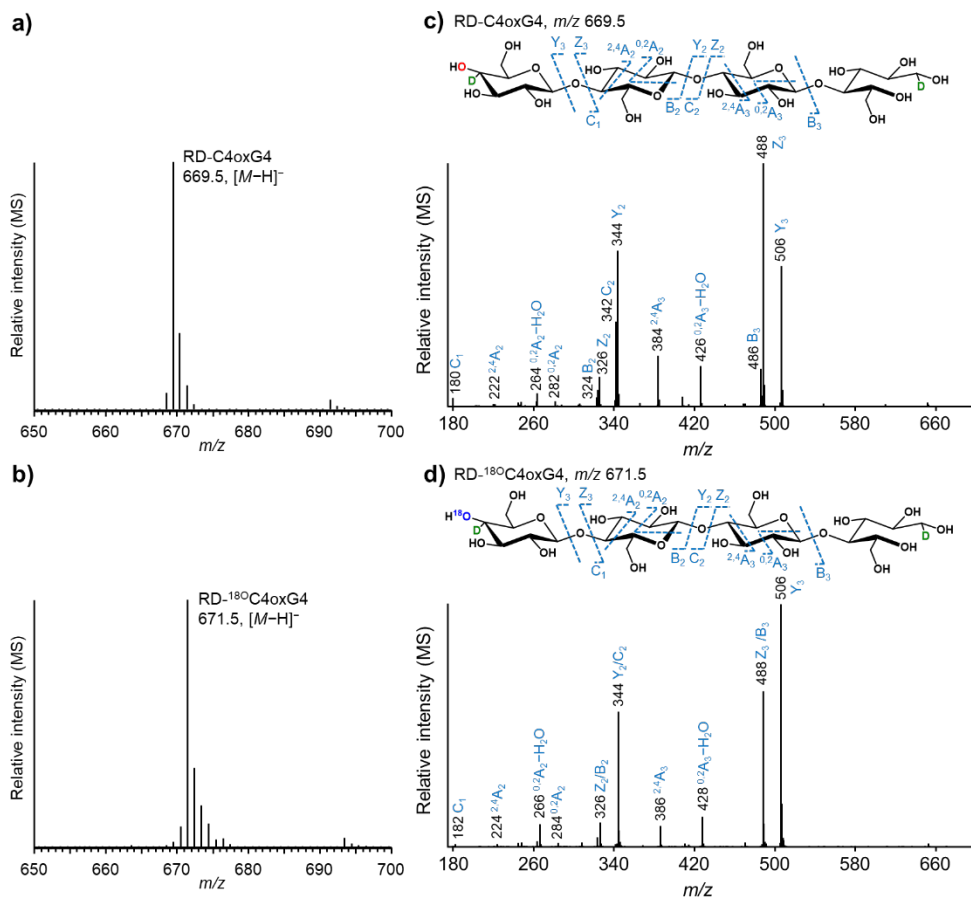


Fig. S8. HILIC-ESI-MS of RD-C4oxG4 (m/z 669.5, a) and RD-¹⁸O-C4oxG4 (m/z 671.5, b). RD-C4oxG4 and RD-¹⁸O-C4oxG4 are examples of reduced C4 oxidized cello-oligosaccharides generated in H₂O and H₂¹⁸O, respectively. Negative ion mode CID-MS² fragmentation patterns of RD-C4oxG4 (m/z 669.5, c) and RD-¹⁸O-C4oxG4 (m/z 671.5, d). The oxygen atom from molecular oxygen or from H₂¹⁸O is indicated in red and blue, respectively. The oxygen atom comes from molecular oxygen (red) and there is slow exchange to ¹⁸O (blue) in H₂¹⁸O. The deuterium atom is indicated in green. Annotation of fragments is according to the nomenclature developed by Domon and Costello (13).

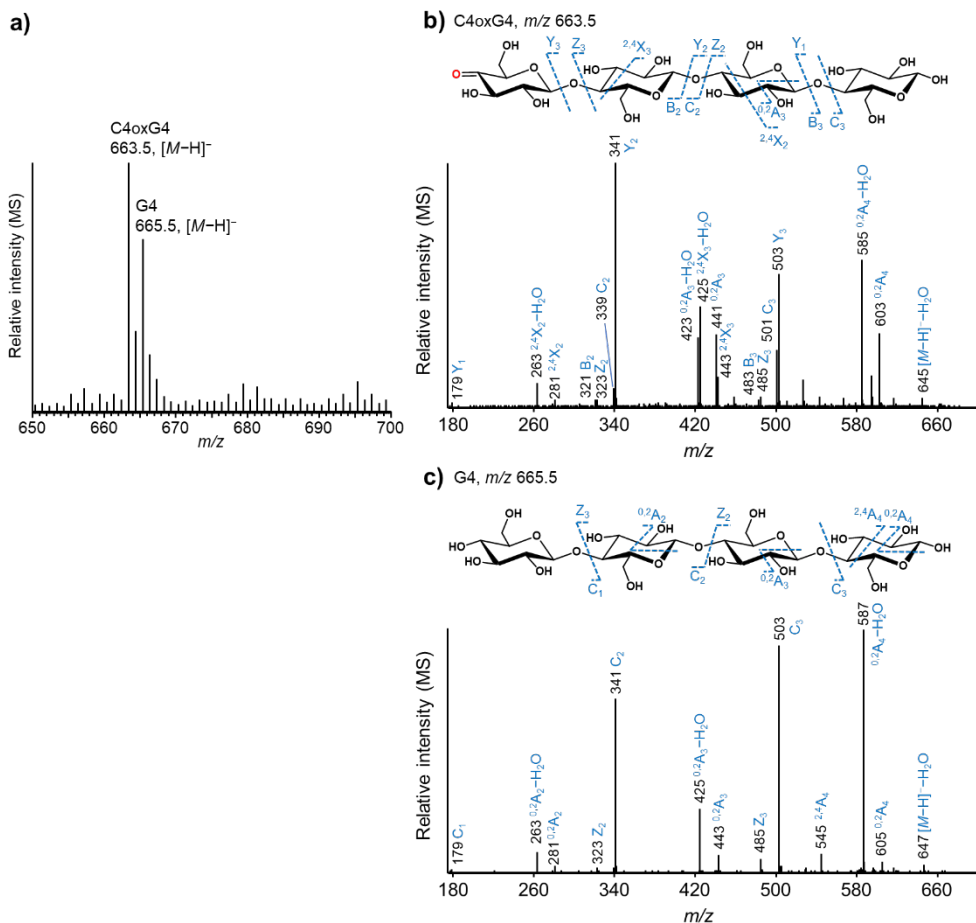


Fig S9. HILIC-ESI-MS of C4oxG4 (m/z 663.5) and G4 (m/z 665.5) (a) generated in H_2O before reduction. Negative ion mode CID-MS² fragmentation patterns of C4oxG4 (m/z 663.5, b) and G4 (m/z 665.5, c). These patterns were comparable as published previously (5). The oxygen atom from molecular oxygen is indicated in red. Annotation of fragments is according to the nomenclature developed by Domon and Costello (13).

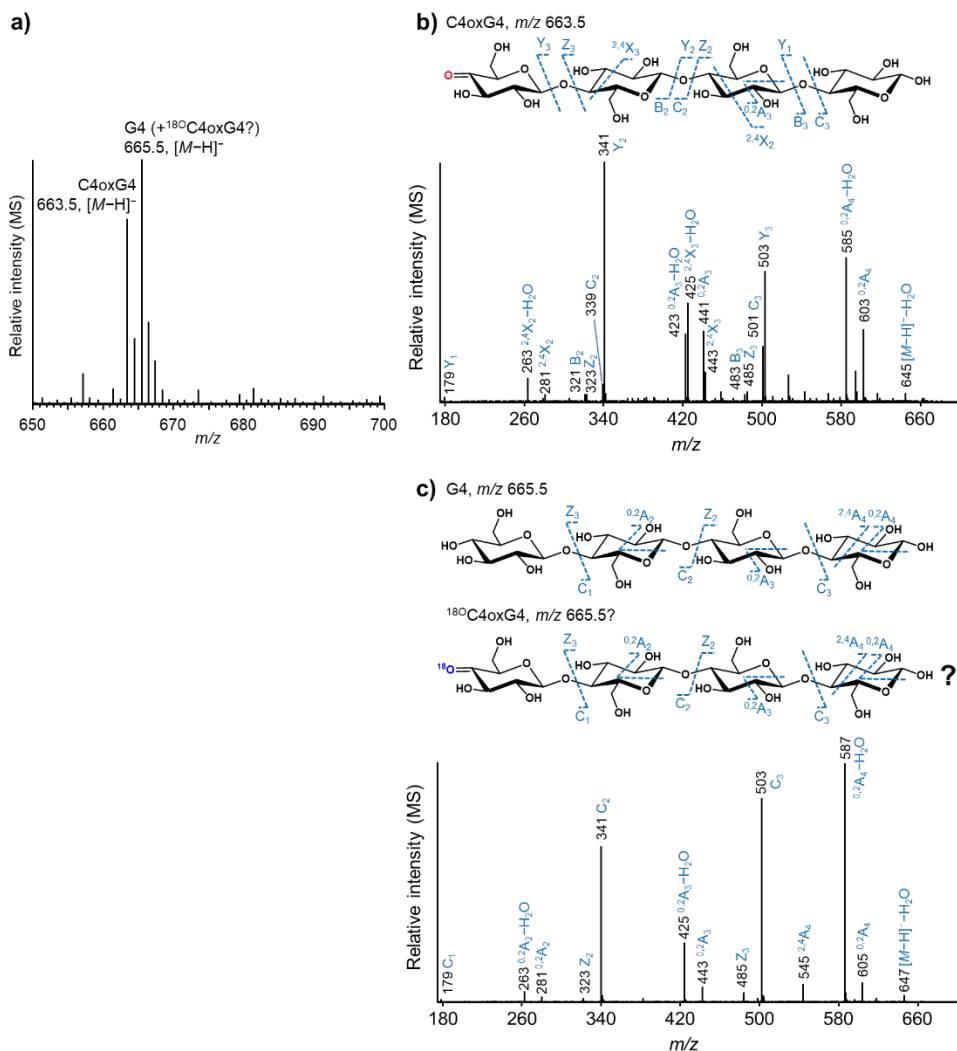


Fig. S10. HILIC-ESI-MS of C4oxG4 (m/z 663.5) and G4 (m/z 665.5) (a) generated in H_2^{18}O before reduction. ^{18}O C4oxG4 could also be present due to the increased relative intensity of m/z 665.5 that is the same as G4. Negative ion mode CID-MS² fragmentation patterns of C4oxG4 (m/z 663.5, b) and G4 (m/z 665.5, c). These patterns were comparable as published previously (5). The oxygen atom from molecular oxygen or from H_2^{18}O is indicated in red and blue, respectively. Annotation of fragments is according to the nomenclature developed by Domon and Costello (13).

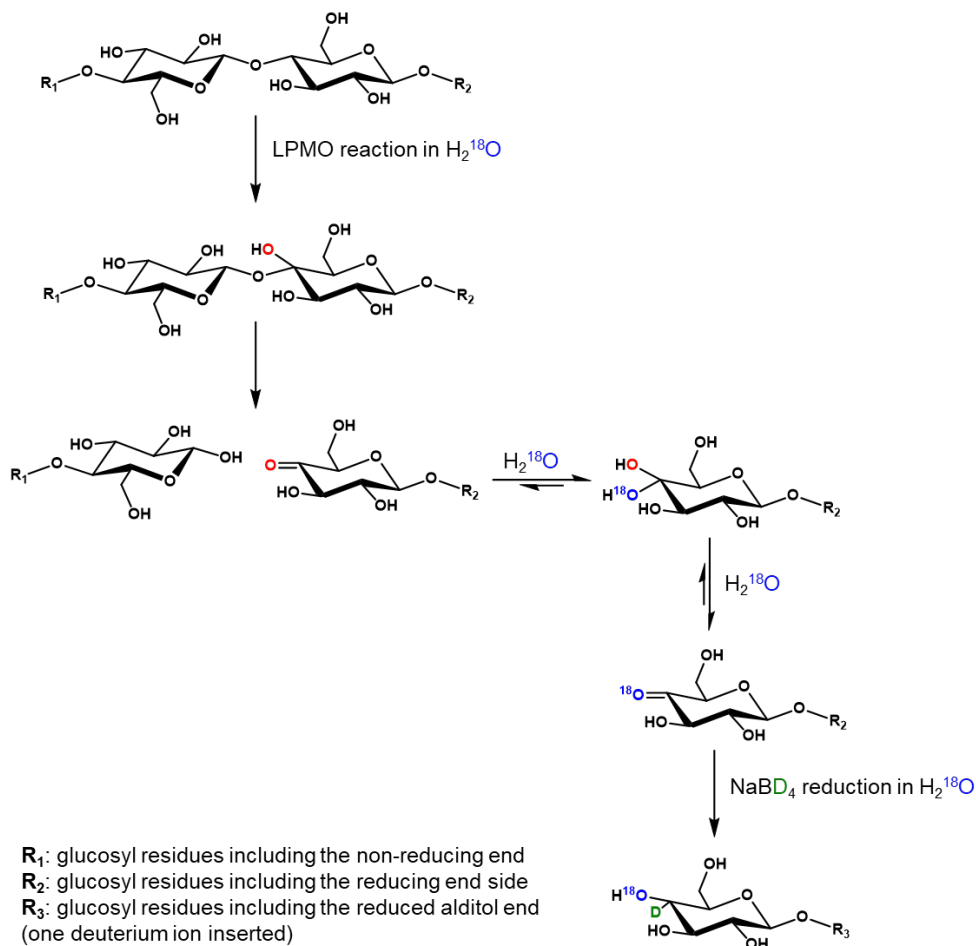


Fig. S11. Proposed route of formation of RD- ^{18}O C4ox. LPMO catalyzes the hydroxylation at the C4 carbon position, which destabilizes the glycosidic bond, leading to the bond cleavage and generation of C4 oxidized cello-oligosaccharides in water (C4 ketone). Here, the oxygen atom (in red) is from dioxygen (O_2) even the reaction is in H_2^{18}O . Due to the equilibrium between C4 ketones and C4 gem-diols, C4 ketones with ^{16}O (red) convert to C4 gem-diol with one ^{16}O and one ^{18}O (blue) in H_2^{18}O . This C4 gem-diol further converts back to C4 ketone but with only ^{18}O at the C4 carbon position. The NaBD_4 reduction in H_2^{18}O initially consumes ^{16}O -C4 ketones and ^{18}O -C4 ketones. However, only ^{18}O -C4 ketones are generated, leading to the presence of predominant of RD- ^{18}O C4ox. Note that RD- ^{18}O C4ox has either glucosyl or galactosyl ends. To simplify the scheme, only one structure is used here.

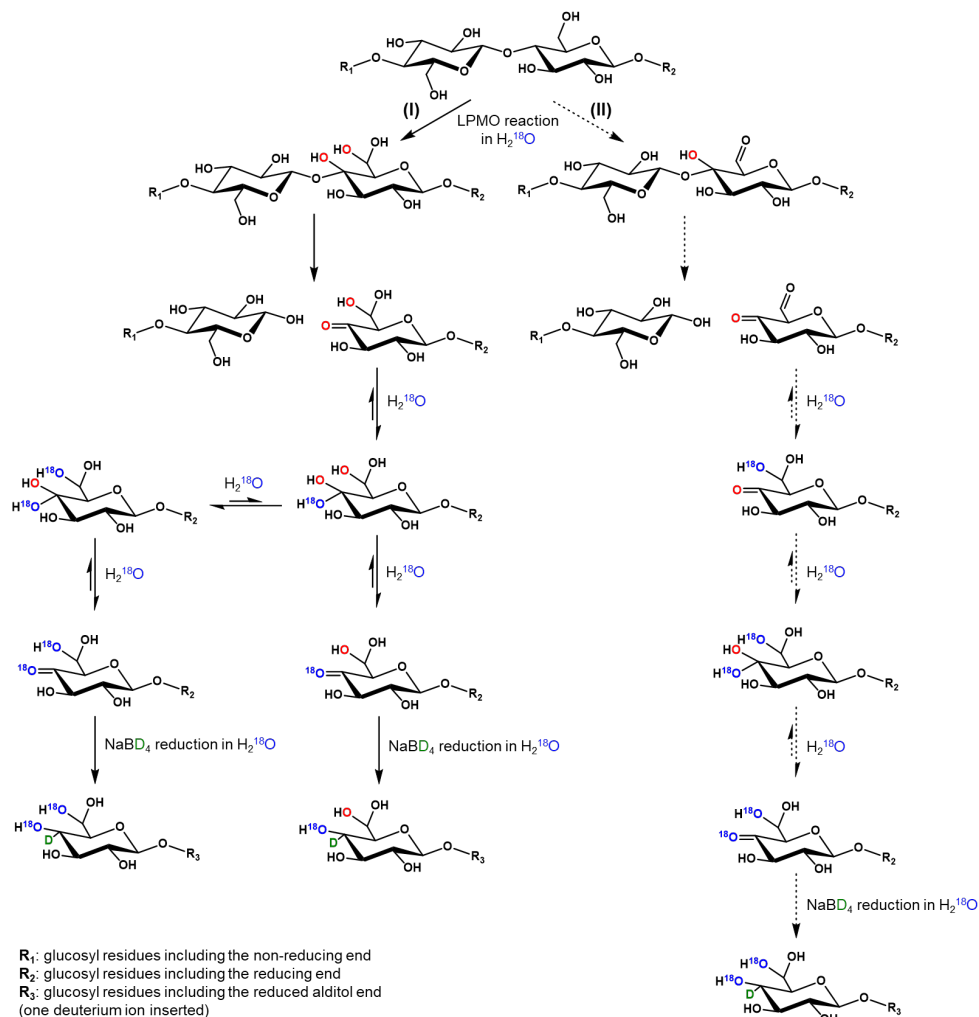


Fig. S12. Proposed routes of formation of RD-¹⁸O₄C₆ox and RD-¹⁸O₄C¹⁸O₆ox. LPMO catalyzes the hydroxylation at the C4 carbon position, which destabilizes the glycosidic bond, leading to the bond cleavage and generation of C4 oxidized cello-oligosaccharides in water (C4 ketone). Here, the oxygen atom (in red) is from dioxygen (O₂) even the reaction is in H₂¹⁸O. In route I, C6 gem-diol is formed by LPMO reaction with one oxygen atom originating from dioxygen (O₂). The conversion of C4-ketone after NaBD₄ reduction in H₂¹⁸O is described in Fig. S11 caption. C6 gem-diol remains after reduction as it cannot be reduced by NaBD₄ (RD-¹⁸O₄C₆ox). In route II, the formation of C4 ketone is the same as in route (a). The C6 carbon is oxidized by LPMO reaction to C6 aldehyde, which is subsequently hydrated to C6 gem-diol. After reduction, RD-¹⁸O₄C¹⁸O₆ox is formed as described in route I. As RD-¹⁸O₄C¹⁸O₆ox can be derived from RD-¹⁸O₄C₆ox, the two-step formation of C6 gem-diol (route II) needs to be further validated.

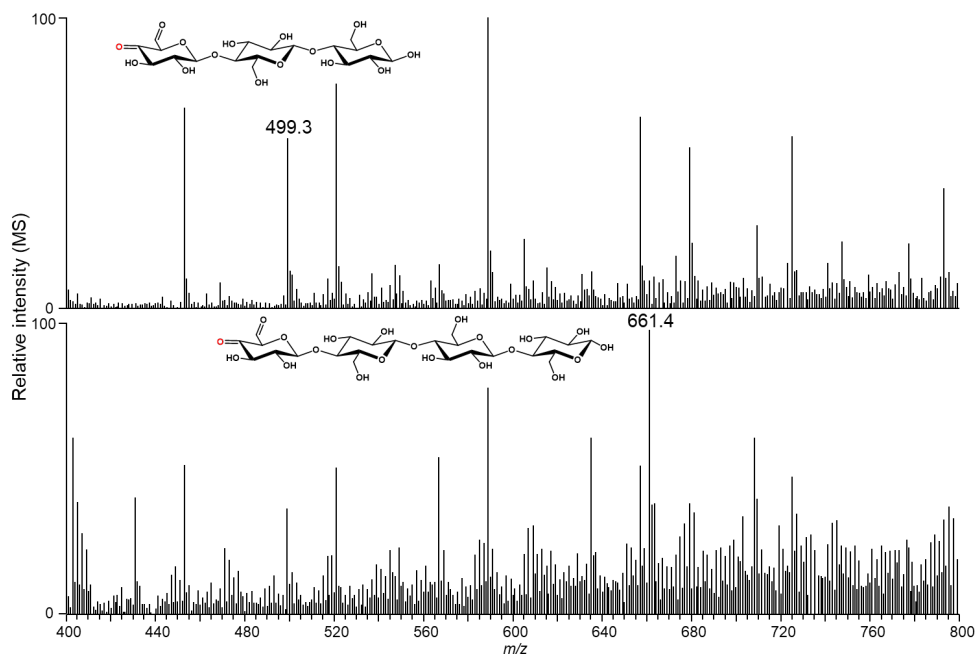


Fig. S13. HILIC-ESI-MS (negative mode; $[M-H]^-$) of possible C4C6oxG3 (m/z 499.3) and C4C6oxG4 (m/z 661.4) having C6 aldehyde structure. No detectable MS signals for m/z values of C4C6oxG2, C4C6oxG5–6 were found. Overall, MS intensity representing C4C6ox having C6 aldehyde structure was too low to confirm or disapprove route II in **Scheme 1**.

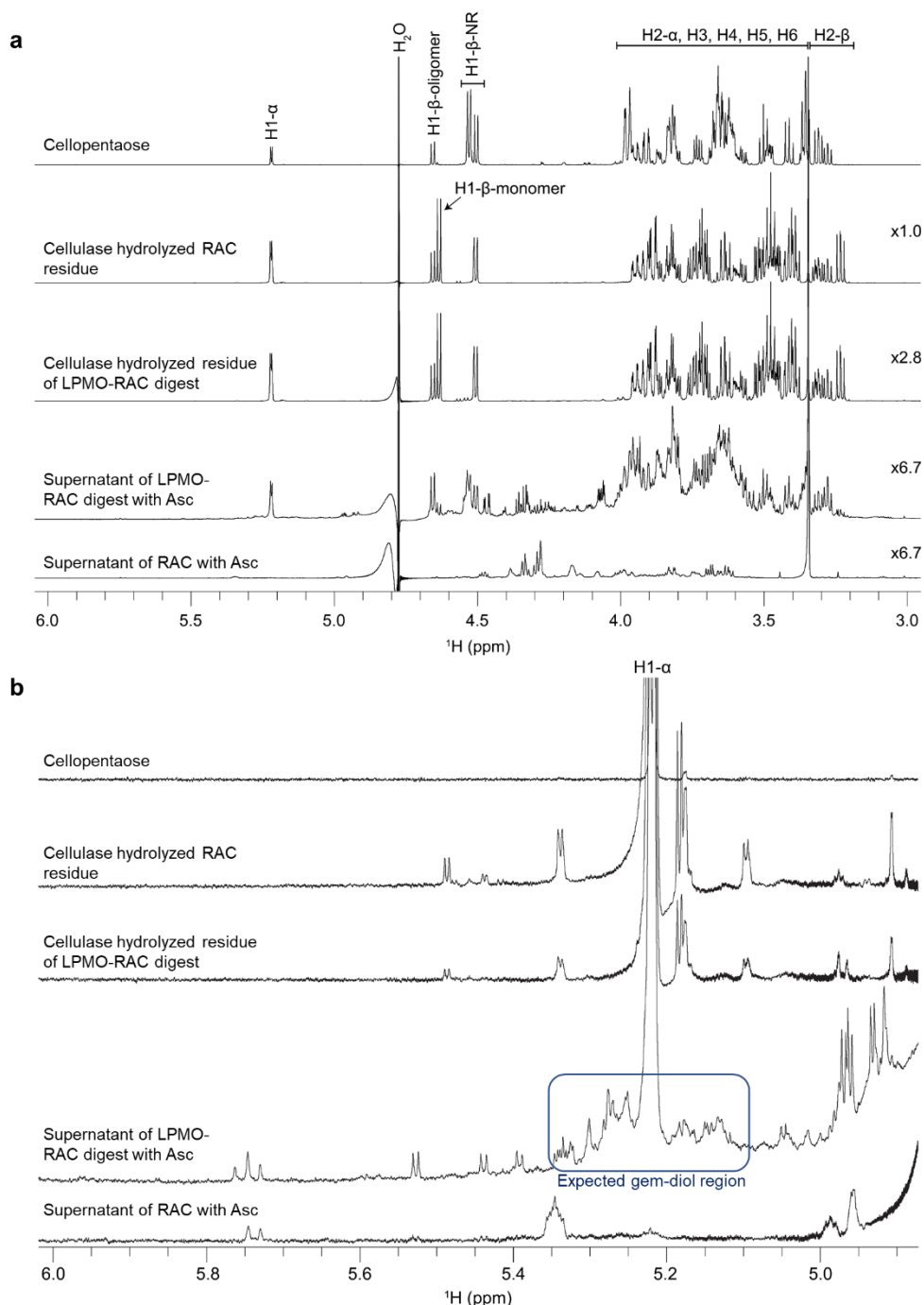


Fig. S14. ^1H NMR spectra (700 MHz, D_2O , water suppressed; for experimental setup see Section 6.1.3.) of LPMO-digested RAC supernatant and residue after cellulase (Celluclast[®] 1.5L) hydrolysis with (a) 3.0–6.0 ppm region and (b) 4.9–6.0 ppm region of a cellopentaose standard, residue of

untreated RAC hydrolyzed by cellulases, residue of LPMO-RAC digest hydrolyzed by cellulases, LPMO-RAC digest supernatant. Both the control samples for cellulase hydrolysis (cellulase sample only) and for the LPMO-RAC digest supernatant (RAC incubated with ascorbic acid (Asc)) showed no interfering peaks in the regions of interest. The spectra in panel (a) were scaled to get an overall comparable intensity among the spectra, the scaling factor is indicated on the right. In panel (b) no scaling factor was applied.

Annotated chemical shifts are in line with literature (14). The presence of a C6 gem-diol has previously been shown to result in a clear downfield shift of the H6 proton as compared to non-oxidized glycosyl units, with H6 chemical shifts reported at 5.27 ppm for the gem-diols of methyl- α -D-mannopyranoside and methyl- α -D-glucopyranoside (15), and 5.11 ppm for the gem-diol of methyl- α -D-galactopyranoside (16). Given the vast amount of different cello-oligosaccharide structures that can include the C6 gem-diol functional groups in our samples, i.e. gem-diols can be present in combination with either C4 ketones or C4 gem-diols, both at various DP's, we expect a range of chemical shifts rather than a single chemical shift. Evidently, several peaks in the expected gem-diol spectral window (panel b) were present in the LPMO-RAC supernatant, which appeared absent in the reaction control samples and absent in both residues after cellulase hydrolysis. We therefore consider it unlikely that specific C6 oxidation occurs but rather seems coupled to LPMO-cleavage of the β -(1 \rightarrow 4)-glucosidic bond.

A signal at 9.36 ppm was detected in the supernatant of the LPMO-RAC digest (not shown), which might correspond to a C6 aldehyde (16, 17), suggesting that C4C6ox having a C6 aldehyde structure might be generated via route II in **Scheme 1**. However, similar as described in Fig. S13, more evidence would be required to prove (or disprove) route II in **Scheme 1**.

6.3. Supplementary references

1. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
2. Sun P, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng*. 2021;9:14124-33.
3. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al., inventors *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
4. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnol*. 2011;7:214-23.
5. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx EJ, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
6. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
7. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
8. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
9. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6:1399-406.
10. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.
11. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SWA, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.

Chapter 4

12. Forsberg Z, Mackenzie AK, Sorlie M, Rohr AK, Helland R, Arvai AS, et al. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2014;111:8446-51.
13. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J*. 1988;5:397-409.
14. Koso T, del Cerro DR, Heikkinen S, Nypelö T, Buffiere J, Perea-Buceta JE, et al. 2D Assignment and quantitative analysis of cellulose and oxidized celluloses using solution-state NMR spectroscopy. *Cellulose*. 2020;27:7929-53.
15. Breton T, Bashiardes G, Léger JM, Kokoh KB. Selective oxidation of unprotected carbohydrates to aldehyde analogues by using TEMPO salts. *Eur J Org Chem*. 2007;2007:1567-70.
16. Parikka K, Tenkanen M. Oxidation of methyl α -D-galactopyranoside by galactose oxidase: Products formed and optimization of reaction conditions for production of aldehyde. *Carbohydr Res*. 2009;344:14-20.
17. Boccia AC, Scavia G, Schizzi I, Conzatti L. Biobased cryogels from enzymatically oxidized starch: Functionalized materials as carriers of active molecules. *Molecules*. 2020;25:2557.

Chapter 5

Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity

Christophe V.F.P. Laurent^{1,2,†}, Peicheng Sun^{3,†}, Stefan Scheiblbrandner¹, Florian Csarman¹, Pietro Cannazza^{1,4}, Matthias Frommhagen³, Willem J.H. van Berkel³, Chris Oostenbrink², Mirjam A. Kabel³ and Roland Ludwig¹

¹Biocatalysis and Biosensing Laboratory, Department of Food Science and Technology, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

²Institute of Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

³Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

⁴Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, Milano, Italy

[†]These authors contributed equally to this work

Abstract

In past years, new lytic polysaccharide monooxygenases (LPMOs) have been discovered as distinct in their substrate specificity. Their unconventional, surface-exposed catalytic sites determine their enzymatic activities, while binding sites govern substrate recognition and regioselectivity. An additional factor influencing activity is the presence or absence of a family 1 carbohydrate binding module (CBM1) connected via a linker to the C-terminus of the LPMO. This study investigates the changes in activity induced by shortening the second active site segment (Seg2) or removing the CBM1 from *Neurospora crassa* LPMO9C. *NcLPMO9C* and generated variants have been tested on regenerated amorphous cellulose (RAC), carboxymethyl cellulose (CMC) and xyloglucan (XG) using activity assays, conversion experiments and surface plasmon resonance spectroscopy. The absence of CBM1 reduced the binding affinity and activity of *NcLPMO9C*, but did not affect its regioselectivity. The linker was found important for the thermal stability of *NcLPMO9C* and the CBM1 is necessary for efficient binding to RAC. Wild-type *NcLPMO9C* exhibited the highest activity and strongest substrate binding. Shortening of Seg2 greatly reduced the activity on RAC and CMC and completely abolished the activity on XG. This demonstrates that Seg2 is indispensable for substrate recognition and the formation of productive enzyme-substrate complexes.

Keywords

Enzyme engineering; Lytic polysaccharide monooxygenase; Phylogenetic analysis; Regioselectivity; Substrate binding; Substrate specificity

1. Introduction

Lignocellulose is the most abundant renewable resource, but because of the many difficulties to process plant biomass, its huge potential cannot be fully utilized yet. The great number of hydrolytic enzymes found in the secretomes of lignocellulolytic microorganisms indicate the difficulties in efficiently degrading the interwoven biopolymers lignin, cellulose and hemicellulose. In recent years it has been found that specialized oxidoreductases support the action of hydrolases on biopolymers. By performing oxidative cleavage of a polysaccharide chain in the presence of an electron donor and an oxygen species, extracellular fungal lytic polysaccharide monooxygenases (LPMOs) increase the substrate accessibility for, and potentiate the action of, other carbohydrate active enzymes (1-8). Since the discovery that LPMOs have an oxidative activity on cellulose, there have been constant additions to the various LPMO families. Until now, the number of organisms from different kingdoms reported to produce LPMOs has multiplied and the number of classes or families into which LPMOs are sorted has increased considerably (1). The classification criteria are either based on the substrate specificity for polysaccharide biopolymers (Enzyme Commission, EC) or on the amino acid sequence, structural features and substrate specificity [Carbohydrate-Active enZyme database (CAZy; www.cazy.org)] (9). The first two LPMO families created in the CAZy database were classified as auxiliary activities (AA) in families AA9 and AA10. AA9 comprises LPMOs which cleave cellulose chains by oxidation of carbons at C1 (EC 1.14.99.54, lytic cellulose monooxygenase (C1-hydroxylating)) or at C4 carbon atom (EC 1.14.99.56, lytic cellulose monooxygenase (C4-dehydrogenating)). The AA10 family comprises LPMOs that act on chitin (EC 1.14.99.53, lytic chitin monooxygenase) and cellulose (both EC 1.14.99.54 and EC 1.14.99.56) (10). Within the past years, new LPMO families have been added: AA11 acting on chitin (11), AA13 acting on starch (EC 1.14.99.55 lytic starch monooxygenase) (12), AA14 acting on xylan and cellulose (13), AA15 acting on cellulose and chitin (14) and AA16 acting on cellulose oxidizing the C1 carbon position (EC 1.14.99.54) (15). The Enzyme Commission has also established a preliminary EC number for lytic xyloglucan monooxygenase (EC 1.14.99.B10).

The catalytic activity of LPMO relies on a mono copper-dependent active site. The copper ion is coordinated by three nitrogen atoms that are part of a histidine-brace composed of two conserved histidine residues (6). The N-terminal His contributes two copper-binding nitrogen atoms, whereas the third copper-binding nitrogen is part of the second His found in a HXGP motif. This conserved structural motif is found in all LPMOs and is located at the base of the immunoglobulin-like β -sandwich fold. In order to initiate the oxidative cleavage of its substrate, the copper center of LPMO needs to undergo a so-called "priming reduction" by either a small molecular reductant or an enzyme like cellobiose dehydrogenase (8). Subsequently, the breakage of the polysaccharide chains by LPMO is accomplished through an H_2O_2 (16) driven mechanism. Despite the ongoing debate of the nature of the co-substrate (O_2 or H_2O_2) (1, 5, 6), it is agreed that LPMOs oxidize glycosidic bonds through a reactive $\text{Cu}^{2+}-\text{O}^{\bullet-}$ species (17). Fast oxidative damage of LPMO has been reported when the enzyme is

not in contact with its substrate (5, 6), which supports the presence of a reactive oxygen species. The disruption of the polysaccharide chain results in an increase of the substrate accessibility and hence potentiates the activity of other carbohydrate active enzymes (1, 3, 6, 8).

LPMOs do not feature a substrate binding pocket, but rather have a flat, slightly grooved binding surface around the exposed active site copper atom (2, 8). However, the exact amino acid residues in this binding surface responsible for substrate recognition and regioselectivity of LPMOs are still unknown. Regioselectivity defines whether LPMOs oxidize glucosyl residues in cellulose at the C1- or C4-position, or both (19). Studies to link sequences and structural features to the regioselectivity of LPMO have been attempted (19-30). So far, although similarities have occurred, such phylogenetic analysis could not unambiguously predict the three LPMO regioselectivities. Therefore, other studies focused on the deletion of amino acids to study regioselectivity. Two independent studies targeted the so-called L2 loop (20) of two different C1/C4-oxidizing LPMOs to render regioselectivity. Whereas in one case a change in the regioselectivity was reported (19), another study resulted in an inactive enzyme variant (27). The latter indicates that mutational studies of the LPMO active site are complicated and difficult to compare. Most likely, sequence and structural differences between LPMOs, i.e., from different organisms, are too large to allow good comparison.

In addition to a catalytic module, around one fifth of fungal LPMOs feature a C-terminal CBM1 connected via a flexible linker of variable length and amino acid composition (2, 31). CBM1 is an important module empowering cellulose binding of many hydrolytic cellulose-degrading enzymes (32). A similar binding role of this CBM1 for LPMOs was demonstrated by isothermal calorimetry (ITC), which measured a forty times lower dissociation constant (K_d) for the complex between the CBM1 of NcLPMO9C and regenerated amorphous cellulose (RAC) than for the complex between the catalytic domain of NcLPMO9C and RAC (24). Interestingly, removal of the CBM1 from the C1/C4-oxidizing *Podospira anserina* LPMO9H, not only negatively affected binding and activity, but also resulted in a shift in regioselectivity towards predominantly C1 oxidation (32). Similarly, other studies reported that truncation of the linker and CBM1 of LPMOs affect both the substrate binding and the formed products (24, 26).

In the current research, the correlation between AA9 LPMO sequences and activity, substrate binding and regioselectivity is studied. Sequence alignments and phylogenetic analyses were used to define structural elements likely to influence substrate binding. Based on this work, NcLPMO9C was subjected to enzyme engineering to remove a segment of the active site or to remove the CBM1. The performance of the generated NcLPMO9C variants was studied by spectrophotometric activity assays, surface plasmon resonance measurements and an HPLC-based product pattern analysis. The results were compared to the corresponding properties of native NcLPMO9C, native NcLPMO9F and native NcLPMO9M, the latter two enzymes having different segments.

2. Materials and methods

2.1. Multiple sequence alignment, phylogenetic analysis and homology modelling

The multiple sequence alignment of fungal AA9 LPMO sequences was calculated with the structure-based MAFFT-DASH algorithm. The alignment and subsequent phylogenetic analysis were based on mature protein sequences selected from the CAZY database with a focus on fungi having multiple LPMOs. Phylogenetic analysis was performed with RaxML-NG (33). ModelTest-NG (32) was used to find the best-fit substitution model. The evolutionary history was inferred by using the Maximum Likelihood method using the Wheelan & Goldman model (35) with frequencies, invariant sites, and the number of gamma distributed sites set to 4. Bootstraps analysis for tree inference was carried out until convergence (cut-off: 0.03; reached after 720 bootstraps) (36). A second tree was established using only the putative surface substrate binding regions of the LPMOs using the same settings and the BLOSUM62 model with frequencies, invariant sites, and the number of gamma distributed sites set to 4. In both cases, the best-scoring most-likelihood tree is shown. The analyzed sequences grouped into three clusters and LPMOs originating from the same fungus were distributed in all three clusters. The phylogenetic trees of fungal LPMOs were linked to already-described substrate types. Homology model of TtLPMO9B (GeneBank ID: AON76800.1) was generated using SWISS-MODEL (35-41) (template: PDB entry 5TKF, sequence identity: 41.86%, QMEAN: -2.03).

2.2. Chemicals and substrates

All chemicals and the substrates carboxymethyl cellulose (CMC) and 2,6-dimethoxyphenol (2,6-DMP) were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless indicated otherwise. Regenerated amorphous cellulose (RAC) was prepared as previously described by Zhang and coworkers (42). Tamarind seed xyloglucan was purchased at Megazyme (Wicklow, Ireland).

2.3. Plasmid design

The cDNA of NcLPMO9M with the native signal peptide (NCU07898, XP_328604.1, EAA33178.1) and the modified NcLPMO9C^{ΔSeg2} with the NcLPMO9C signal peptide were synthesized by BioCat GmbH (Heidelberg, Germany) and inserted into the pPICZ A vector by means of restriction enzymes BstBI and NotI. To generate the NcLPMO9C^{ΔSeg2} variant, nucleotides coding for residues 79–98 were replaced by the ones coding for residues 84–88 of NcLPMO9F (amino acid numbering including signal peptide). To generate the expression plasmids coding for the NcLPMOs 9C^{ΔCBM} and 9C^{ΔSeg2, ΔCBM}, variants with the truncated linker and CBM1, prolines 228 and 213 respectively were mutated into an ochre stop codon using the PCR primers P228X_F and P228X_R (Table. S7). Successful mutations were confirmed by Sanger sequencing at Microsynth Austria GmbH (Vienna, Austria) using the inhouse sequencing primers 5-AOX1 and 3-AOX1.

2.4. Enzyme production and purification

NcLPMO9C and NcLPMO9F were produced by Kittl et al. (43). NcLPMO9M, NcLPMO9C^{Δseg2}, NcLPMO9C^{ΔCBM} and NcLPMO9C^{ΔSeg2, ΔCBM} were recombinantly expressed in *Pichia pastoris* X-33 as previously described (43). The enzyme production was carried out in a 5 L fermenter [Eppendorf (Hamburg, Germany) BioFlo 120 system] with the addition of 5 μM CuSO₄ and the enzyme was subsequently purified by column chromatography as published (43). In brief, ammonium sulfate salt was added to the clear supernatant until a concentration of 30% was reached. Subsequently, the supernatant was loaded onto a 600 mL Phenyl-Sepharose Fast Flow column which was equilibrated with 50 mM sodium acetate, pH 5.0 containing 30% ammonium sulfate. The conductivity was matched to the supernatant by adding sodium chloride. A linear gradient from 30–0% ammonium sulfate over 3 column volumes was used to elute the proteins. Fractions containing LPMO activity were pooled and diafiltered with a Vivaflow crossflow module [MWCO 10000, Millipore (Burlington, MA, USA)]. A 60 mL Q15-Source column was equilibrated with 20 mM Tris/HCl buffer, pH 8.0 and the pooled fractions were loaded. As previously described, the LPMOs eluted in the flow-through. To reduce the volume the flow-through was diafiltrated with the same Vivaflow crossflow module as described above. NcLPMO9C and NcLPMO9C^{ΔCBM} were further purified with a 500 mL Superdex™ 75 column equilibrated with 20 mM potassium phosphate buffer, pH 6.0. Finally, the activity of the purified LPMOs was confirmed by the recently published LPMO activity assay (44). Table S2 summarizes the purification steps and yields. The enzyme purity was verified by SDS-PAGE and peptide mapping as described below (Fig. S7, S9 and Table S8). The SDS-PAGE was prepared as follows. Lanes and molecular weights of the marker proteins (Precision Plus Protein Unstained Standards, Bio-Rad, Hercules, CA, USA) are indicated. The precast gel (4%–20% Mini-PROTEAN® TGX Stain-Free™ Precast Gel, Bio-Rad, , Hercules, CA, USA) was stained with colloidal Coomassie blue [Bio-Safe™ Coomassie G.250 Stain, BioRad (Hercules, CA, USA)] for 1 h and destained overnight.

2.5. Protein identification: peptide mapping analysis

The sample was digested in solution. The protein mixture was reduced with dithiothreitol, S-alkylated with iodoacetamide and digested with Sequencing Grade Modified Trypsin [Promega (Madison, WI, United States)]. The digested samples were loaded on a BioBasic (Markham, ON, Canada) C18 column [BioBasic-18, 150 × 0.32 mm, 5 μm, Thermo Scientific (Waltham, MA, USA)] using 80 mM ammonium formate buffer as the aqueous solvent. A gradient from 96.5% solvent A and 3.5% solvent B (solvent A: 80 mM ammonium formate buffer at pH 3.0, B: 80% acetonitrile and 20% A) to 40% B in 45 min was applied, followed by a 15-min gradient from 40% B to 95% B, at a flow rate of 6 μL min⁻¹. Detection was performed with a QTOF MS (Bruker maXis 4G) equipped with the standard ESI source in positive ion, DDA mode (= switching to MSMS mode for eluting peaks). MS-scans were recorded (range: 150–2200 m/z) and the 6 highest peaks were selected for fragmentation. Instrument calibration was performed using ESI calibration mixture (Agilent, Santa Clara, CA, USA). The analysis files were converted (using Data Analysis, Bruker, Billerica, MA,

USA) to mgf files, which are suitable for performing a MS/MS ion search with Global Proteome Machine [GPM (Rockville, MD, USA); X! Tandem embedded]. The files were searched against a homemade database containing the target sequences.

2.6. LPMO activity assay

The activity and catalytic properties of the LPMOs were determined with the 2,6-DMP activity assay for LPMO (44). The activity of the LPMOs was measured at pH 6.0 in a 100 mM sodium acetate buffer, pH 6.0 in the presence of 2 mM 2,6-DMP. To determine the Michaelis Menten kinetics, H₂O₂ was added to the reactions in the following final concentrations: 1.562, 3.125, 6.25, 12.5, 25, 50, 100 and 500 μ M. The reactions were incubated in triplicates at 30 °C for 10 min before adding LPMO to the reactions. The absorbance at 469 nm was recorded in 1 mL cuvettes for 300 s using a PerkinElmer (Waltham, MA, USA) LAMBDA 35 Spectrophotometer.

2.7. Enzyme incubations

Substrates (RAC, CMC, and XG) were suspended in 50 mM ammonium acetate buffer, pH 5.0 to a concentration of 2 g L⁻¹. Subsequently, enzymes were added to a final concentration of 1.25 μ M. In order to reduce the LPMO copper center, 1 mM ascorbic acid was added to the incubations. Control reactions were performed without the addition of ascorbic acid. All reactions were incubated at 30 °C in a head-over-tail rotator at 20 rpm (time curves; 5 mL total volume) or in an Eppendorf (Hamburg, Germany) Thermomixer comfort at 800 rpm (single 24 h reactions, 200 μ L total volume). An incubation in an Eppendorf (Hamburg, Germany) ThermoMixer® C at 97 °C for 10 min was performed to stop the enzyme reactions. Subsequently, the supernatant was recovered after centrifugation in a Hermle (HERMLE Labortechnik GmbH, Wehingen, Germany) Z 233 MK-2 centrifuge at 22,000 \times g (rotor: 220.87 VO5/6) and stored at -20 °C until further usage.

2.8. Differential scanning calorimetry

Differential scanning calorimetry (DSC) was performed on a MicroCal PEAQ-DSC Automated (Malvern Panalytical, Malvern, Worcestershire, UK) equipped with a 96-well plate autosampler holding at 4 °C. To prevent the solutions from boiling, the experiments were performed under increased pressure (~4.2 bar). At least four buffer runs (i.e. buffer in sample as well as reference cell) were performed at the beginning of an experiment to establish the thermal history of the cells. NcLPMOs (15 μ M in 50 mM potassium phosphate buffer, pH 6.0) were heated from 20 to 90 °C with a temperature ramp of 1 K min⁻¹ in high feedback mode. Since the rescans (identical settings) showed some unfolding events, the baseline correction was performed with the buffer. The data analysis was performed using the MicroCal PEAQ-DSC Software version 1.4 (Malvern Panalytical, Malvern, Worcestershire, UK). Correction for sample baseline was done using the spline function. A non-two-state thermal unfolding model was fitted to the data points after subtraction of the buffer baselines and normalization for the protein concentration.

2.9. Electronic circular dichroism (ECD) spectroscopy

A Chirascan circular dichroism spectrometer (Applied Photophysics, Leatherhead, Surrey, UK) was used to collect ECD spectra. The instrument was flushed with a nitrogen flow of 5 L min⁻¹ throughout all experiments. The samples were analyzed at 30 °C using a 0.1 mm path length quartz cell. Spectra were recorded at a wavelength range of 190–280 nm with scan speed set to 5 s nm⁻¹ at a bandwidth of 1 nm. Thermal denaturation experiments were performed from 40–90°C with a constant heating rate of 1 °C min⁻¹ at a wavelength range of 200–250 nm and a scan speed of 2 s nm⁻¹ at a bandwidth of 1 nm. All measurements were performed in 50 mM phosphate buffer, pH 6.0 with a LPMO concentration of 3 mg mL⁻¹. Data were analyzed with the PRO-DATA SX software version 2.2.17 (Applied Photophysics, Leatherhead, Surrey, UK). Transition midpoint temperatures were calculated from double sigmoidal non-linear curve fits at a wavelength of 229 nm, corresponding to the highest delta ellipticity of the thermal denaturation experiment. Secondary structure was predicted using the BeStSel web server (<http://bestsel.elte.hu>, accessed 25.11.2019) from the far UV wavelength range of 200–250 nm and a scale factor of 1 (45).

2.10. HPAEC analysis for profiling oligosaccharides

Degraded xyloglucan oligosaccharides were analyzed by high performance anion exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) on a ICS5000 (Thermo Scientific, Waltham, MA, USA) system equipped with a CarboPac PA-1 column (2 mm ID × 250 mm) in combination with a CarboPac PA guard column (2 mm ID × 50 mm). The mobile phases used were (A) 0.1 M NaOH, (B) 1 M NaOAc in 0.1 M NaOH and the column temperature was 20 °C. The elution program applied has previously been described (46). Samples were diluted 5-fold prior to analysis. Standard cello-oligosaccharides [DP2–6; Sigma-Aldrich (St. Louis, MO, USA)] were mixed, each in a concentration of 2.5 µg mL⁻¹ and used for calibration.

2.11. HPSEC analysis for molecular weight distribution of (degraded) xyloglucan

Xyloglucan and corresponding digests were analyzed by high performance size exclusion chromatography (HPSEC) for their molecular weight distribution. The analysis was performed on an Ultimate 3000 system (Thermo Scientific, Waltham, MA, USA) coupled to a Shodex RI-101 detector (Showa Denko K.K., Tokyo, Japan). Three TSK-Gel columns (SuperAW4000, SuperAW3000, SuperAW2500; 6 mm × 150 mm per column) in series combined with a TSK Gel super AW guard column (6 mm ID × 40 mm) (Tosoh Bioscience, Tokyo, Japan) were used with the column temperature of 55 °C. Supernatant (10 µL) of LPMO-xyloglucan digests was injected and eluted with 0.2 M NaNO₃ at a flow rate of 0.6 mL min⁻¹. Pullulans (Associated Polymer Labs Inc., New York, USA) in the range of 0.4–708 kDa were used as calibrants.

2.12. Binding studies on cellulose

Microcrystalline cellulose (MCC) for column chromatography (Merck, Darmstadt, Germany) was dissolved in N,N-dimethylacetamide/LiCl to prepare cellulose films for SPR measurements. LiCl was dried at 200 °C for 24 h and afterwards stored in the desiccator over silica prior to use. Anhydrous N,N-dimethylacetamide (DMAc) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Poly(diallyldimethylammoniumchloride) solution (PDADMAC) with an average molecular weight of 400,000 to 500,000 g mol⁻¹ and a concentration of 20 wt.% dissolved in water was obtained from Sigma-Aldrich (St. Louis, MO, USA). Activation and dissolution of cellulose was performed by the procedure adapted from the procedure described by Raj et al. (47). In brief, 2 g of MCC were swollen in 100 mL HQ-water constantly stirred at 22 °C for 18 h. The swollen cellulose fibers were further incubated in 25 mL methanol for 45 min at 40 °C twice before they were swollen in 25 mL DMAc at 22 °C for 45 min for four times. Between the washing and swelling steps, the fibers were recovered using a paper filter. The same filter was used throughout the different steps to minimize losses. Activated cellulose was dried under vacuum at 60 °C for 24 h and stored over silica prior to further use. Water free LiCl was dissolved to a final concentration of 7 wt.% in anhydrous DMAc at 40 °C. This solvent system was used to dissolve activated cellulose by stirring at 22 °C for 18 h to obtain a clear cellulose solution.

SPR sensors were prepared using the SIA Kit Au (GE Healthcare, Chicago, IL, USA). The method for the preparation of cellulose films on gold surfaces was modified from the procedure described by Szech and Riegler (48). Gold-coated targets were cleaned using alkaline Piranha solution containing NH₄OH/H₂O₂/H₂O in a 1:1:3 ratio at 75 °C for 15 min and rinsed with HQ-water and dried at 80 °C. PDADMAC was used as adhesion promoter. The cleaned targets were covered with a solution of 100 mg L⁻¹ PDADMAC and 0.01 M NaCl for 30 min and rinsed with HQ-water. Afterwards, the targets were dried at 80 °C and stored over silica prior to their modification with cellulose. The cellulose film was deposited on the gold surface by spin-coating. For that, 80 µL of 0.5 wt.% cellulose in DMAc/LiCl were deposited at 3000 rpm for 3 min and the solvent was removed by drying at 160 °C. The cellulose film was extensively washed in HQ-water to remove LiCl and the target was dried at 160 °C. Finally, the sensor chip was assembled according the instructions of the manufacturer.

Surface plasmon resonance spectroscopy was performed with a Biacore T200 system (GE Healthcare, Chicago, IL, USA). All experiments were performed at 30 °C at a flow rate of 10 µL min⁻¹ in Dulbecco's phosphate buffered saline supplemented with 0.05 wt.% Tween 20. Binding kinetics were determined using single-cycle-kinetic measurements with five sequential injections of the analyte (LPMO) with an association time of 60 s and a final dissociation of 180 s without regeneration between the injections of one cycle. Between the cycles, the surface was regenerated using 4 M MgCl₂ for 60 s and 10 mM glycine/HCl (pH 2.5) for 30 s at the same flow rate. For the initial equilibration of the surface three cycles of blank injections with running buffer and the same regeneration procedure were applied. Each measurement was performed

in three independent runs on independent flow channels. Data were evaluated using the BiaEvaluation Software version 3.1 (GE Healthcare, Chicago, IL, USA) and SigmaPlot version 12.0 (Systat Software Inc., San Jose, CA, USA). To study the binding behavior of LPMO in its reduced state to the cellulosic substrate the running buffer, solutions used for regeneration of the surface and sample solutions were degassed applying vacuum, purged with nitrogen and finally with argon immediately before the measurement to remove oxygen. The running buffer and the sample solutions were supplemented with 10 mM gallic acid to reduce the active site copper.

3. Results

3.1. Clustering of AA9 LPMOs is based on extended catalytic site segments

A multiple sequence alignment of 101 putative fungal AA9 LPMO amino acid sequences was compiled with the structure-based MAFFT-DASH algorithm (Fig. S1). Subsequent phylogenetic analysis was performed with RaxML-NG using the maximum likelihood algorithm (Fig. S2). The phylogenetic analysis was based on mature protein sequences selected from the CAZy database with a focus on fungi having putative multiple LPMOs. The analyzed sequences grouped into three clusters and LPMOs originating from the same fungus distributed over all three clusters. For example, *NcLPMO9F* is found in Cluster 1, *NcLPMO9C* in Cluster 2, and *NcLPMO9M* in Cluster 3. The main difference between the three clusters was the length of the two different regions located within the first third of the sequence alignment. Similar observations have been made previously and based on structural data these regions, together with two additional ones, were assigned to loops (L2, L3, LS and LC) close to LPMO's catalytic site (19, 20, 24, 27). However, since these regions sometimes include secondary structure elements such as α -helices or β -strands, we prefer the term segment instead of loop. Five sequence segments (Seg1–Seg5) were defined based on the alignment and are indicated in various LPMO structures selected from the three clusters (**Fig. 1**, Fig. S3 and Table S1). Although the exact designation differs, because our definition is sequence-based, Seg1–3 and Seg5 are comparable to the previously defined L2, L3, LS and LC regions, respectively. Seg4 is short and has not been mentioned in previous literature. LPMOs found in Cluster 1 have a short Seg1 and Seg2 exemplified by *NcLPMO9F* (**Fig. 2a**). LPMOs in Cluster 2 have an extended Seg2 (*NcLPMO9C*, **Fig. 2c**), whereas LPMOs in Cluster 3 have an extended Seg1 (*NcLPMO9M*, **Fig. 2b**). More examples of segments of LPMOs with a resolved structure are shown in Fig. S4.

3.2. Analysis of active site segments hints toward regioselectivity

By analyzing the sequence alignment and the resulting phylogenetic tree we observed that the LPMO clustering roughly reflects the length and composition of the segments. To remove the influence of other positions, a sequence alignment of only the sequences defined as segments was generated (Fig. S5) and a phylogenetic tree inferred. A comparison of the phylogenetic trees obtained from the alignment of full-length sequences and the "segments only" alignment highlights small differences between

these two approaches (Fig. S2). One difference is the regrouping of a small branch of LPMO sequences containing an elongated Seg2 and a cysteine in Seg1 and Seg3 from Custer 1 to Cluster 2.

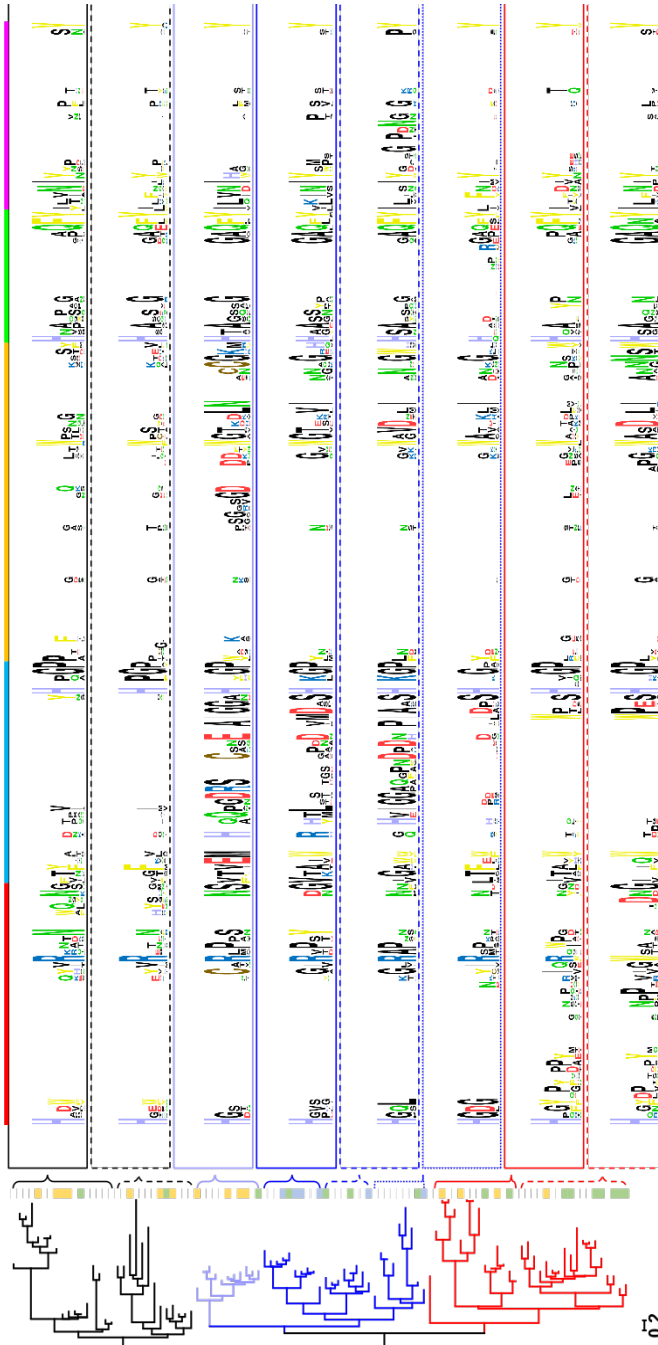


Fig. 1. Phylogenetic tree showing the clustering of AA9 lytic polysaccharide monooxygenases (LPMOs) (Cluster 1, black; Cluster 2, blue; Cluster 3, red) and sequence logos of distinct clades within these clusters for the segment sequences. The colored boxes beside the clades indicate the regioselectivity of characterized LPMOs according to Frommhaugen et al. (49): C1, yellow; C4, blue; C1/C4, green. The colored boxes above the sequence logos indicate the extent of each segment: Seg1, red; Seg2, blue; Seg3, orange; Seg4, green; Seg5, purple, and correlate with the color code used in Fig. 2.

This shows that the sequences in the segments affect the clustering in the phylogenetic analysis the most although covering only 33–41% of the total sequence (not including linker or CBM when present), whereas the rest of the positions found in the core and substrate averted surfaces contributes little to the final result. An overview of the amino acids involved in the five segments is given in Fig. S5. The clusters are relatively well-supported by high bootstrap values, but the phylogenetic distance between the three clusters is relatively small in comparison to the distances between the clades in a cluster. This points to an early acquisition of the differentiating positions (mostly the segments) and an extended evolution of LPMOs within the clusters without changing the segment length. However, the amino acid composition within the segments varies considerably as is obvious from the relatively few conserved positions in the sequence logos in **Fig. 1**. Additionally, most of the conserved positions are at the start and end of the segments and few are found in the middle. The inspection of twelve fungal genomes (*Arthrobotrys oligospora*, *Fusarium fujikuroi*, *F. graminearum*, *F. venenatum*, *Leptosphaeria maculans*, *Neurospora crassa*, *Podospora comata*, *P. anserina*, *Pyricularia oryzae*, *Serendipita indica*, *Thermothelomyces thermophilus*, *Thermothielavioides terrestris*) containing 10 or more *lpmo* genes showed that every fungus in this group encodes LPMOs from all three clusters. This supports the important role of the segments as a basis for the evolution of LPMOs.

When adding the published regioselectivities of the used LPMOs (49) no straightforward correlation of the preferred cleavage position and the clustering was obtained, although certain trends can be found for these three clusters. In Cluster 3, featuring an elongated Seg1, C1/C4-oxidizing LPMOs are predominant (9) together with a small subset of C1-oxidizing LPMOs (4). In Cluster 1, with neither Seg1 or Seg2 elongated, both C1- (7) and C1/C4-oxidizing (2) LPMOs are present, however, here the number of C1-oxidizing LPMOs is higher. In Cluster 2, featuring an elongated Seg2, C4-oxidizing LPMOs are dominant (7) with a small number of C1/C4-oxidizing LPMOs (3). Certainly, the small number of sequences with reported data on the regioselectivity and substrate specificity reduce the predictive power of such an analysis, but the distorted distribution supports the hypothesis that the evolution of sequences within the clusters went a long way after the elongation of the segments and led to a wide diversification of the regioselectivity. The even more interesting question of how the segments and their evolution influences the substrate specificity can, unfortunately, not be answered at present, due to a lack of published substrate specificities for characterized LPMOs.

An interesting observation in regard to Cluster 2 is that in some sequences a cysteine is present in Seg2. LPMOs with this cysteine residue have a predominantly C1-oxidizing regioselectivity and would in this regard better fit to Cluster 1 than to Cluster 2. Homology models suggest that this cysteine can form a disulfide bond with another cysteine in Seg3 (Fig. S6). When this bond is present it pulls the elongated Seg2 away from the catalytic site and mimics a short Seg2 that is found in Cluster 1, which contains predominantly C1-oxidizing LPMOs. This can be viewed as an example of evolutionary adaption within a cluster.

To test the cluster-function relationship (e.g., a potential relationship between Seg2 and a C4 regioselectivity) and the role of the CBM1, we generated a variant of NcLPMO9C found in Cluster 2 and replaced the sequence of Seg2 with that of Seg2 from NcLPMO9F. With that, the LPMO^{ΔSeg2} variant was created which is more related to LPMOs from Cluster 1. In addition, we also removed the C-terminal linker and CBM1 from the wild-type NcLPMO9C and from the NcLPMO^{ΔSeg2} variant resulting in NcLPMO9C^{ΔCBM} and NcLPMO9C^{ΔSeg2, ΔCBM}, respectively. In detail, NcLPMO9C^{ΔSeg2} was generated by replacing the residues 63–82 of NcLPMO9C with residues 67–71 of NcLPMO9F. Rather than deleting the indicated residues from NcLPMO9C we opted for a replacement to avoid a structural strain on the histidine brace forming the active site. Furthermore, Seg2 of NcLPMO9C contains a short β-strand which is not present

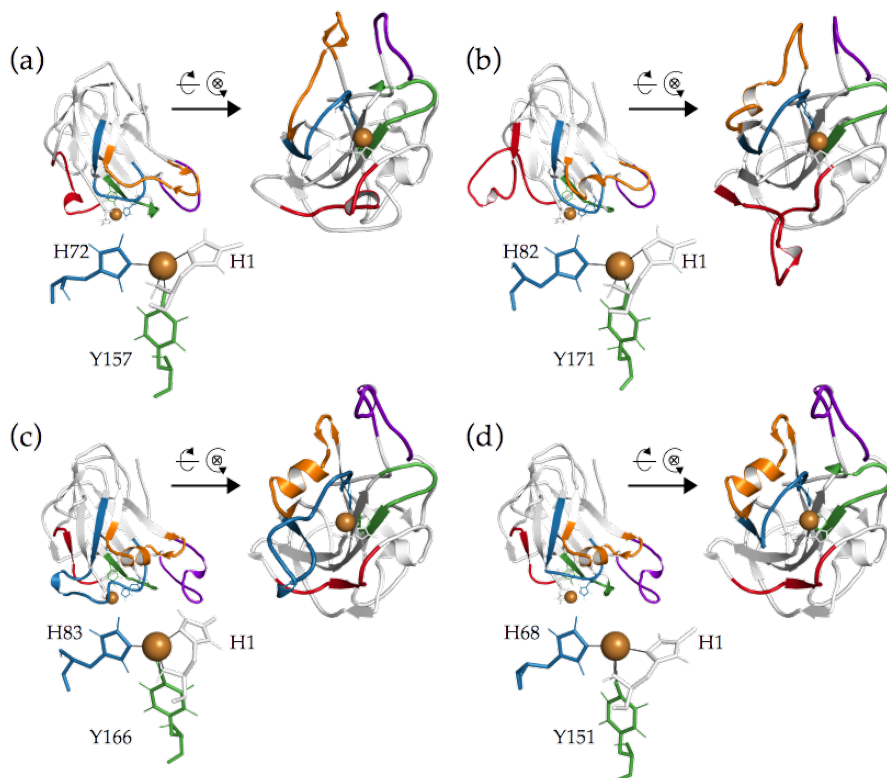


Fig. 2. Cartoon representation of (a) NcLPMO9F (PDB ID: 4QI8), (b) NcLPMO9M (PDB ID: 4EIS), (c) NcLPMO9C (PDB ID: 4D7U) and (d) NcLPMO9C^{ΔSeg2}. The copper atom in the active site is shown as a brown sphere and its coordinating residues are shown in stick representation. Segments are colored: Seg1, red; Seg2, blue; Seg3, orange; Seg4, green; Seg5, purple. The residues within these segments are listed in Table S1. The 3D models have been structurally aligned onto each other and the secondary structure elements were determined using the "Define Secondary Structure of Proteins" (DSSP) algorithm (50) as implemented in GROMOS++ (51). The model is shown from two directions, the arrow indicates the LPMO model that has been rotated by 90° to show the active site region from the front.

in *NcLPMO9F* (Fig. S3). A secondary structure analysis revealed that this short β -strand should not be present in *NcLPMO9C* ^{Δ Seg2}. Additionally, we truncated the linker including the CBM1 from both *NcLPMO9C* and *NcLPMO9C* ^{Δ Seg2} by inserting a stop codon at positions P228 and P213, respectively, which resulted in variants *NcLPMO9C* ^{Δ CBM} and *NcLPMO9C* ^{Δ Seg2, Δ CBM}. Additionally, *NcLPMO9C*, *NcLPMO9F* and *NcLPMO9M* were produced for reference experiments.

The activity of the recombinantly expressed and chromatographically purified *NcLPMOs* and *NcLPMO9C* variants (Fig. S7 and Table S2) was tested with a spectrophotometric assay measuring LPMO activity with 2,6-dimethoxyphenol (2,6-DMP) and H_2O_2 (44). Steady-state kinetic constants of LPMOs and LPMO variants for H_2O_2 were determined at pH 6.0 in a sodium acetate buffer (Table 1 and Fig. S8). Based on this assay, all enzymes and variants were active, which indicates that the catalytic site is functional in all produced LPMOs. The measured specific activities of *NcLPMO9C* and its variants diverge by a factor of three indicating minor effects of the introduced mutations on the catalytic site. The shortened Seg2 in *NcLPMO9C* ^{Δ Seg2} and *NcLPMO9C* ^{Δ Seg2, Δ CBM} lowers the catalytic efficiency by a factor of two. Removal of the CBM in *NcLPMO9C* ^{Δ CBM} and *NcLPMO9C* ^{Δ Seg2, Δ CBM} leads to only small deviations in the catalytic efficiency from the corresponding wild-type and variant enzymes. *NcLPMO9M* also showed activity with kinetic constants for H_2O_2 similar to *NcLPMO9C*. *NcLPMO9F* was previously shown to have a specific activity of $2.2 \pm 0.2 \text{ U g}^{-1}$ in a 100 mM sodium succinate/phosphate buffer at pH 7.5 (44). The amino acid sequence of all produced LPMOs and LPMO variants was confirmed by liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS, Fig. S9).

Table 1. Kinetic constants with standard error of the mean of *NcLPMOs* and variants for H_2O_2 using 2 mM 2,6-DMP as substrate and reductant in 100 mM sodium acetate buffer, pH 6.0. ($n = 3$).

Enzyme	K_M [μM]	Spec. act. with 100 $\mu\text{M H}_2\text{O}_2$ [U g^{-1}]	kcat [s^{-1}]	kcat / K_M [$\mu\text{M}^{-1} \text{s}^{-1}$]
<i>NcLPMO9C</i>	2.8 ± 0.1	28.8 ± 0.2	0.57 ± 0.03	0.20 ± 0.02
<i>NcLPMO9C</i> ^{ΔCBM}	1.4 ± 0.1	13.0 ± 0.1	0.39 ± 0.01	0.28 ± 0.03
<i>NcLPMO9C</i> ^{ΔSeg2}	5.4 ± 0.1	42.4 ± 0.2	0.55 ± 0.01	0.10 ± 0.01
<i>NcLPMO9C</i> ^{ΔSeg2, ΔCBM}	4.0 ± 0.1	41.1 ± 0.3	0.37 ± 0.01	0.09 ± 0.01
<i>NcLPMO9M</i>	7.0 ± 0.3	39.1 ± 0.3	0.40 ± 0.01	0.06 ± 0.01

3.3. Thermal stability of purified LPMOs

The thermal stability of *NcLPMOs* and variants was measured to ensure their proper functioning during further assays. Differential scanning calorimetry (DSC) was used to observe endothermic transitions. *NcLPMO9C* and *NcLPMO9M* showed sharp transitions and a similar transition midpoint temperature (T_m) of 62.3 and 63.1 °C, respectively (Fig. 3, Fig. S10 and S11). In contrast, the thermogram of *NcLPMO9F* revealed a much broader transition with a shoulder at 57.9 °C (T_{m1}) and a maximum at 65.8 °C (T_{m2}) (Fig. S12). The T_m of *NcLPMO9C* and T_{m2} of *NcLPMO9F* were close to previously

reported values (43). The highest T_m was found for NcLPMO9M which has an additional third disulfide bridge (Fig. S3).

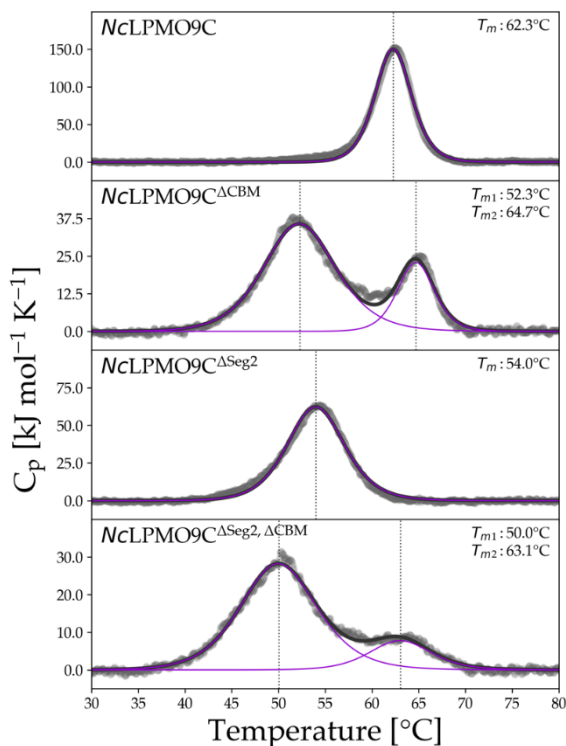


Fig. 3. DSC thermograms of NcLPMO and variants. The NcLPMO9C variants lacking linker and a family 1 carbohydrate binding module (CBM1) are designated by Δ CBM and variants with a shortened Seg2 and indicated by Δ Seg2. Solid black lines show the fitted curves to the raw data (grey scatter plots). The transition midpoint temperature (T_m) is indicated by a dotted line.

As a result of shortening Seg2 in NcLPMO9C Δ Seg2, a decreased T_m was observed and a broader transition compared to the wild-type, indicating a destabilization of the protein fold (**Fig. 3** and Fig. S13). The removal of the linker and CBM1 in NcLPMO9C Δ CBM and NcLPMO9C Δ Seg2, Δ CBM resulted in two partially separated, broadened endothermic transitions (Fig. S14 and S15). Most of the unfolding occurred in the T_{m1} region, representing 75% and 82% of the total enthalpy, respectively. Whereas the T_{m2} of both truncated variants was close to the T_m of the wild-type, destabilization of the protein fold by the shortening of Seg2 and removal of the CBM1 was evident. The T_m or T_{m1} of NcLPMO9C Δ Seg2, NcLPMO9C Δ CBM, and NcLPMO9C Δ Seg2, Δ CBM decreased in relation to the wild-type by 8.3, 10, and 12.3 °C, respectively.

An explanation for the strong destabilizing effect of the removal of the linker and CBM1 on the thermal stability of the resulting variants is that the termination site of the linker is directly after a cysteine. Subsequent linker residues might be involved in the correct folding of the protein and the formation of the C-terminal disulfide bridge in NcLPMO9C (Fig. S3). The truncation directly after this cysteine could hinder the formation of this disulfide bridge. Support for this hypothesis comes from the presence of T_{m2} in NcLPMO9C Δ CBM and NcLPMO9C Δ Seg2, Δ CBM, which likely represents the smaller

fraction of LPMO with a disulfide bridge. To proof this hypothesis, we incubated LPMOs with 5 mM tris(2-carboxyethyl)phosphine (TCEP) before a DSC scan. Unfortunately, the addition of TCEP reduced all disulfide bonds in the LPMOs, which resulted in thermograms that barely differed from the buffer signal, but demonstrate the important role of disulfide bonds to maintain the native protein conformation. Circular dichroism spectroscopy was used to verify DSC experiments and obtain insights into the changes of secondary structure elements during unfolding (**Fig. 4**). The change in ellipticity at 229 nm indicates also one T_m at 65.6 °C for *NcLPMO9C* and two T_m values for *NcLPMO9C*^{ΔCBM} and thereby support the DSC experiments. The temperature-induced change in ellipticity between 40–90 °C (Fig. S16) shows that the unfolding of secondary structure elements begins approximately 5 °C below the T_m for both *NcLPMO9C* and *NcLPMO9C*^{ΔCBM}. Antiparallel β -sheets are the main secondary structure element found in LPMO. The extent of unfolding of the antiparallel β -sheets is similar for both enzymes indicating the same secondary structure composition, but starts already at a lower temperature for *NcLPMO9C*^{ΔCBM} (~50 °C) than for *NcLPMO9C* (~60 °C). A possible explanation is again the lower stability of the *NcLPMO9C*^{ΔCBM} fraction that has no intact, second disulfide bond in the catalytic domain. Highly interesting is the partial refolding of LPMO, which we observed in DSC rescans (Fig. S10–S15). The ratio of ΔH_{cal}^{rescan} over ΔH_{cal}^{scan} was very low for *NcLPMO9M* (0.07), but moderate for *NcLPMO9F* (0.38) indicating that the unfolding process was not fully reversible (Fig. S17). However, *NcLPMO9C* had a ratio of 0.78 which is close to the threshold classifying reversible unfolding (52). The removal of Seg2 and the CBM1 decreased this ratio (again additively), indicating their stabilizing role for the protein fold.

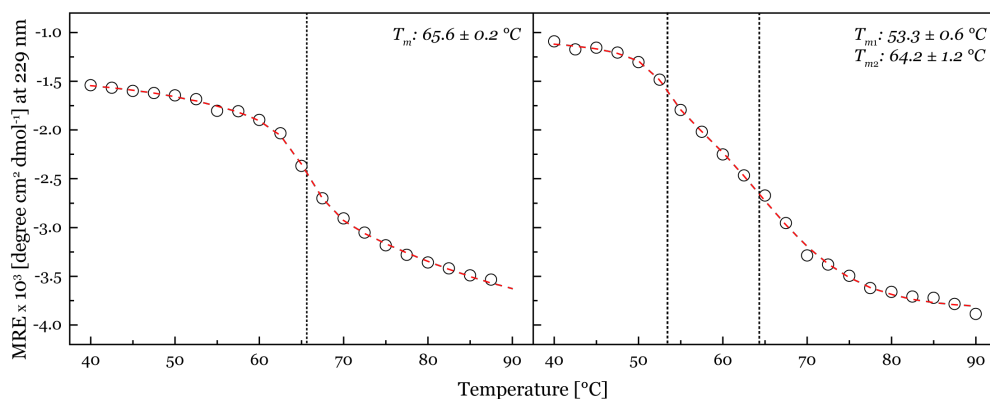


Fig. 4. Thermal unfolding followed by circular dichroism spectroscopy. A double sigmoidal curve was fitted to the changes of ellipticity at 229 nm to calculate the transition midpoint temperatures of 3 mg mL⁻¹ *NcLPMO9C* (left panel) and 3 mg mL⁻¹ *NcLPMO9C*^{ΔCBM} (right panel). (Mean residue ellipticity, MRE).

3.4. LPMO activity on cellulosic substrates

To test the enzymatic activity of the produced LPMOs and their variants on cellulosic substrates, all were incubated with RAC for 24 h at 30 °C in a constantly rotating orbital shaker. In agreement with their classifications, the use of high-performance

anion exchange chromatography (HPAEC) showed that *NcLPMO9C* released C4-oxidized products and *NcLPMO9F* released C1-oxidized products, next to non-oxidized products (Fig. S18 and S19). RAC incubated with *NcLPMO9M* led to the release of the previously described C1- and C4-oxidized and non-oxidized products. Oxidized oligosaccharides were identified based on elution time of previously developed methods (53-55). The same regioselectivity was observed when the soluble substrate carboxymethyl cellulose was incubated with the LPMOs (CMC, Fig. S20 and S21). CMC incubations with *NcLPMO9C*, *NcLPMO9F*, and *NcLPMO9M* resulted in lower amounts of oxidized products, more specifically cello-oligosaccharides (degree of polymerization (DP) ranging from 2 to 6) compared to corresponding RAC incubations. (**Fig. 5**, Fig. S22, Table S4 and S5). Notably, *NcLPMO9M* and *NcLPMO9F* released considerably less amounts of non-oxidized and oxidized products from CMC compared to *NcLPMO9C*.

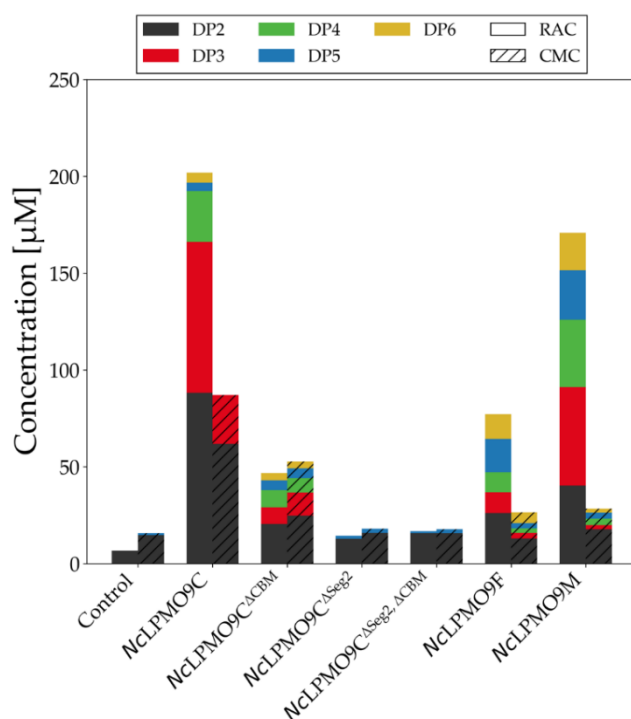


Fig. 5. Quantification of released soluble cello-oligosaccharides from the incubation of 2 mg mL⁻¹ RAC or carboxymethyl cellulose (CMC) with 1.25 μ M *NcLPMOs* and variants in the presence of 1 mM ascorbic acid after 24 h. The quantification of soluble cello-oligosaccharides and the calculation is defined in Table S3. To exclude the impact of hydrolytic background activity, soluble cello-oligosaccharides that were released from the incubation of RAC or CMC with *NcLPMOs* in the absence of ascorbic acid were subtracted from the above-mentioned quantified values (Fig. S22). Buffer was added to the control reactions instead of LPMO. DP: degree of polymerization.

The removal of the CBM1 in *NcLPMO9C* Δ CBM resulted in a decreased activity towards RAC compared to the wild-type (**Fig. 5**). The total release of soluble non-oxidized products by *NcLPMO9C* Δ CBM was 23% for RAC and 60% for CMC compared to *NcLPMO9C* (set as 100%). This decrease in the release of non-oxidized (**Fig. 5**) and oxidized (Fig. S18 and S20) oligosaccharides shows that the CBM1 contributes to the substrate recognition and binding as prerequisite of catalysis. This is especially true for RAC. Interestingly, the preferential conversion of RAC over CMC by *NcLPMO9C* was not found for *NcLPMO9C* Δ CBM. Based on the amounts of the released products one could argue that the removal of the CBM1 no longer allows the enzyme to distinguish between RAC

and CMC. Unfortunately, influence of the modified Seg2 on the LPMO substrate specificity and regioselectivity could not be evaluated due to the low activity of *NcLPMO9C*^{ΔSeg2} and *NcLPMO9C*^{ΔSeg2, ΔCBM}. Nevertheless, the low activity of *NcLPMO9C*^{ΔSeg2} and *NcLPMO9C*^{ΔSeg2, ΔCBM} indicates that Seg2 is a crucial component for substrate recognition in this Cluster 2 LPMO.

3.5. LPMO activity on xyloglucan

Since *NcLPMO9C* is active on tamarind seed xyloglucan (XG) (56), we were interested in how the *NcLPMO9C* variants performed on this soluble hemicellulose under quasi-homogeneous conditions. A release of non-oxidized and oxidized XG oligosaccharides from XG in presence of reduced *NcLPMO9C* was observed (Fig. S23 and S24). Similar to the results with RAC and CMC, the amount of non-oxidized and oxidized XG-oligomers released by *NcLPMO9C*^{ΔCBM} was reduced compared to its wild-type. A qualitative difference of oxidized XG-oligomers was observed for peaks with an elution time between 25–45 min between *NcLPMO9C*^{ΔCBM} and *NcLPMO9C*. No substantial differences of released XG-oligomers of the control and *NcLPMO9C*^{ΔSeg2} or *NcLPMO9C*^{ΔSeg2, ΔCBM} were found, showing that the Seg2 variants lost the capacity of oxidatively degrading XG. The molecular weight distribution (hydrodynamic volume) of XG incubated with the LPMOs was analyzed over time by using high-performance size exclusion chromatography using a refractive index detector (HPSEC-RI). In the absence or presence of ascorbic acid, XG incubated without LPMOs had a molecular weight distribution in the range between 20 kDa to above 700 kDa (Fig. S25 and S26). With the incubation of XG with *NcLPMO9C* or *NcLPMO9C*^{ΔCBM} in the presence of ascorbic acid, the hydrodynamic volume of XG decreased dramatically over time leading to the formation of XG-oligomers with a molecular weight distribution between 20 kDa to 0.4 kDa (Fig. S27–S30).

For a better overview of the product size distribution over time, three elution windows were defined for integration (elution times defined in Table S6). Window 1 at the beginning of the elution covers the biggest XG molecules with a molecular weight between 21.2 and >1000 kDa. Window 2 includes middle-sized XG molecules with a molecular mass range of 7–21.2 kDa and Window 3 represents small-sized reaction products with a molecular mass from 0.35–7 kDa (Fig. S26, S28 and S30). Remarkably, the presence of ascorbic acid had an immediate, molecular-mass decreasing effect in Window 1 (Fig. S25), which was not observed in the absence of ascorbic acid (Fig. S25, S27 and S29).

Fig. 6 (reference experiments in Fig. S31) shows the normalized, integrated intensities of the above-described windows. In general, *NcLPMO9C* and *NcLPMO9C*^{ΔCBM} quickly reduced the amount of large-size XG molecules found in Window 1 and accumulated mid-sized XG molecules in Window 2, which were then further converted to the even smaller, final XG products found in Window 3. *NcLPMO9C* consumed the mid-sized XG molecules fully after 4 h, whereas *NcLPMO9C*^{ΔCBM} stopped after 4 h to convert mid-sized XG molecules into small-sized final products indicating a stop of the conversion process. Fitting areas of Window 3 to an exponential function (Fig. S32), one observes after only 4 h a plateau for *NcLPMO9C*^{ΔCBM} whereas *NcLPMO9C* continues to increase

until 8 h. According to the linear slope between 0–4 h, the production rate of small-sized XG molecules is similar for *NcLPMO9C* and *NcLPMO9C*^{ΔCBM} and shows the CBM1 is not essential for the activity on XG. However, the CBM1 might increase the binding affinity to XG to avoid self-inactivation of LPMO catalytic site, leading to a longer reaction time. Furthermore, the lack of activity of *NcLPMO9C*^{ΔSeg2} and *NcLPMO9C*^{ΔSeg2, ΔCBM} on XG indicates that Seg2 is also important for the substrate recognition of XG.

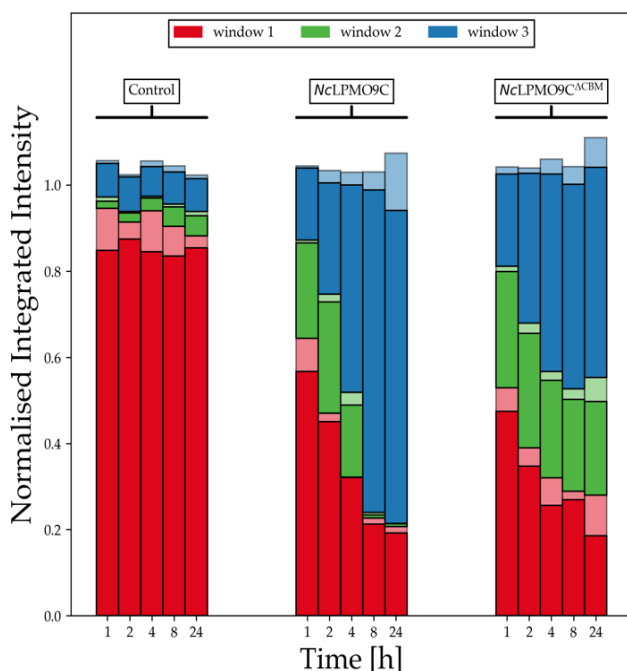


Fig. 6. Normalized integrated intensity of high performance size exclusion chromatography coupled to a refractive index detector (HPSEC-RI) chromatograms over time. Integration windows 1, 2 and 3 were defined as ranging from 7.4–11.0, 11.0–12.0, and 12.00–14.75 min (Table S6). The upper boundaries of the windows were excluded from the integration. Data for two replicates are shown and have been normalized to the sum of the average data points per time point. The opaque bars and the bars with a reduced transparency are indicating the minimum values and maximum values from two independent measurements.

3.6. LPMO binding to cellulose

Conversion experiments clearly showed the correlation between the ability of LPMO to bind to its substrate and the exerted activity. To investigate the effect of the introduced changes on the substrate binding affinity of *NcLPMO9C* and its variants, K_d values were determined by SPR measurements (**Fig. 7**). Microcrystalline cellulose was dissolved in *N,N*-dimethylacetamide and LiCl and spin-coated onto the SPR gold probes to produce a thin layer of cellulose. By measuring the amount of adsorbed LPMO using single-cycle-kinetic measurements with five sequential LPMO injections, K_d values were calculated for *NcLPMO9C* and its variants (**Fig. 7**). Kinetic titration experiments were chosen over multicycle experiments to avoid the modification of the deposited cellulose film by surfactants or high ion concentrations. The K_d determined for the complex between *NcLPMO9C* and the deposited cellulose was $0.94 \pm 0.27 \mu\text{M}$, an about 10-times higher affinity than estimated with *NcLPMO9C*^{ΔSeg2} ($K_d = 11.7 \pm 2.76 \mu\text{M}$). This showed clearly that the binding site of *NcLPMO9C*^{ΔSeg2} has a greatly decreased substrate affinity and it also showed the synergistic effect of the substrate binding site

and the CBM1. The two other studied wild-type *Nc*LPMOs have no CBM1, which allowed us to study the affinity of a single LPMO binding site (although with different active site segments). By applying the SPR methodology described above we determined for the Cluster 1 *Nc*LPMO9F a K_d of $25.6 \pm 6.87 \mu\text{M}$ and for Cluster 3 *Nc*LPMO9M a K_d of $53.3 \pm 7.87 \mu\text{M}$ for the complex with the cellulose substrate (**Fig. 7**). Keeping in mind that the cellulose on the SPR probe is only an approximation of its naturally available form, it still demonstrated that the presence of a CBM1 directs the LPMO to its substrate and decreases the fraction of unbound, diffusible LPMO.

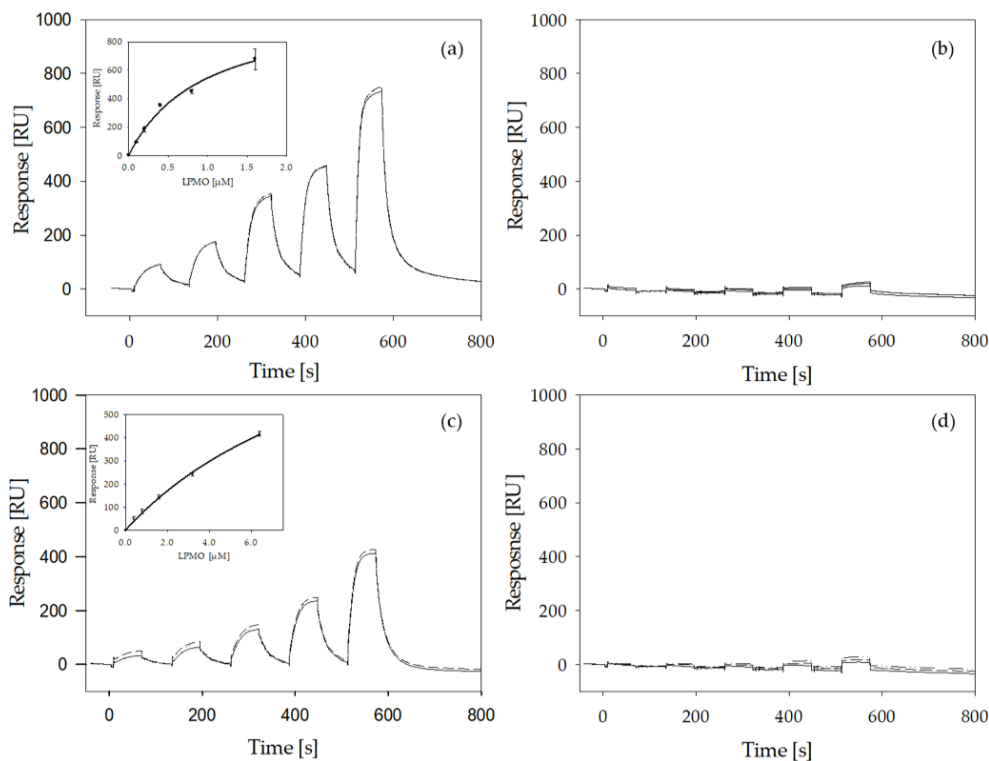


Fig. 7. Sensograms of surface plasmon resonance measurements interaction studies of LPMO on gold-modified cellulose surfaces measured by single-cycle kinetics. Five injections with increasing concentrations of the analyte were performed. (a) *Nc*LPMO9C (measured concentration: 0.1, 0.2, 0.4, 0.8 and 1.6 μM), (b) *Nc*LPMO9C ΔCBM , (c) *Nc*LPMO9C ΔSeg2 , (d) *Nc*LPMO9C ΔSeg2 , ΔCBM (all with concentrations of: 0.4, 0.8, 1.6, 3.2 and 6.4 μM). All runs were performed in triplicates. Insets in (a) and (c) show LPMO concentration vs. response units and were fitted with SigmaPlot 12.

To quantify how much the removal of the CBM1 and the shortening of Seg2 decreased the affinity we tried to determine the K_d values for *Nc*LPMO9C ΔCBM and *Nc*LPMO9C ΔSeg2 , ΔCBM , but did observe only very weak binding (**Fig. 7b and 7d**). Unfortunately, this did not allow us to perform reliable calculations. This demonstrates that the binding site of *Nc*LPMO9C, neither intact nor with the shortened Seg2, has a high affinity towards the employed cellulosic substrate. We can therefore only indirectly deduce that the

loss of activity observed for *NcLPMO9C*^{ΔSeg2} (**Fig. 5**) comes from a reduced affinity of the binding site, because *NcLPMO9C*^{ΔCBM} is active on RAC, CMC, and XG.

Investigating the binding behavior of LPMO in its reduced state to its cellulosic substrate by SPR measurements, using 10 mM gallic acid as a reducing agent, showed no significant difference of the binding affinity for *NcLPMO9C* compared to the oxidized form (Fig. S33). For *NcLPMO9C*^{ΔCBM}, a higher response could be observed in its reduced state. Unfortunately, the determination of the K_d values in the reduced state was not possible. This would require an oxygen-free environment, which unfortunately is a major hurdle in SPR measurements. Moreover, less efficient regeneration of the cellulosic surface in the presence of gallic acid did also hinder reliable experiments.

4. Discussion

Previous studies to infer the regioselectivity of LPMOs from phylogenetic trees (19, 21, 22, 24, 29, 30) resulted in the currently described four types of LPMOs (types 1, 2, 3, and 3*). The amino acid segments around the LPMO catalytic site are flexible which gives them the necessary mobility to adapt to different substrates (Fig. S34). Intriguingly, this is not changing the site-specific oxidation of the substrate. Type 1 and type 2 LPMOs both hold a short Seg1 and are supposed to have a C1 or C4 regioselectivity, respectively. The elongated Seg2 in type 2 LPMOs has been suggested to be responsible for the preference in C4 regioselectivity (19). Type 3 LPMOs were found to have an elongated Seg1 and to produce both 4-ketoaldoses and aldonic acids. However, type 3* LPMOs are defined as a subgroup of type 3, solely producing C1-oxidized products (19). This was attributed to the lack of conserved residues within Seg1. Based on our sequence alignment we found that not all C1-oxidizing LPMOs in Cluster 3 were lacking the conserved residues described by Vu and coworkers (i.e., CAP71839.1 *T. anserina* S mat+ LPMO9F, CAP66744.1 *T. anserina* S mat+ LPMO9D) (19). Therefore, we suggest that the classification of LPMO regioselectivity solely based on phylogenetic clusters is not reliable enough. One example are LPMOs within Cluster 2 that are featuring a cysteine in Seg2 (i.e., AON76800.1 *T. thermophilus* LPMO9B). A potential disulfide bond to a cysteine in Seg3 could result in a simulated short Seg2 (Fig. S6). Furthermore, according to Bey and coworkers (57), *T. anserina* S mat+ LPMO9B (CAP68375.1) is switching between the generation of C1/C4- and C1-oxidation products when reduced by ascorbic acid or cellobiose dehydrogenase (CDH), respectively. Since CDH is a natural reductant we suggest that the latter LPMO should belong to type 1 LPMOs rather than type 2. This is highlighting the complexity of the regioselectivity in LPMOs and hinting that the heterogenous distribution might be due to non-optimal experimental conditions. The large phylogenetic distances between LPMO sequences within a cluster shows that the evolutionary adaption process was done after the division into clusters. This leads to highly diverse substrate specificities and regioselectivities. In contrast, the phylogenetic distances between the three clusters are relatively short. Nevertheless, more reliable data from characterized LPMOs are necessary to provide a final proof of the correlation between LPMO phylogeny and regioselectivity.

A previous study showed that it is possible to alter the LPMO regioselectivity by altering Seg1 of the active site (19). A second study did not reproduce this result for a different LPMO (*Hj*LPMO9A) (27) but resulted in an inactive enzyme. Presumably the loss of activity in the second study was linked to a high strain on the catalytic site after the deletion of Seg1. For this reason, we did not simply shorten Seg2 from *Nc*LPMO9C but replaced it with Seg2 from *Nc*LPMO9F. Despite this strategy, the expressed variant *Nc*LPMO9C^{ΔSeg2} showed no activity or affinity towards any of the tested polysaccharide substrates. Since Seg2 is close to the catalytic site, its shortening could influence the integrity of the catalytic site as reported by Danneels and coworkers (27). However, shortening Seg2 in *Nc*LPMO9C did not result in the complete inactivation of the catalytic activity as demonstrated by the ability to determine the kinetic constants of LPMO variants for H₂O₂. The obtained K_M - and k_{cat} -values of *Nc*LPMO9C and the variant LPMOs were similar to previously published values (44). In contrast to the recently published work by Chalak and coworkers (32), we were able to successfully produce an active LPMO variant with a truncated linker and CBM1 right after the catalytic domain. However, this modification resulted in a partial reduction of the thermostability. It seems that a truncation right after the cysteine 227 of *Nc*LPMO9C is leading to an unequal folding of the protein. This is supported by the appearance of two distinct T_m values in DSC experiments. Most likely the LPMO linker is contributing to the correct folding of the LPMO and formation of the C-terminal disulfide bond back to the catalytic domain of *Nc*LPMO9C. Tanghe and coworkers have shown that disulfide bridges increase the thermal stability of a bacterial LPMO (44). The observed partial refolding of LPMOs has been previously reported (59, 60). It was suggested that this effect is connected to the disulfide bonds (59). Moreover, a similar signal collapse near 205 nm in the *Nc*LPMO9C CD spectrum was observed by Frommhagen and coworkers for two *Myceliophthora thermophila* LPMOs (61). In their study the changes in ellipticity were suggested to show that the LPMOs adopted a fibrillar-like state during heat treatment. It is known that a significantly negative spectrum around 200 nm is pointing towards a poly (L-proline) type II helix (62). The observed more negative ellipticity above 70 °C in the ECD spectra (**Fig. 4**) therefore suggests that there is no complete unfolding, but possibly a formation of fibrils.

When looking at the regioselectivity on cellulosic substrates, *Nc*LPMOs 9F, 9C and 9M share the same product pattern as previously published (19). In contrast to Chalak and coworkers (32) no change in the regioselectivity of the enzyme was observed when removing the CBM1. However, recent efforts to elucidate the role of a family 2 CBM linked to LPMO10C from *Streptomyces coelicolor* revealed no change in regioselectivity upon truncation of the CBM2 from the enzyme (26). Furthermore, the same study also showed that the quantity of formed products was decreased. This finding is further supported by our conversion experiments with *Nc*LPMO9C^{ΔCBM}. Since *Nc*LPMO9C^{ΔCBM} is lacking the CBM1 it is conceivable that the latter domain helps the enzyme to stay close to the substrate and therefore promotes the formation of even smaller products. This effect has been described by Courtade et al. (26). For the conversion of XG, the presence of the CBM1 is not necessary since similar conversion rates of *Nc*LPMO9C and *Nc*LPMO9C^{ΔCBM} were observed. However, the conversion of XG

stopped earlier when employing *NcLPMO9C*^{ΔCBM}, maybe due to a faster deactivation. The lack of activity of *NcLPMO9C*^{ΔSeg2} and *NcLPMO9C*^{ΔSeg2, ΔCBM} on XG indicates that Seg2 is important for the substrate recognition of XG.

The role of the CBM1 of *NcLPMO9C* on substrate binding was already investigated by performing isothermal titration calorimetry (ITC) experiments with both RAC and XG (24). The authors of the study demonstrated that the loss of the CBM1 resulted in a significantly higher *K_d* values for both substrates. The SPR experiments performed during this study were confirming this result. Nevertheless, we were not able to detect an interpretable signal in the absence of the CBM1. However, when turning to the *NcLPMOs* 9F and 9M we were able to calculate dissociation constants despite the lack of a carbohydrate binding module. This is pointing towards an optimization of the catalytic site for binding without a CBM1 and might explain the difference in binding affinity between *NcLPMO9C* and *NcLPMO9C*^{ΔSeg2}. A recent study provided evidence that the reduction of the catalytic site copper is improving the binding affinity of *NcLPMO9C* with a more than two times lower *K_d* (9.5 and 4.4 μM for the oxidized and the reduced LPMO, respectively) (60). In preliminary SPR experiments with reduced enzymes we were able to confirm that the binding to a cellulosic substrate was greatly improved.

5. Conclusions

In conclusion, we could show that the shortening of Seg2 of *NcLPMO9C* reduced thermal stability and the affinity to polysaccharide substrates. Moreover, our results confirm that the C-terminal linker between the *NcLPMO9C* and its CBM1 plays a role in the thermal stability. Possibly, the linker is promoting the formation of a disulfide bridge which is stabilizing the protein fold. Furthermore, we could demonstrate that the CBM1 strongly contributes to substrate binding and does not influence the regioselectivity of the reaction.

6. References

1. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink, VGH. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219–222.
2. Harris PV, Welner D, McFarland KC, Re E, Poulsen J-CN, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry*. 2010;49:3305–3316.
3. Tan TC, Kracher D, Gandini R, Sygmond C, Kittl R, Haltrich D, Hällberg BM, Ludwig R, Divne C. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat Commun*. 2015;6:7542.
4. Beeson WT, Phillips CM, Cate JHD, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134: 890–892.
5. Hangasky JA, Iavarone AT, Marletta MA. Reactivity of O₂ versus H₂O₂ with polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2018;115:4915–4920.
6. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen J-CN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A*. 2011;108:15079–15084.
7. Cannella D, Möllers KBB, Frigaard NU, Jensen PEE, Bjerrum MJJ, Johansen KS, Felby C. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. *Nat Commun*. 2016;7:11134.

8. Kracher D, Scheiblbrandner S, Felice AKG, Breslmayr E, Preims M, Ludwicka K, Haltrich D, Eijssink VGH, Ludwig R. Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science*. 2016;352:1098.
9. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res*. 2014;42:D490-D495.
10. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZY database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6:41.
11. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol*. 2014;10:122-126.
12. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2014;111:13822-13827.
13. Couturier M, Ladevèze S, Sulzenbacher G, Ciano L, Fanuel M, Moreau C, Villares A, Cathala B, Chaspoul F, Frandsen KE, et al. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol*. 2018;14:306-310.
14. Sabbadin F, Hemsworth GR, Ciano L, Henrissat B, Dupree P, Tryfona T, et al. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun*. 2018;9:756.
15. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevret D, et al. AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels*. 2019;12:55.
16. Bissaro B, Rohr AK, Muller G, Chylenski P, Skaugen M, Forsberg Z, et al. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat Chem Biol*. 2017;13:1123-8.
17. Chylenski P, Bissaro B, Sorlie M, Rohr AK, Varnai A, Horn SJ, et al. Lytic polysaccharide monooxygenases in enzymatic processing of lignocellulosic biomass. *ACS Catal*. 2019;9:4970-91.
18. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijssink VG, van Aalten DM. Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem*. 2005;280:11313-9.
19. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136:562-5.
20. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure*. 2012;20:1051-61.
21. Patel I, Kracher D, Ma S, Garajova S, Haon M, Faulds CB, Berrin J-G, Ludwig R, Record E. Salt-responsive lytic polysaccharide monooxygenases from the mangrove fungus *Pestalotiopsis* sp. NCi6. *Biotechnol Biofuels*. 2016;9:108.
22. Lenfant N, Hainaut M, Terrapon N, Drula E, Lombard V, Henrissat B. A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res*. 2017;448:166-74.
23. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, et al. Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem*. 2013;288:12828-39.
24. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.
25. Danneels B, Tanghe M, Joosten HJ, Gundinger T, Spadiut O, Stals I, Desmet T. A quantitative indicator diagram for lytic polysaccharide monooxygenases reveals the role of aromatic surface residues in H₂LPMO9A regioselectivity. *PLOS ONE*. 2017;12:e0178446.
26. Courtade G, Forsberg Z, Heggset EB, Eijssink VGH, Aachmann FL. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J Biol Chem*. 2018;293:13006-15.
27. Danneels B, Tanghe M, Desmet T. Structural features on the substrate-binding surface of fungal lytic polysaccharide monooxygenases determine their oxidative regioselectivity. *Biotechnol J*. 2019;14:1800211.
28. Forsberg Z, Bissaro B, Gullesen J, Dalhus B, Vaaje-Kolstad G, Eijssink VGH. Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J Biol Chem*. 2018;293:1397-412.
29. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6:1399-406.

30. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels*. 2015;8:90.
31. Jagadeeswaran G, Gainey L, Mort AJ. An AA9-LPMO containing a CBM1 domain in *Aspergillus nidulans* is active on cellulose and cleaves cello-oligosaccharides. *AMB Express*. 2018;8:171.
32. Chalak A, Villares A, Moreau C, Haon M, Grisel S, d'Orlando A, Herpoël-Gimbert I, Labourel A, Cathala B, Berrin JG. Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol Biofuels*. 2019;12:206.
33. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. 2019;35: 4453–4455.
34. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol*. 2019;37:291–294.
35. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 2001;18:691–699.
36. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol*. 2010;17:337–354.
37. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep*. 2017;7:10480.
38. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 2011, 27, 343–350.
39. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 2009;30: S162–S173.
40. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res*. 2017;45:D313–D319.
41. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296–W303.
42. Zhang Y-HP, Cui J, Lynd LR, Kuang LR. A transition from cellulose swelling to cellulose dissolution by O-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules*. 2006;7:644–648.
43. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
44. Breslmayr E, Hanzek M, Hanrahan A, Leitner C, Kittl R, Santek B, et al. A fast and sensitive activity assay for lytic polysaccharide monooxygenase. *Biotechnol Biofuels*. 2018;11:79.
45. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee Y-H, Goto Y, Réfrégiers M, Kardos J. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res*. 2018, 46, W315–W322.
46. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
47. Raj G, Balnois E, Baley C, Grohens Y. Probing interfacial interactions in natural fibre reinforced biocomposites using colloidal force microscopy. In *Natural Polymers*. Volume 1: Composites; John MJ, Sabu T, Eds. The Royal Society of Chemistry: Cambridge UK, 2012; pp. 183–203 ISBN 978-1-84973-402-8.
48. Szech R, Riegler H. Molecularly smooth cellulose surfaces for adhesion studies. *J Colloid Interface Sci*. 2006;301:376–385.
49. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol*. 2018;9:1080.
50. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983;22:2577–2637.
51. Eichenberger AP, Allison JR, Dolenc J, Geerke DP, Horta BAC, Meier K, Oostenbrink C, Schmid N, Steiner D, Wang D, et al. GROMOS++ software for the analysis of biomolecular simulation trajectories. *J Chem Theory Comput*. 2011;7:3379–3390.

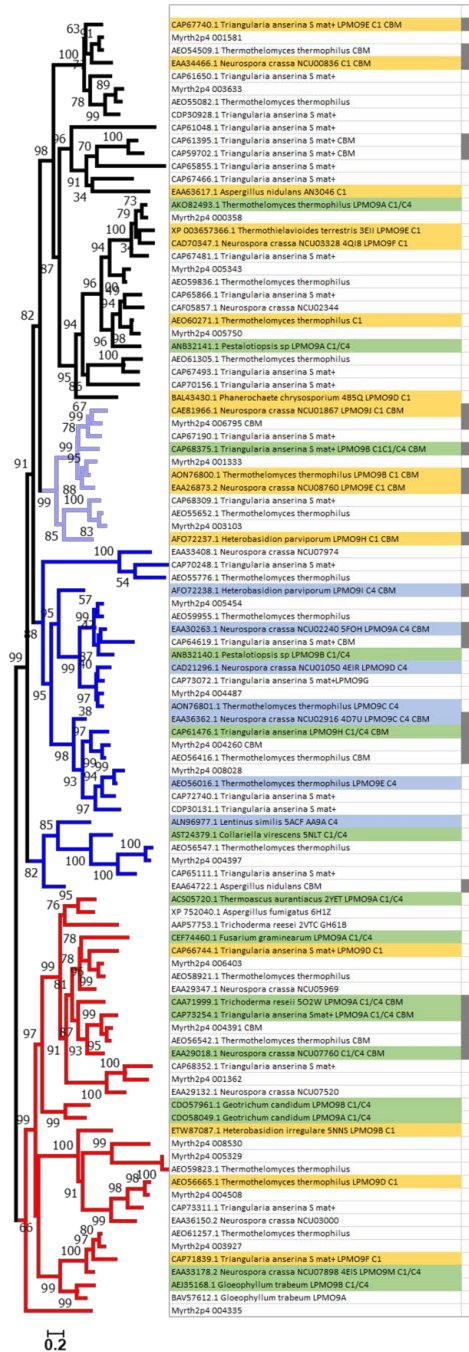
52. Durowoju IB, Bhandal KS, Hu J, Carpick B, Kirkitadze M. Differential scanning calorimetry—a method for assessing the thermal stability and conformation of protein antigen. *J Vis Exp*. 2017;121:e55262.
53. Westereng B, Ishida T, Vaaje-Kolstad G, Wu M, Eijsink VGH, Igarashi K, Samejima M, Ståhlberg J, Horn SJ, Sandgren M. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLOS ONE*. 2011;6:e27807.
54. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, et al. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20:1479-83.
55. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
56. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A*. 2014;111:6287-92.
57. Bey M, Zhou S, Poidevin L, Henrissat B, Coutinho PM, Berrin JG, et al. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospira anserina*. *Appl Environ Microbiol*. 2013;79:488-96.
58. Tanghe M, Danneels B, Last M, Beerens K, Stals I, Desmet T. Disulfide bridges as essential elements for the thermostability of lytic polysaccharide monooxygenase LPMO10C from *Streptomyces coelicolor*. *Protein Eng Des Sel*. 2017;30:401-408.
59. Singh RK, Blossom BM, Russo DA, van Oort B, Croce R, Jensen PE, Felby C, Bjerrum MJ. Thermal unfolding and refolding of a lytic polysaccharide monooxygenase from *Thermoascus aurantiacus*. *RSC Adv*. 2019;9:29734-29742.
60. Kracher D, Andlar M, Furtmüller PG, Ludwig R. Active-site copper reduction promotes substrate binding of fungal lytic polysaccharide monooxygenase and reduces stability. *J Biol Chem*. 2018;293:1676-1687.
61. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol*. 2017;102:1281-95.
62. Rodger A, Ismail MA. Introduction to circular dichroism (Chapter 4). In *Spectrophotometry and Spectrofluorimetry*; Gore MG, Ed. Oxford University Press: Avon, UK, 2000; pp. 99-139 ISBN 978-0199638130.

7. Supplementary information

Fig. S1 is available online (open access). <https://www.mdpi.com/1422-0067/20/24/6219>

Fig. S1. Full length multiple sequence alignment without linker region and CBM and associated phylogenetic tree of AA9 LPMOs. All available AA9 LPMO sequences of *Neurospora crassa*, *Podospora anserina* S mat+ (syn. *Triangularia anserina* S mat+), *Thermothelomyces thermophilum* (syn. *Myceliophthora thermophila*), *Crassiparpon hotsonii* (syn *Myriococcum thermophilum*) as well as all AA9 LPMOs listed under the headers “characterized” and “structure” in the CAZy database were aligned using the MAFFT-DASH algorithm based on the protein secondary structure. Prior to phylogenetic analysis the C-terminal linker region and the CBM1 were deleted from the alignment. Phylogeny was inferred by using the RaxML-NG algorithm applying the Wheelan & Goldman model with frequencies and gamma distribution. The color code of the tree corresponds to the top line of the alignment and indicates the respective segments: Seg1, red; Seg2, blue; Seg3, orange; Seg4, green; Seg5, magenta. Gray squares next to the sequence name column indicate the presence of a CBM1. The color code used to color sequence names indicates the regioselectivity of AA9 LPMOs as previously published (1): C1, yellow; C4, blue; C1/C4, green. Amino acids considered relevant for catalysis or substrate interaction are colored in the multiple sequence alignment: His, light blue; Arg and Lys, blue; Asp and Glu, red; Asn and Gln, green; Tyr and Phe, yellow.

full length



segments only

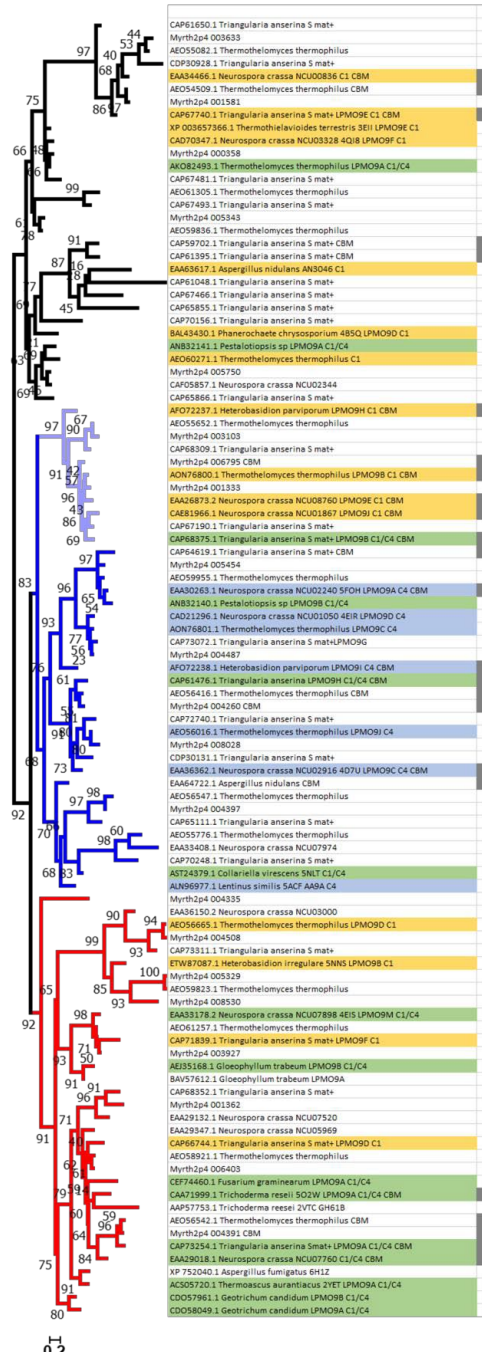


Fig. S2. Comparison of the phylogenetic trees inferred from full length (left) and "segments only" (right) multiple sequence alignments. The three main clusters are color coded (Cluster 1, black;

Cluster 2, blue; Cluster 3) and a subgroup of sequences changing from Cluster 1 to Cluster 2 when phylogenetic inference is not based on the full dataset, but the sequence information is highlighted in pale blue. Taxa color code: yellow, C1-oxidizing; blue, C4-oxidizing; green, C1/C4-oxidizing. Gray boxes besides the taxa names indicate the presence of a CBM.

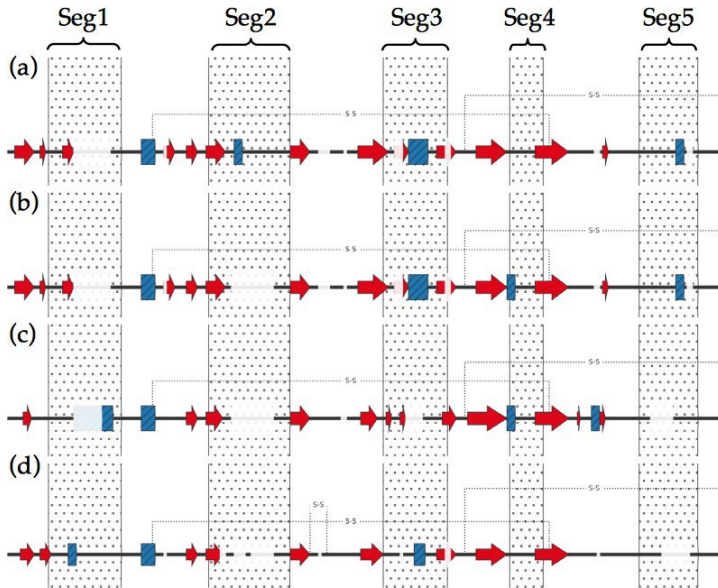


Fig. S3. Schematic representation of the secondary structure elements of the catalytic domains. (a) *NcLPMO9C*, (b) *NcLPMO9C*^{ΔSeg2}, (c) *NcLPMO9F* and (d) *NcLPMO9M*. α -Helices and β -strands are shown in blue and red, respectively. Disulfide bonds are depicted as dotted lines. Transparent regions are indicating gaps in the multiple sequence alignment and dotted regions are highlighting the Seg1–5 as defined as given in Table S1. The secondary structure elements were determined using the DSSP algorithm (2) as implemented in GROMOS++ (3).

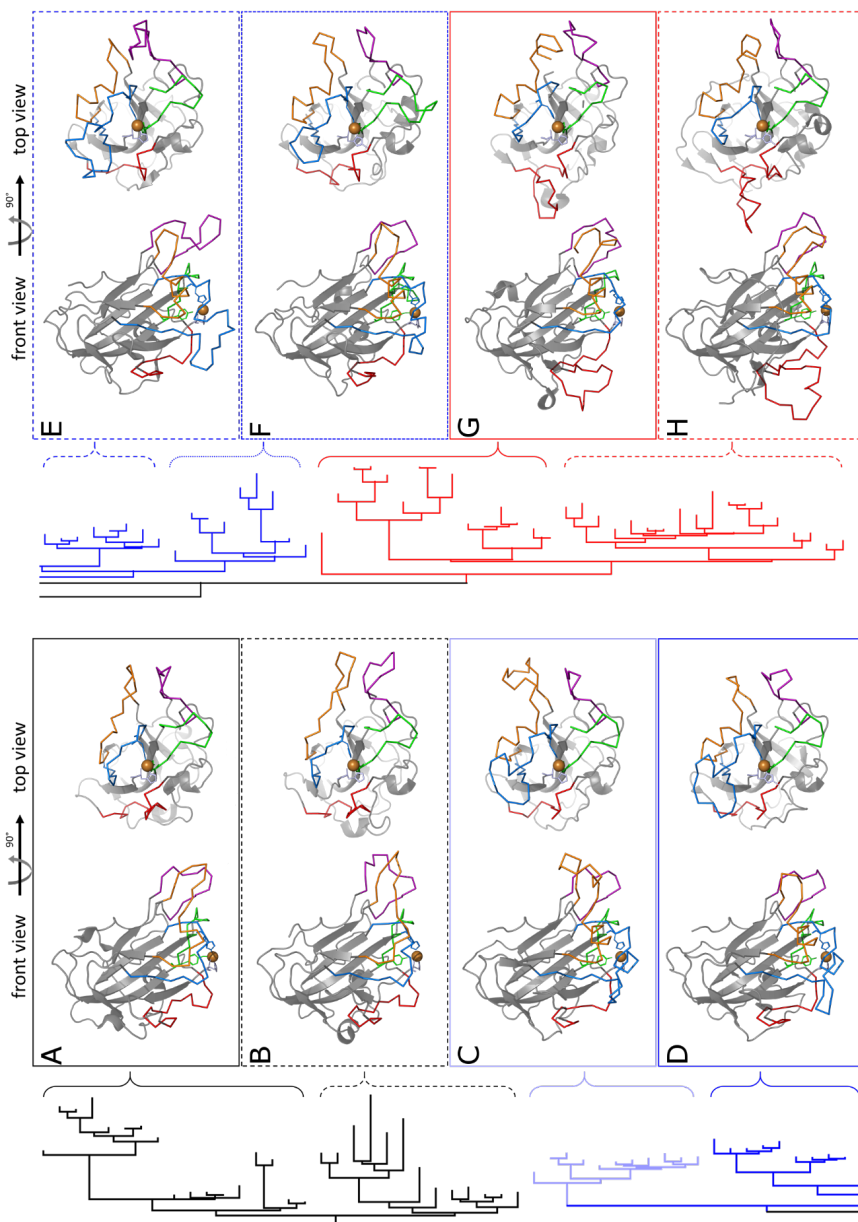


Fig. S4. Front and top view of crystal structures of selected AA9 LPMOs. The crystal structures and homology models of eight LPMOs is shown in "front view", at which the copper atom and the putative, planar-like substrate surface binding region is facing the bottom, and "top view", at which the previous visualization is rotated by 90°. Each structure represents one of eight distinct clades of the phylogenetic tree. (A) *NcLPMO9F* (GeneBank: CAD70347.1; PDB: 4QI8), (B) *PcLPMO9D* (GeneBank: BAL43430.1; PDB: 4B5Q), (C) *TtLPMO9B* (GeneBank: AON76800.1; SwissModel template: 5TKF with 41.68 % sequence identity), (D) *NcLPMO9D* (GeneBank: CAD21296.1; PDB: 4EIR), (E) *NcLPMO9C* (GeneBank: EAA36362.1; PDB: 4D7U), (F) *CvAA9A* (GeneBank: AST24379.1; PDB: 5NLT), (G) *HiLPMO9B* (GeneBank: ETW87087.1; PDB: 5NNS),

and (H) TrLPMO9A (GeneBank: CAA71999.1; PDB: 5O2W). Segments 1 to 5 are colored according to the previously introduced color code (S1: red; S2: blue; S3: orange; S4: green; S5: purple). The catalytic copper atom is shown as an orange sphere.

Fig. S5 is available online (open access). <https://www.mdpi.com/1422-0067/20/24/6219>

Fig. S5. "Segments-only" multiple sequence alignment. All available AA9 LPMO sequences of *Neurospora crassa*, *Podospira anserina* S mat+ (syn. *Triangularia anserina* S mat+), *Thermothelomyces thermophilum* (syn. *Myceliophthora thermophila*), *Crassiparson hotsonii* (syn. *Myriococcum thermophilum*) as well as all AA9 LPMOs listed under the headers "characterized" and "structure" in the CAZy database were aligned using the MAFFT-DASH algorithm based on the protein secondary structure. The multiple sequence alignment was then assembled in a way that left only the indicated segments (Seg1–5) in the alignment. The phylogeny was inferred by the RaxML-NG algorithm applying the Whelan & Goldman model with frequencies and gamma distribution. The color code of the tree corresponds to the top line of the alignment and indicates the position/possession of the respective segments: Seg1, red; Seg2, blue; Seg3, orange; Seg4, green; Seg5, magenta. Gray squares next to the sequence name column indicate the presence of a CBM1. The color code used to color sequence names indicates the regioselectivity of AA9 LPMOs as previously published (1): C1, yellow; C4, blue; C1/C4, green. Amino acids considered relevant for catalysis or substrate interaction are colored in the multiple sequence alignment: His, light blue; Arg and Lys, blue; Asp and Glu, red; Asn and Gln, green; Tyr and Phe, yellow.

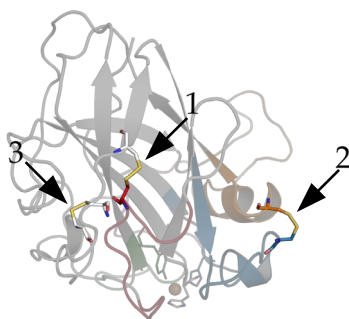


Fig. S6. Cartoon representation of the homology model of AON76800.1 *Thermothelomyces thermophilus* LPMO9B. The segments are colored red (Seg1), blue (Seg2), orange (Seg3), green (Seg4) and magenta (Seg5). The disulfide bonds are highlighted by stick representation. The numbers 1–3 indicate the disulfide bonds between the residues 18–49, 38–173 and 70–133, respectively. The second disulfide bond is binding and shortening Seg2 and thereby mimicking a short Seg2. The homology model was generated using SWISS-MODEL (4–8) (template: PDB entry 5TKF, sequence identity: 41.86%, QMEAN: -2.03).

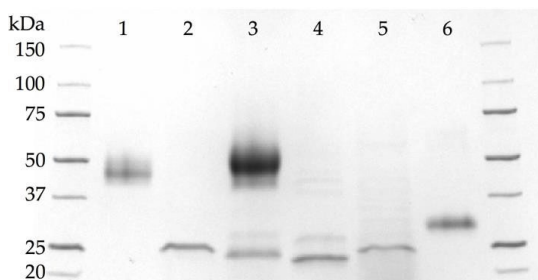


Fig. S7. SDS-PAGE of purified LPMOs and LPMO variants. NcLPMO9C (Lane 1), NcLPMO9C Δ CBM (Lane 2), NcLPMO9C Δ Seg2 (Lane 3), NcLPMO9C Δ Seg2, Δ CBM (Lane 4), NcLPMO9F (Lane 5) and NcLPMO9M (Lane 6). Lanes and molecular weights of the marker proteins (Precision Plus Protein Unstained Standards, Bio-Rad) are indicated. The precast gel (4–20% Mini-PROTEAN[®] TGX Stain-Free[™] Precast Gel,

BioRad) was stained with colloidal Coomassie blue (Bio-Safe[™] Coomassie G250 Stain, BioRad) for 1 h and destained overnight.

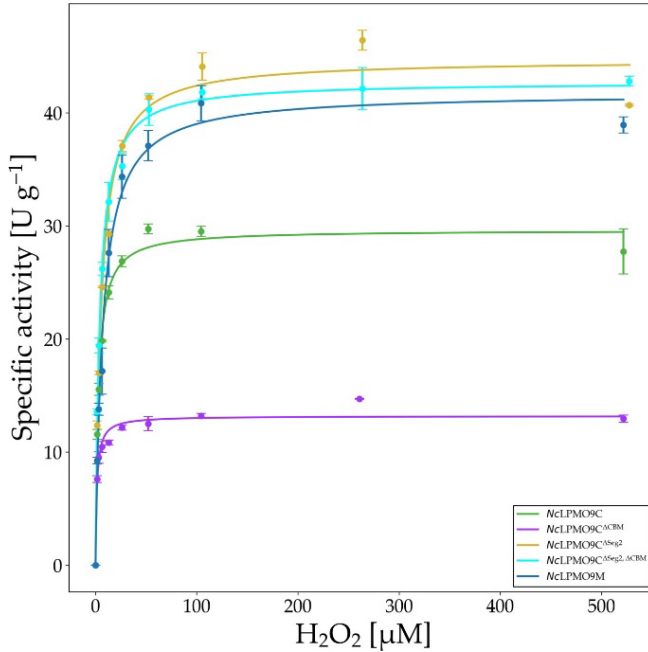


Fig. S8. Determination of kinetic constants of NcLPMO9M, NcLPMO9C and NcLPMO9C variants for H_2O_2 . As substrate 2 mM 2,6-DMP in 100 mM sodium acetate buffer, pH 6.0 was used. The LPMO activity assay was performed as previously published (9). The calculated K_M and k_{cat} values are given in **Table 1**. Error bars show the standard deviation of three replicates.

(a)	1	HTIFQKVS	VNGADGQQLK	GI	RAPANNNP	PVTDVMSSDII	CNAV	TMKDSNV	LTVPAGAK	VGH	60		
	61	FWGHEIGGA	AGPNDADNP	IAASHKGP	IMVYLAKV	DNAATTGT	SGLKWF	KVAEAGL	SNGKW		120		
	121	AVDDL	IANNWSYFD	MPTC	IA	PGQYLM	RAELIALHN	AGSQAQ	FYIGCAQ	INVTGGGSA	180		
	181	SPSNTV	SFPGAYSAS	DPGIL	INIYGG	SGKTDNGG	KPYQIP	GPALFT	CPAGGSGGSS	PAPA	240		
	241	TTASTPK	PTSASAPK	PVSTTAST	PKPTNGS	SGSGTGA	AHSTKCGG	SKPAAT	TKASNPQ	PTN	300		
	301	GGNSAV	RAAALYG	QC	GGGKWT	GPTSCAS	GTCKF	SNDW	YSQCLP		360		
(b)	1	HTIFQKVS	VNGADGQQLK	GI	RAPANNNP	PVTDVMSSDII	CNAV	TMKDSNV	LTVPAGAK	VGH	60		
	61	FWGHEIGGA	AGPNDADNP	IAASHKGP	IMVYLAKV	DNAATTGT	SGLKWF	KVAEAGL	SNGKW		120		
	121	AVDDL	IANNWSYFD	MPTC	IA	PGQYLM	RAELIALHN	AGSQAQ	FYIGCAQ	INVTGGGSA	180		
	181	SPSNTV	SFPGAYSAS	DPGIL	INIYGG	SGKTDNGG	KPYQIP	GPALFT	CPAGGSGGSS	PAPA	240		
												240	
(c)	1	HTIFQKVS	VNGADGQQLK	GI	RAPANNNP	PVTDVMSSDII	CNAV	TMKDSNV	LTVPAGAK	VGH	60		
	61	FWAPS	VYHKGPI	IMVYLAKV	DNAATTGT	SGLKWF	KVAEAGL	SNGKW	AVDDL	IANNWSYFD	120		
	121	MPTC	IA	PGQYLM	RAELIALHN	AGSQAQ	FYIGCAQ	INVTGGG	SASPNTV	SFPGAYSAS	180		
	181	DPGIL	INIYGG	SGKTDNGG	KPYQIP	GPALFT	CPAGGSGG	SPAPATT	ASTPKPTS	ASAPK	240		
	241	PVSTTAST	PKPTNGS	SGSGTGA	AHSTKCGG	SKPAAT	TKASNPQ	TNGGNSAV	RAAALYG	QC	300		
	301	GGKWT	GPTSCAS	GTCKF	SNDW	YSQCLP					360		
(d)	1	HTIFQKVS	VNGADGQQLK	GI	RAPANNNP	PVTDVMSSDII	CNAV	TMKDSNV	LTVPAGAK	VGH	60		
	61	FWAPS	VYHKGPI	IMVYLAKV	DNAATTGT	SGLKWF	KVAEAGL	SNGKW	AVDDL	IANNWSYFD	120		
	121	MPTC	IA	PGQYLM	RAELIALHN	AGSQAQ	FYIGCAQ	INVTGGG	SASPNTV	SFPGAYSAS	180		
	181	DPGIL	INIYGG	SGKTDNGG	KPYQIP	GPALFT	CPAGGSGG	SPAPATT	ASTPKPTS	ASAPK	240		
												240	
(e)	1	HYTFPKV	WNSGTTAD	WQYVRR	ADNWQNN	GFVDNVS	QQIRCF	QSTHSPA	QSTLSVA	AGT	60		
	61	TITYG	AAFSVYH	PGPMQ	FYLARV	PDGQDINS	TGEGAW	FKIYHEQ	PTFGS	QLTWSSNGK	120		
	121	SSFPV	KIPISCK	SGSYLL	RAEHIGH	VAQSSGA	AFYISCA	QLSITG	GGSTEP	GANYKVS	180		
	181	FPGAY	KASDPGIL	LININYP	VPVTSY	KNPGPS	VFCTC				240		
												240	
(f)	1	HGFVD	NATIGGQ	FYQFYQ	QYQDPY	MGSPDP	IRSKIP	GNPGV	EDVTS	LAIQC	NADSAPAK	60	
	61	LHASA	AGSTV	TLRW	TIWPD	SHVGP	VITYMAR	CPDTC	QDWT	PSASD	KVWF	KIEGGR	120
	121	TSNV	WAATPL	MTAP	ANYEAI	PSCLK	PGYLV	RHEI	IALHS	AYSYP	GAFY	PGCHQLQVT	180
	181	GGG	TKTP	SSGLV	SFPGAY	KSTD	PGVTD	YDAY	QAATY	TIPG	PAVFTC		240
													240

Fig. S9. Analysis of NcLPMO variants by liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS) confirmed their mutations. Section a, NcLPMO9C; Section b, NcLPMO9C Δ^{CBM} ; Section c, NcLPMO9C Δ^{Seg2} ; Section d, NcLPMO9C $\Delta^{\text{Seg2}, \Delta^{\text{CBM}}}$; Section e, NcLPMO9F and Section f, NcLPMO9M. Sequence sections colored in red were identified by the software Global Proteome Machine (GPM) by performing MS/MS ion searches against a homemade database containing the host cell proteome and all target sequences. Amino acids marked in green are marked as "unlikely to find" due to too short or too long peptides (or glycosylation).

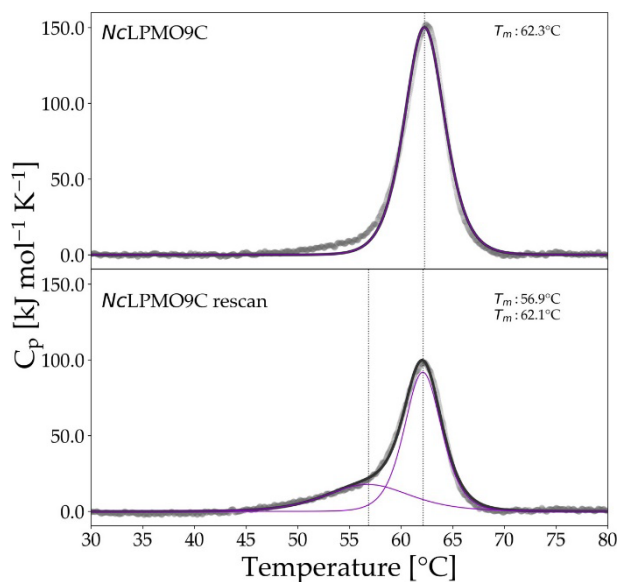


Fig. S10. Differential scanning calorimetry (DSC) thermograms of the scan and rescan of NcLPMO9C. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted curves are shown as a purple line and the transition midpoint temperature (T_m) is indicated by a vertical dotted line. Starting the scan at 20°C , a temperature ramp of 1 K min^{-1} was applied until reaching 90°C . For the rescan, the sample was cooled down with a temperature ramp of 1 K min^{-1} immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A $15\text{ }\mu\text{M}$ NcLPMO9C concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

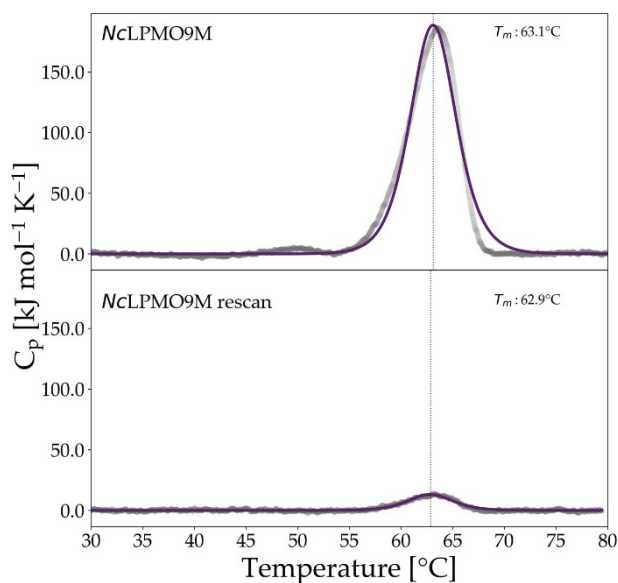


Fig. S11. DSC thermograms of the scan and rescan of NcLPMO9M. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted T_m curves are shown as a purple line and the transition midpoint temperatures are indicated by a vertical dotted line. Starting the scan at 20°C , a temperature ramp of 1 K min^{-1} was applied until reaching 90°C . For the rescan, the sample was cooled down with a temperature ramp of 1 K min^{-1} immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A $15\text{ }\mu\text{M}$ NcLPMO9C concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

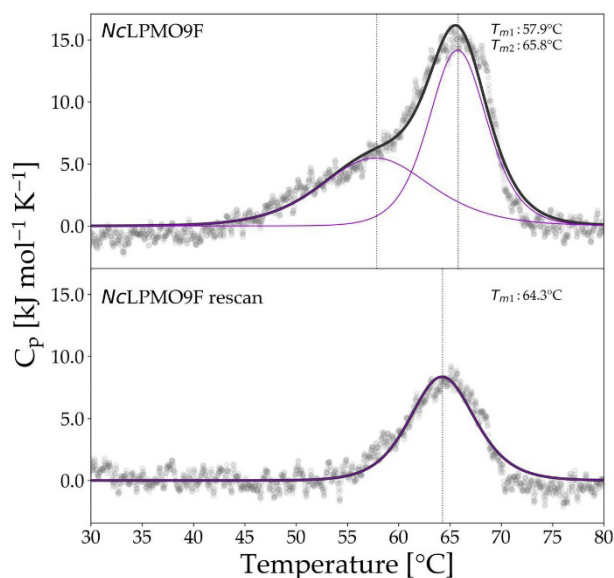


Fig. S12. DSC thermograms of the scan and rescan of *NcLPMO9F*. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted T_m curves are shown as a purple line and the transition midpoint temperatures are indicated by a vertical dotted line. Starting the scan at 20°C , a temperature ramp of 1 K min^{-1} was applied until reaching 90°C . For the rescan, the sample was cooled down with a temperature ramp of 1 K min^{-1} immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A $15\text{ }\mu\text{M}$ *NcLPMO9C* concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

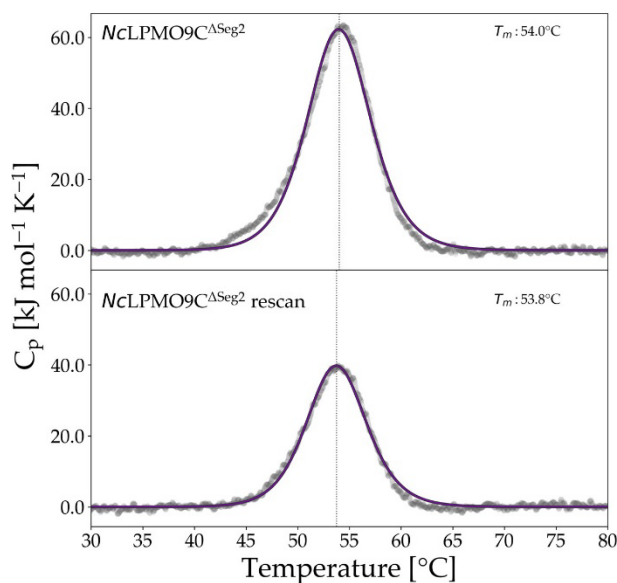


Fig. S13. DSC thermograms of the scan and rescan of *NcLPMO9C^{ΔSeg2}*. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted T_m curves are shown as a purple line and the transition midpoint temperatures are indicated by a vertical dotted line. Starting the scan at 20°C , a temperature ramp of 1 K min^{-1} was applied until reaching 90°C . For the rescan, the sample was cooled down with a temperature ramp of 1 K min^{-1} immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A $15\text{ }\mu\text{M}$ *NcLPMO9C* concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

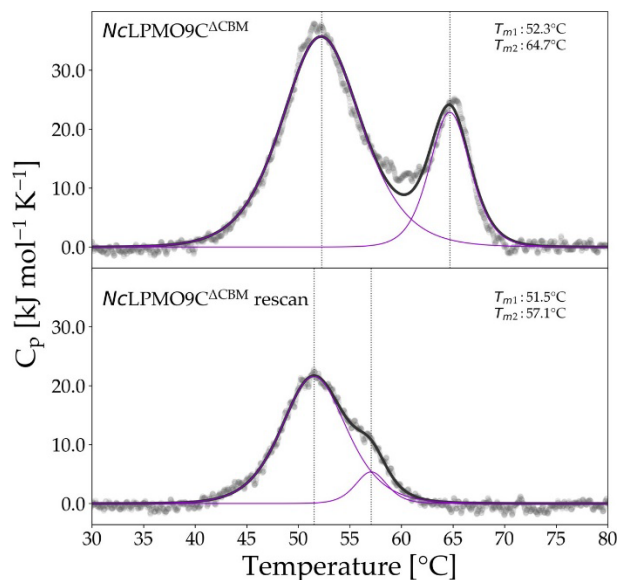


Fig. S14. DSC thermograms of the scan and rescan of $NcLPMO9C^{\Delta CBM}$. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted T_m curves are shown as a purple line and the transition midpoint temperatures are indicated by a vertical dotted line. Starting the scan at 20 °C, a temperature ramp of 1 K min⁻¹ was applied until reaching 90 °C. For the rescan, the sample was cooled down with a temperature ramp of 1 K min⁻¹ immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A 15 μM $NcLPMO9C$ concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

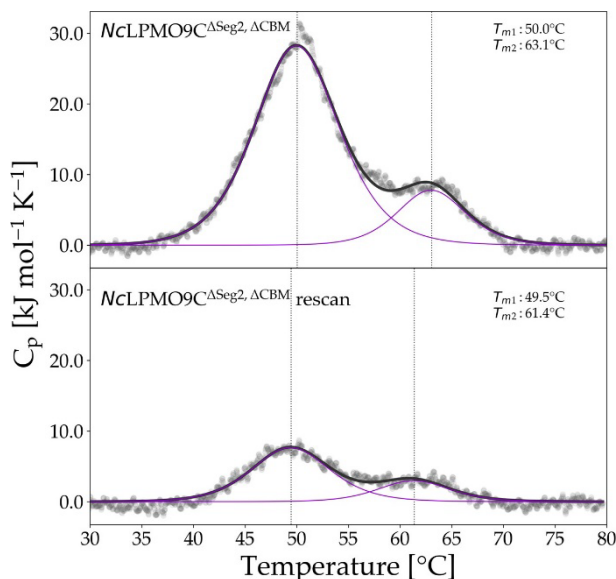


Fig. S15. DSC thermograms of the scan and rescan of $NcLPMO9C^{\Delta Seg2, \Delta CBM}$. Baseline subtracted raw data are shown as a grey scatter plot and the fitted curve is shown in black. The fitted T_m curves are shown as a purple line and the transition midpoint temperatures are indicated by a vertical dotted line. Starting the scan at 20 °C, a temperature ramp of 1 K min⁻¹ was applied until reaching 90 °C. For the rescan, the sample was cooled down with a temperature ramp of 1 K min⁻¹ immediately after reaching 90°C and when reaching 20°C the rescan was started without delay. A 15 μM $NcLPMO9C$ concentration in 50 mM potassium phosphate buffer, pH 6.0 was applied to the instruments measurement cell.

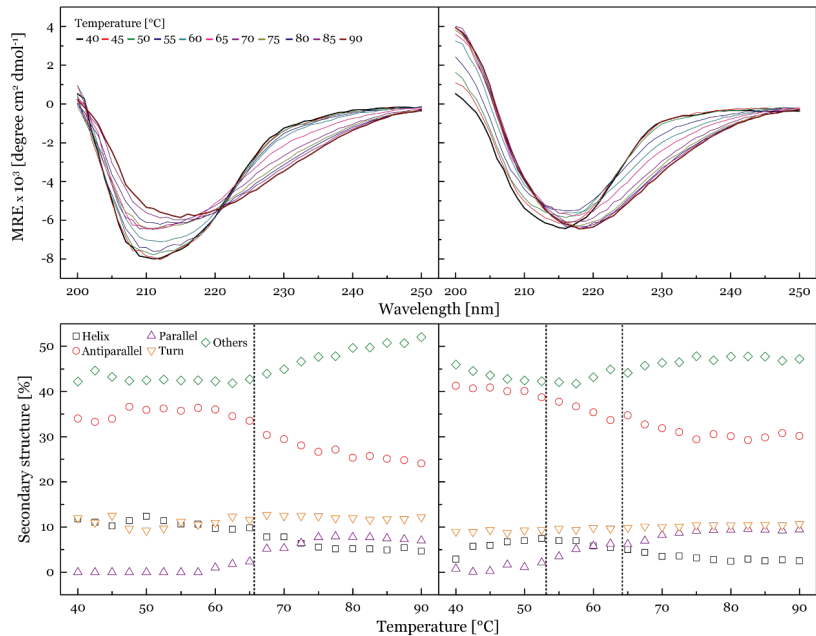


Fig. S16. Monitoring of thermal denaturation by circular dichroism (CD). ECD spectra of 3 mg mL⁻¹ NcLPMO9C (left panels) and NcLPMO9C Δ CBM (right panels) at different temperatures (top panels) for a wavelength range between 200–250 nm. Using the BestSel web server (<http://bestsel.elte.hu>) the secondary structure was predicted for all different temperatures (lower panels) (11).

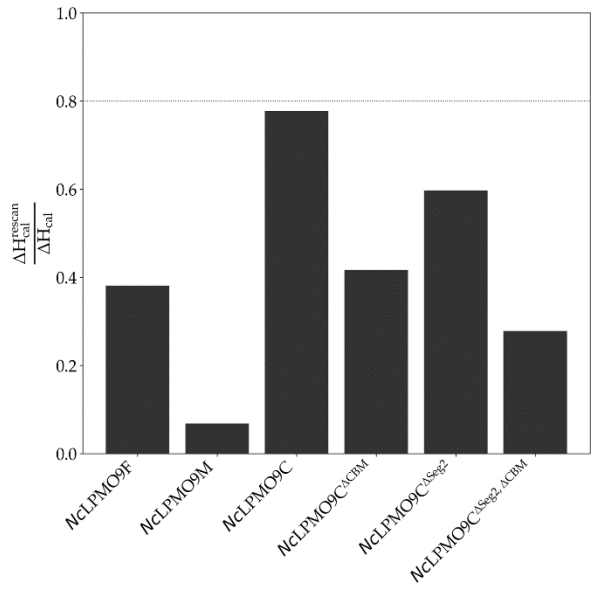


Fig. S17. Bar plot showing the ratio of ΔH_{cal} of the DSC rescan and ΔH_{cal} of the first scan. This ratio is used to assess possible refolding of the protein (10). The horizontal line indicates the threshold for reversible unfolding. Data were calculated from the thermograms shown in Fig. S10–S15.

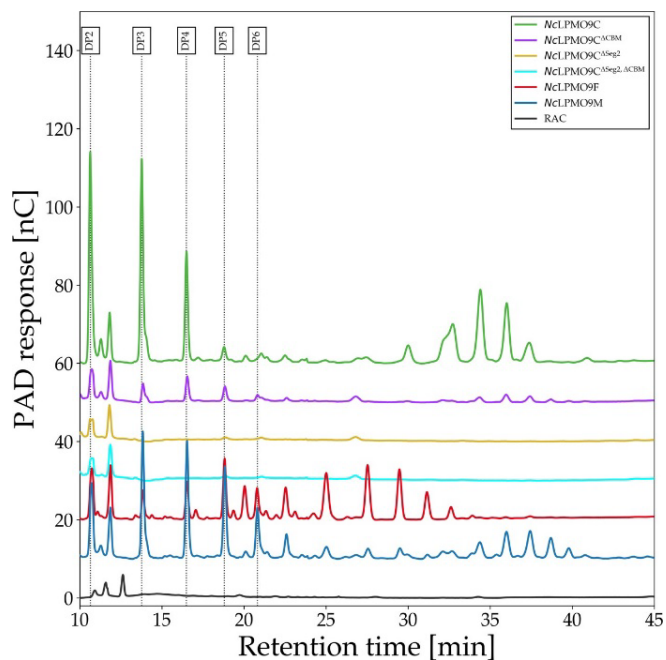


Fig. S18.

HPAEC chromatogram of products released from the reaction of 2 g L^{-1} RAC with $1.25 \text{ }\mu\text{M}$ of the indicated LPMOs in the presence of 1 mM ascorbic acid after 24 h of incubation at $30 \text{ }^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, $\text{pH } 5.5$. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10^{-14} (12).

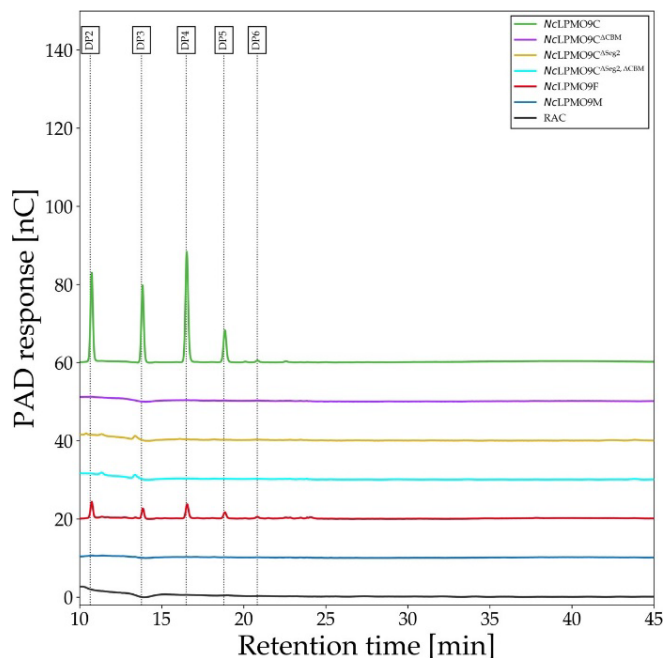


Fig. S19.

HPAEC chromatogram of products released from the reaction of 2 g L^{-1} RAC with $1.25 \text{ }\mu\text{M}$ of the indicated LPMOs in the absence of ascorbic acid after 24 h of incubation at $30 \text{ }^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, $\text{pH } 5.5$. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10^{-14} (12).

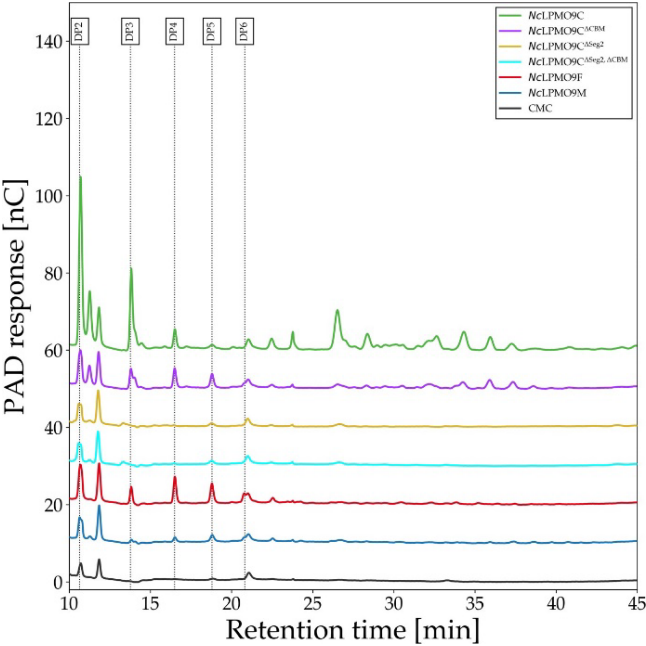


Fig. S20.

HPAEC chromatogram of products released from the reaction of 2 g L⁻¹ CMC with 1.25 μM of the indicated LPMOs in the presence of 1 mM ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10⁻¹⁴ (12).

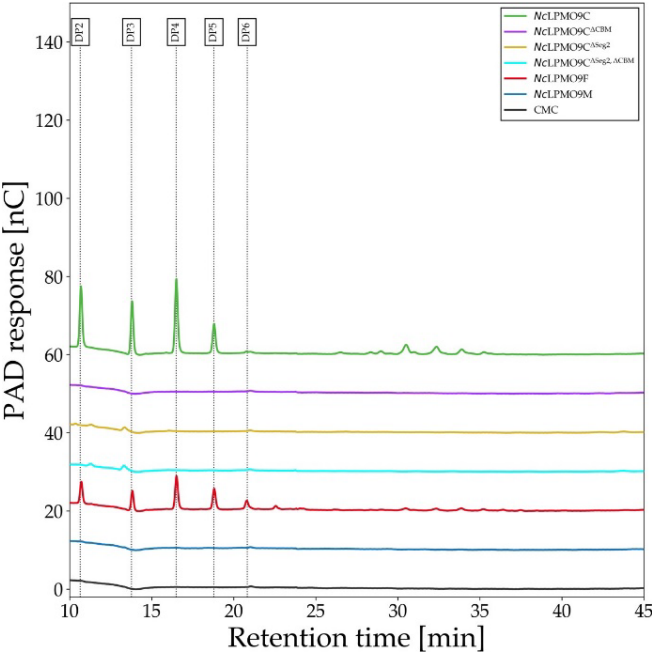


Fig. S21.

HPAEC chromatogram of products released from the reaction of 2 g L⁻¹ CMC with 1.25 μM of the indicated LPMOs in the absence of ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10⁻¹⁴ (12).

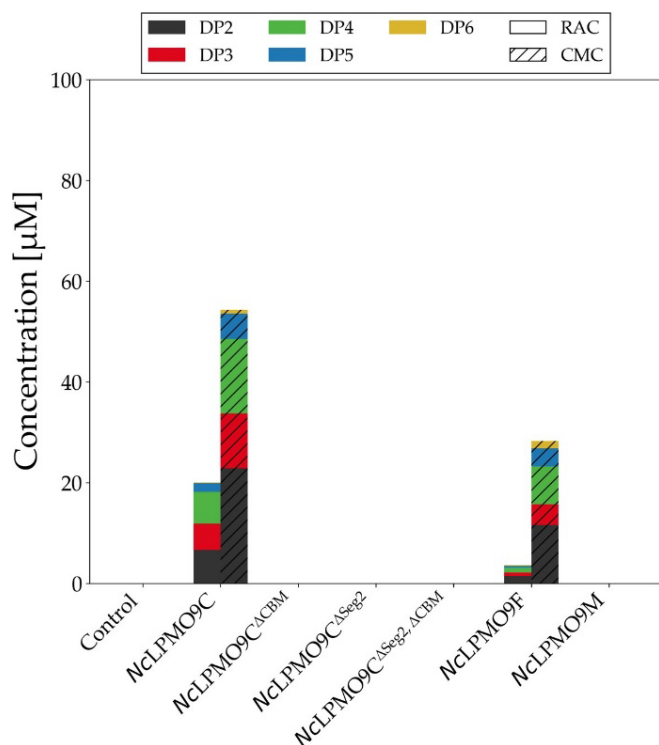


Fig. S22. Quantification of cello-oligosaccharides in the supernatant released in a batch conversion by 1.25 μM of LPMO and LPMO variants after a 24 h incubation at 30 $^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, pH 5.5 on 2 g L^{-1} RAC or CMC in the absence of ascorbic acid. The integration windows for product quantitation were defined as described in Table S3.

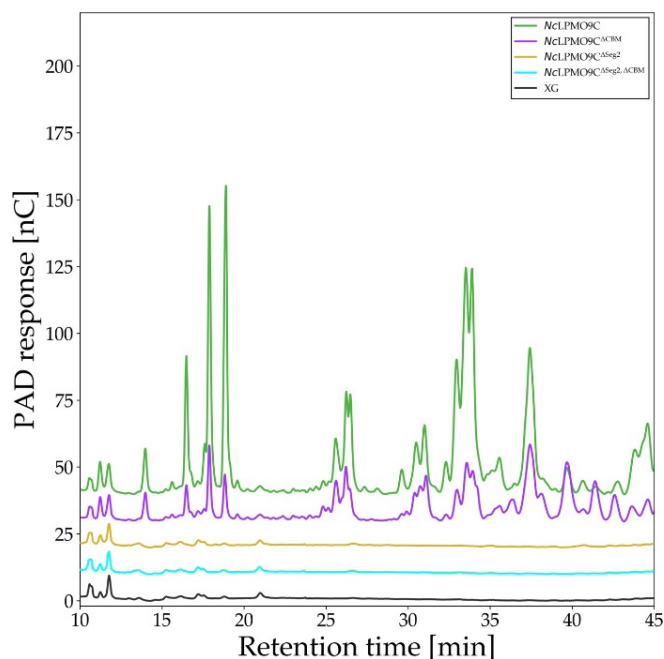


Fig. S23.

HPAEC chromatogram of xyloglucan oligosaccharides released from the reaction of 2 g L^{-1} XG with 1.25 μM of the indicated LPMOs in the presence of 1 mM ascorbic acid after 24 h of incubation at 30 $^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10^{-14} (12).

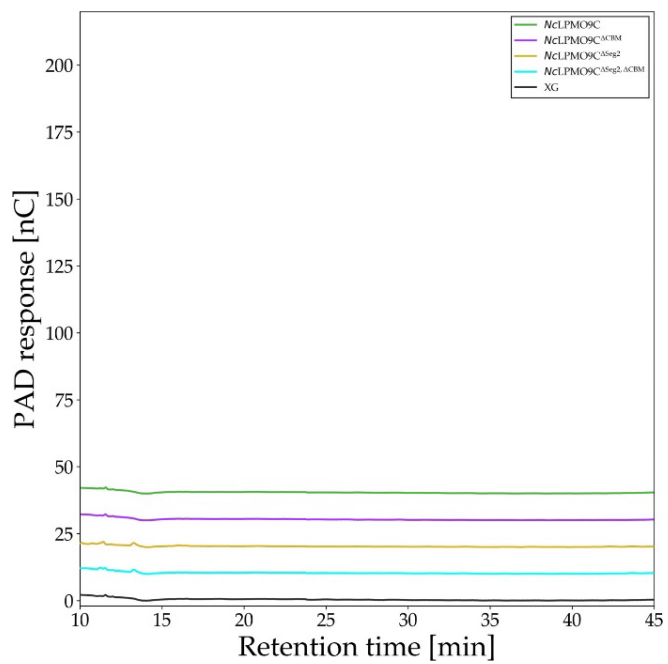


Fig. S24.

HPAEC chromatogram of xyloglucan oligosaccharides released from the reaction of 2 g L⁻¹ XG with 1.25 μM of the indicated LPMOs in the absence of ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 3rd degree and a tolerance of 10⁻¹⁴ (12).

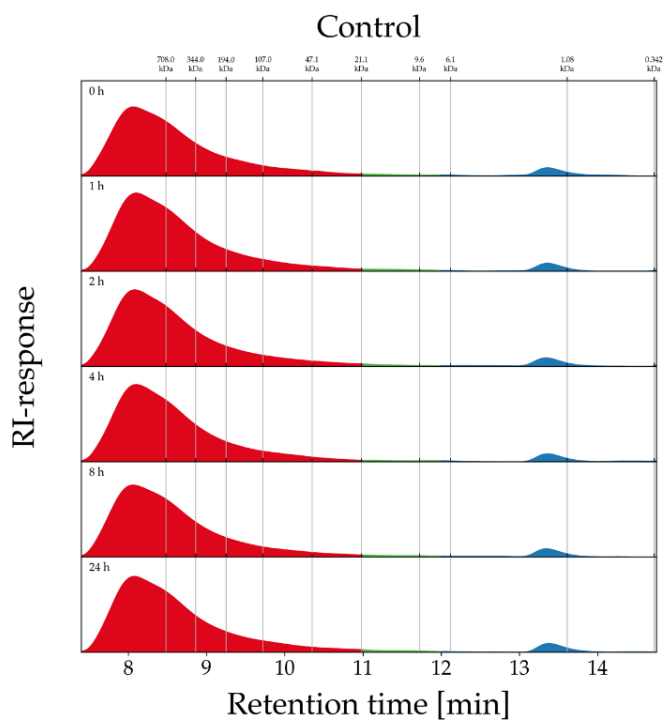


Fig. S25.

HPSEC chromatogram of products released from the reaction of 2 g L⁻¹ XG without LPMO in the absence of ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2nd degree and a tolerance of 10⁻¹⁴ (12).

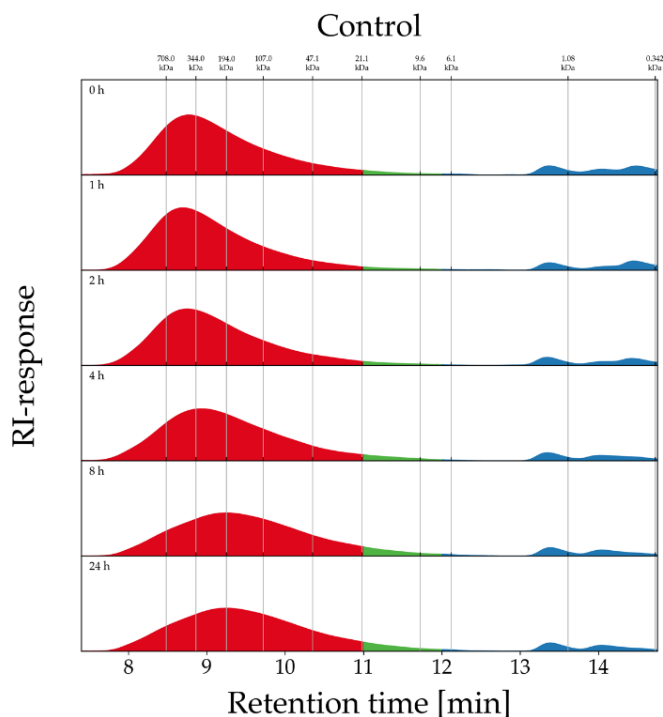


Fig. S26.

HPSEC chromatogram of products released from the reaction of 2 g L^{-1} XG without LPMO in the presence of 1 mM ascorbic acid after 24 h of incubation at $30 \text{ }^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, $\text{pH } 5.5$. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2^{nd} degree and a tolerance of 10^{-14} (12).

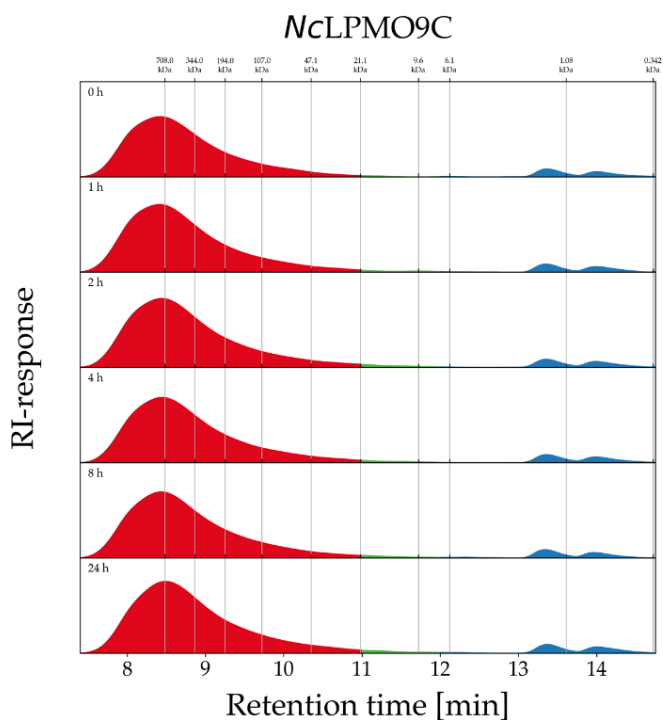


Fig. S27. HPSEC chromatogram of products released from the reaction of 2 g L^{-1} XG with $1.25 \text{ } \mu\text{M}$ NcLPMO9C in the absence of ascorbic acid after 24 h of incubation at $30 \text{ }^{\circ}\text{C}$ in 50 mM ammonium acetate buffer, $\text{pH } 5.5$. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2^{nd} degree and a tolerance of 10^{-14} (12).

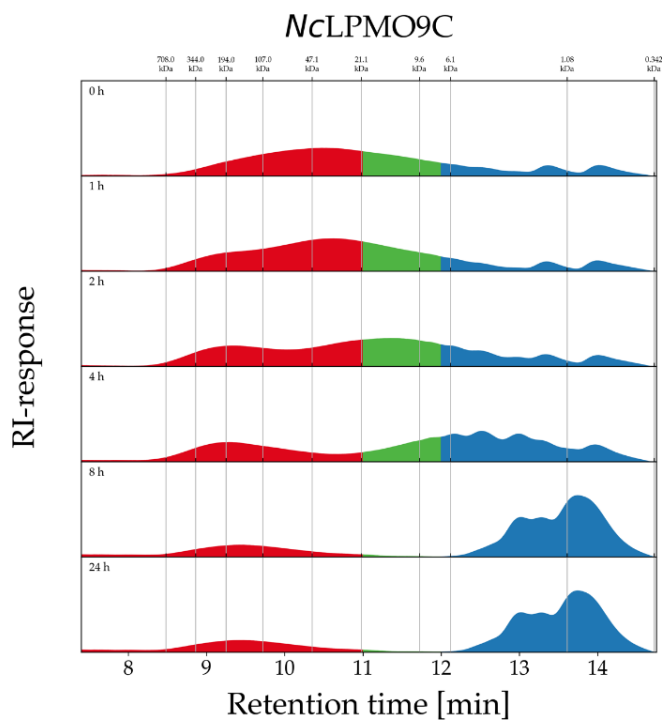


Fig. S28.

HPSEC chromatogram of products released from the reaction of 2 g L^{-1} XG with $1.25 \text{ } \mu\text{M}$ NcLPMO9C in the presence of 1 mM ascorbic acid after 24 h of incubation at $30 \text{ } ^\circ\text{C}$ in 50 mM ammonium acetate buffer, $\text{pH } 5.5$. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2^{nd} degree and a tolerance of 10^{-14} (12).

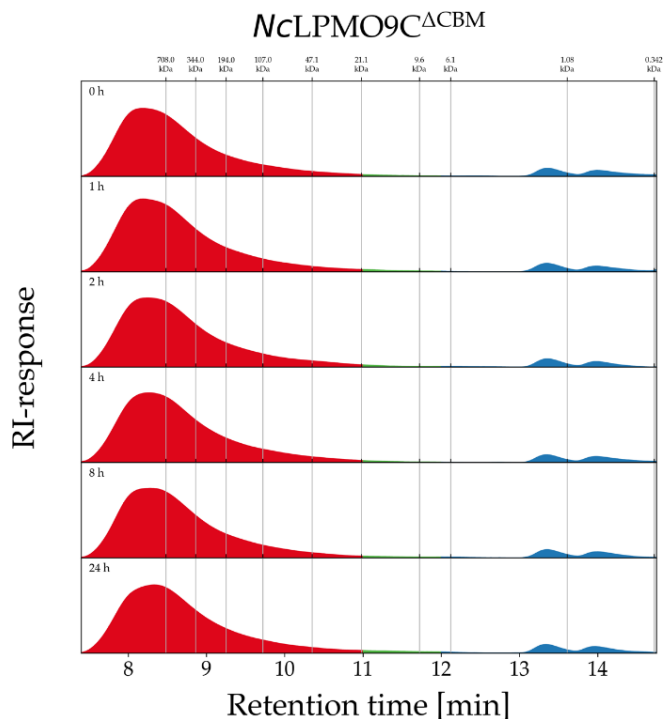


Fig. S29.

HPSEC chromatogram of products released from the reaction of 2 g L⁻¹ XG with 1.25 μM NcLPMO9C^{ΔCBM} in the absence ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2nd degree and a tolerance of 10⁻¹⁴ (12).

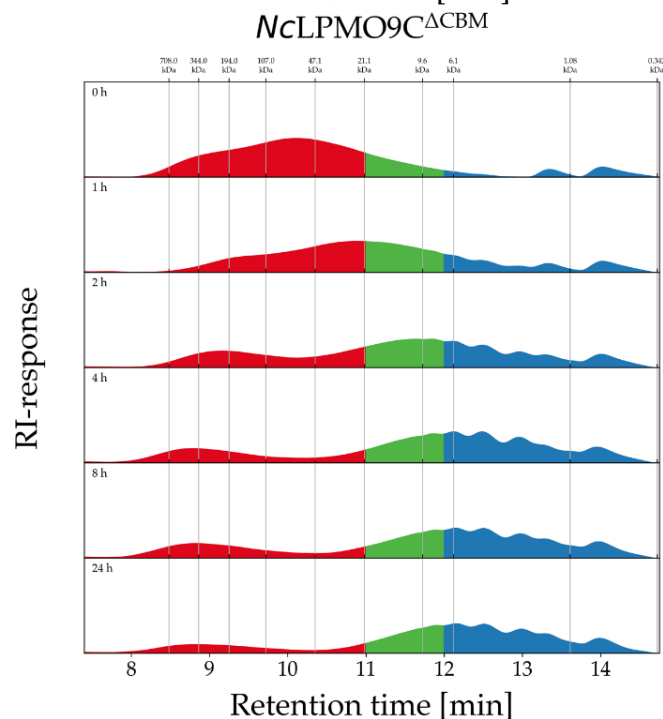


Fig. S30.

HPSEC chromatogram of products released from the reaction of 2 g L⁻¹ XG with 1.25 μM NcLPMO9C^{ΔCBM} in the presence of 1 mM ascorbic acid after 24 h of incubation at 30 °C in 50 mM ammonium acetate buffer, pH 5.5. The data are baseline corrected using a baseline detection algorithm with a polynomial of the 2nd degree and a tolerance of 10⁻¹⁴ (12).

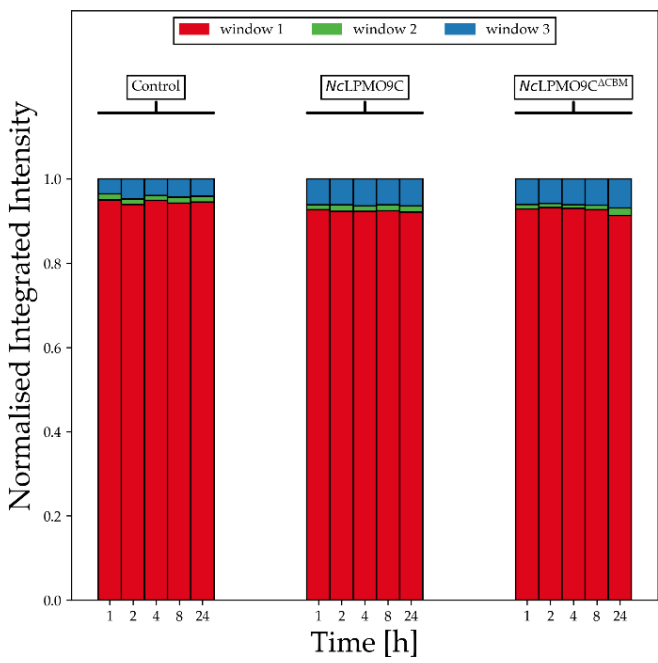


Fig. S31.

Normalized integrated intensity of negative controls without ascorbic acid. The area under the curves from Fig. S27–S29 were integrated. The integrated peaks within the elution windows as described in Table S4 are showing no hydrolytic background of the LPMO preparations.

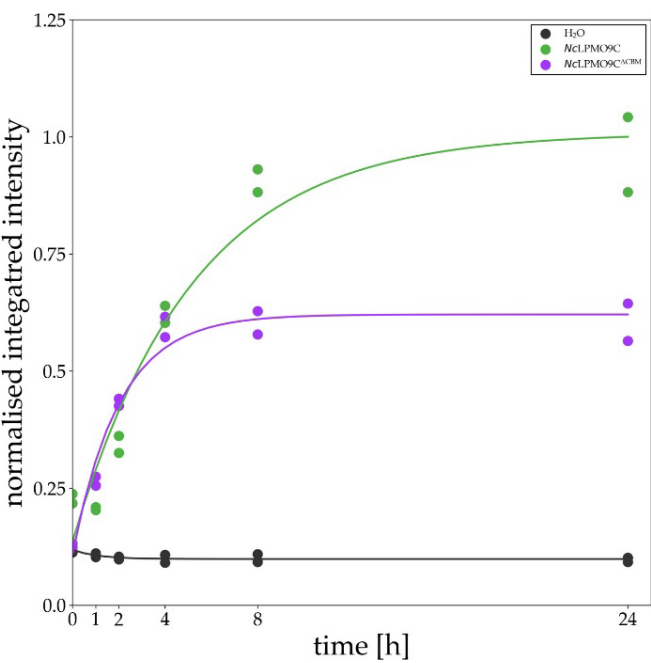


Fig. S32.

Normalized, integrated Elution window 3 over the conversion time. Curves are normalized to the maximum of the NcLPMO9C curve. The curves are fitted to the data points using the following function: $y = ae^{-bx} + c$.

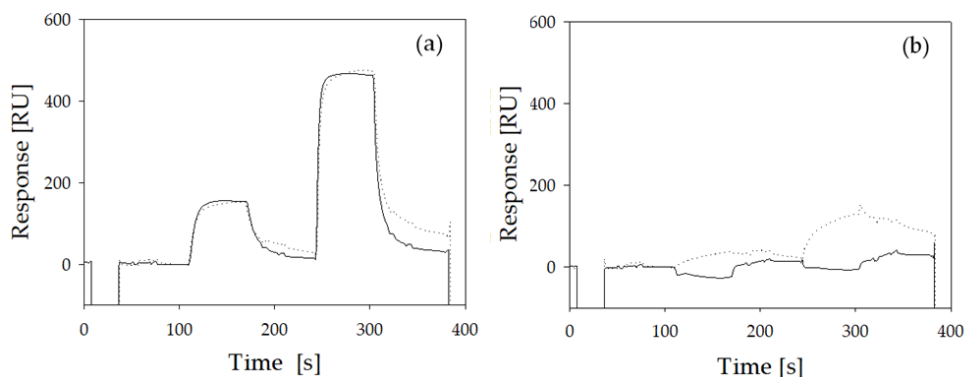


Fig. S33. Sensograms of LPMO binding to cellulose surface in its oxidized and reduced state. The binding behavior of *NcLPMO9C* (a) and *NcLPMO9C^{ACBM}* (b) injected in two different concentrations (400 nM and 1600 nM) to the cellulosic substrate was recorded in the presence of oxygen (black line) and in the absence of oxygen supplemented with 10 mM gallic acid as reductant (dotted line). Sensograms were adjusted to the same response at the begin of the first injection.

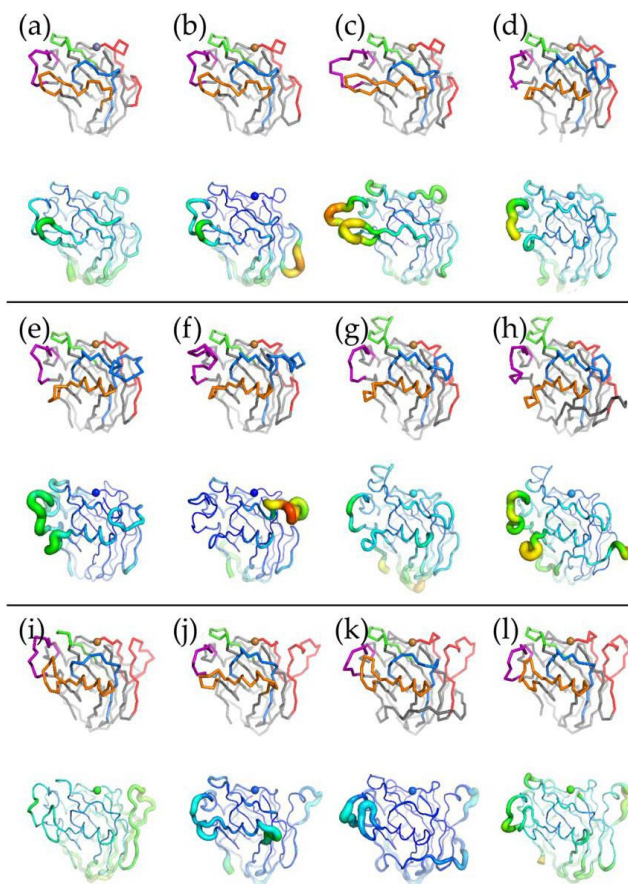


Fig. S34. Representation of the mobility of the amino acid segment around the LPMO catalytic sites. Crystal structures of (a) *TtLPMO9E* (GeneBank: CP_003657366.1; PDB: 3EII), (b) *NcLPMO9F* (GeneBank: CAD70347.1; PDB: 4QI8), (c) *PcLPMO9D* (GeneBank: BAL43430.1; PDB: 4B5Q), (d) *NcLPMO9A* (GeneBank: EAA30263.1; PDB: 5FOH), (e) *NcLPMO9D* (GeneBank: CAD21296.1; PDB: 4EIR), (f) *NcLPMO9C* (GeneBank: EAA36362.1; PDB: 4D7U), (g) *CvAA9A* (GeneBank: AST24379.1; PDB: 5NLT), (h) *LsAA9A* (GeneBank: ALN96977.1; PDB: 5ACF), (i) *HiLPMO9B* (GeneBank: ETW87087.1; PDB: 5NNS), (j) *NcLPMO9M* (GeneBank: EAA33178.2; PDB: 4EIS), (k) *TrLPMO9A* (GeneBank: CAA71999.1; PDB: 5O2W), (l) *TaLPMO9A* (GeneBank: ACS05720.1; PDB: 2YET). Top row: ribbon in gray and segments 1–5 in red, blue, orange, green, and magenta, respectively. Bottom row: flexibility indicated via “b-factor putty”, that is, from blue via green to red and from narrow to wide tubes corresponds to low to high flexibility of the respective regions.

Chapter 5

Table S1. Definitions of segments around the catalytic site of NcLPMOs. The segment definitions for NcLPMO9C^{ΔCBM} and NcLPMO9C^{ΔSeg2}, ΔCBM are identical to NcLPMO9C and NcLPMO9^{ΔSeg2}, respectively.

Enzymes	Segments				
	Seg1	Seg2	Seg3	Seg4	Seg5
NcLPMO9F	A15-G30	I62-P75	P107-F123	H146-Y157	I192-Y204
NcLPMO9C	G15-N27	V58-P86	L115-S132	H155-Y166	I199-Y217
NcLPMO9M	Q12-G40	V71-P85	R118-N136	S161-Y171	V205-Y215
NcLPMO9 ^{ΔSeg2}	G15-N27	V58-P71	L100-S117	H140-Y151	I184-Y202

Table S2. Purification scheme of recombinantly produced LPMOs. Culture supernatants of *P. pastoris* fermentations have been purified by hydrophobic interaction chromatography (HIC), anion exchange chromatography (AEC), and size exclusion chromatography (SEC). The total protein concentration was determined with the Bradford method except for the pools indicated with an "*", which were measured with the bicinchoninic acid (BCA) assay. To the previously published purification protocol of NcLPMO9C (13) a SEC step was added. NcLPMO9F was produced by Kittl et al. (13).

Purification step	Total protein [g]	Yield [%]
<i>NcLPMO9^{ΔSeg2}</i>		
Culture supernatant	6.02	100
HIC	0.84	14
AEC	0.80	13
<i>NcLPMO9M</i>		
Culture supernatant	4.48	100
HIC	0.49	11
AEC	0.32	7
<i>NcLPMO9C^{ΔCBM}</i>		
Culture supernatant	6.72	100
HIC	3.70	55
AEC	1.04	15
SEC*	0.28	4
<i>NcLPMO9C^{ΔSeg2}, ΔCBM</i>		
Culture supernatant	3.84	100
HIC	0.11	2.9
AEC	0.10	2.5
<i>NcLPMO9C</i>		
Stock*	0.13	100
SEC*	0.01	8
<i>NcLPMO9F</i>		
Stock*	0.005	

Table S3. Elution windows defined to integrate detected products in HPAEC chromatograms used for RAC and CMC conversion experiments. Values are in minutes. DP = Degree of polymerization (e.g. cellobiose id DP2).

DP2	DP3	DP4	DP5	DP6	Total
[10.40, 11.00]	[13.60, 14.25]	[16.00, 17.00]	[18.50, 19.40]	[20.50, 21.36]	[4.90, 45.00]

Table S4. Quantification of cello-oligosaccharides in the supernatant released by 1.25 μM enzymes after a 24 h incubation on 2 mg mL⁻¹ RAC or CMC in the presence of 1 mM ascorbic acid. The integration times of product elution windows are defined as described in Table S3.

Enzyme	DP2	DP3	DP4	DP5	DP6	Total
RAC + ascorbic acid						
Control	6.6	0.0	0.0	0.0	0.0	6.6
NcLPMO9F	26.2	10.7	10.3	17.2	12.8	77.2
NcLPMO9C	88.2	77.9	26.4	4.2	5.2	201.9
NcLPMO9C ΔCBM	20.4	8.6	9.0	5.1	3.6	46.7
NcLPMO9 ΔSeg2	12.8	0.0	0.0	1.5	0.0	14.4
NcLPMO9C ΔSeg2 , ΔCBM	15.8	0.0	0.0	1.0	0.0	16.8
NcLPMO9M	40.3	50.9	34.8	25.4	19.4	170.8
CMC + ascorbic acid						
Control	14.8	0.0	0.0	1.0	0.0	15.8
NcLPMO9F	13.0	2.8	2.3	2.7	5.6	26.5
NcLPMO9C	61.9	25.2	0.0	0.0	0.0	87.1
NcLPMO9C ΔCBM	24.8	12.0	7.4	5.2	3.5	52.8
NcLPMO9 ΔSeg2	16.0	0.0	0.0	2.1	0.0	18.1
NcLPMO9C ΔSeg2 , ΔCBM	15.7	0.0	0.0	2.1	0.0	17.8
NcLPMO9M	17.8	2.1	3.2	3.2	2.0	28.4

Chapter 5

Table S5. Quantification of cello-oligosaccharides in the supernatant released by 1.25 μM enzymes after a 24 h incubation on 2 mg mL^{-1} RAC or CMC in the absence of 1 mM ascorbic acid. The integration times of product elution windows are defined as described in Table S3.

Enzyme	DP2	DP3	DP4	DP5	DP6	Total
RAC, no ascorbic acid						
Control	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9F</i>	1.5	0.7	0.9	0.4	0.1	3.6
<i>NcLPMO9C</i>	6.7	5.2	6.3	1.7	0.2	20.1
<i>NcLPMO9C</i> ΔCBM	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9</i> ΔSeg2	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9C</i> ΔSeg2 , ΔCBM	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9M</i>	0.0	0.0	0.0	0.0	0.0	0.0
CMC, no ascorbic acid						
Control	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9F</i>	11.6	4.1	7.5	3.6	1.5	28.3
<i>NcLPMO9C</i>	22.9	10.9	14.8	5.0	0.7	54.3
<i>NcLPMO9C</i> ΔCBM	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9</i> ΔSeg2	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9C</i> ΔSeg2 , ΔCBM	0.0	0.0	0.0	0.0	0.0	0.0
<i>NcLPMO9M</i>	0.0	0.0	0.0	0.0	0.0	0.0

Table S6. Elution windows used to integrate XG conversion products in HPSEC chromatograms. Values shown are in min. The upper boundaries were excluded from the integrations as indicated by the bracket "[)".

Elution window 1	Elution window 2	Elution window 3
[7.40, 11.00]	[11.00, 12.00]	[12.00, 14.75]

Table S7. Nucleotide sequences of primer pair used to generate *NcLPMO9C* ΔCBM and *NcLPMO9C* ΔSeg2 , ΔCBM .

Primer name	Sequence (5'→3')
P213X_F	GTCCAGCTTTGTTCACTTGT <u>TA</u> AGCTGGTGGTTC
P213X_R	CAAAGCTGGACCTGGAATTGGTATGGCTTAC

Table S8. Statistics of the peptide mapping with liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS). Results with 1 unique peptide have been removed from the table. Two human proteins have been removed. Rows that were showing the used protease were also deleted from the table. The likelihood that a peptide is a random match is indicated the log(e) score. E.g., the log(e) of the first row corresponds a 1 in 10419.3 chance that this result is a false positive. The coverage indicates sequence coverage.

	Coverage (%)	Log (e) score	Unique peptides	UniProt ID
NcLPMO9C	53	-419.3	34	NcLPMO9C
NcLPMO9C ^{ΔCBM}	88	-622.4	50	NcLPMO9C ^{ΔCBM}
NcLPMO9C ^{ΔSeg2}	72	-668.6	53	NcLPMO9C ^{ΔSeg2}
	45	-63.6	6	C4R8X7 (<i>P. pastoris</i>)
	62	-570.3	47	NcLPMO9C ^{ΔSeg2} , ^{ΔCBM}
	60	-149.4	13	C4R8X7 (<i>P. pastoris</i>)
	48	-72.4	7	C4R300 (<i>P. pastoris</i>)
	36	-72.3	8	A0A1B2JFA7 (<i>P. pastoris</i>)
NcLPMO9C ^{ΔSeg2} , ^{ΔCBM}	33	-214.1	19	C4R9F6 (<i>P. pastoris</i>)
	25	-93.5	10	C4QW09 (<i>P. pastoris</i>)
	23	-68.6	8	C4R862 (<i>P. pastoris</i>)
	22	-70.9	7	A0A1B2J744 (<i>P. pastoris</i>)
	21	-65.4	7	C4QY07 (<i>P. pastoris</i>)
	21	-63.3	7	C4R7U0 (<i>P. pastoris</i>)
	71	-244.7	27	NcLPMO9F
	71	-62	7	A0A1B2J759 (<i>P. pastoris</i>)
	68	-359.3	33	A0A1B2JHZ1 (<i>P. pastoris</i>)
	39	-163.5	18	A0A1B2JGG7 (<i>P. pastoris</i>)
	39	-41.4	6	A0A1B2JCS3 (<i>P. pastoris</i>)
NcLPMO9F	38	-67.5	11	A0A1B2JEK0 (<i>P. pastoris</i>)
	28	-44.2	6	A0A1B2JGF8 (<i>P. pastoris</i>)
	17	-40.2	6	A0A1B2JIZ2 (<i>P. pastoris</i>)
	17	-38.8	6	A0A1B2J5L5 (<i>P. pastoris</i>)
	10	-54.8	6	A0A1B2J5W9 (<i>P. pastoris</i>)
	8.8	-35.8	5	A0A1B2J841 (<i>P. pastoris</i>)
NcLPMO9M	84	-369	32	NcLPMO9M

7.1. Supplementary references

1. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol.* 2018;9:1080.
2. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983;22:2577–2637.
3. Eichenberger AP, Allison JR, Dolenc J, Geerke DP, Horta BAC, Meier K, Oostenbrink C, Schmid N, Steiner D, Wang D, et al. GROMOS++ software for the analysis of biomolecular simulation trajectories. *J Chem Theory Comput.* 2011;7:3379–3390.
4. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep.* 2017;7:10480.
5. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27:343–350.
6. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis.* 2009;30:S162–S173.
7. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* 2017;45:D313–D319.
8. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46:W296–W303.
9. Breslmayr E, Hanzek M, Hanrahan A, Leitner C, Kittl R, Santek B, et al. A fast and sensitive activity assay for lytic polysaccharide monooxygenase. *Biotechnol Biofuels.* 2018;11:79.
10. Durowoju IB, Bhandal KS, Hu J, Carpick B, Kirkitadze M. Differential scanning calorimetry—a method for assessing the thermal stability and conformation of protein antigen. *J Vis Exp.* 2017;121:e55262.
11. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee Y-H, Goto Y, Réfrégiers M, Kardos J. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* 2018;46:W315–W322.
12. Negri, L.H.; Vestri, C. peakutils: v1.1.0 2017.
13. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels.* 2012;5:79.

Chapter 6

Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation

Peicheng Sun^{1,†}, Christophe V.F.P. Laurent^{2,3,†}, Stefan Scheiblbrandner², Matthias Frommhagen¹, Dimitrios Kouzounis¹, Mark G. Sanders¹, Willem J.H. van Berkel¹, Roland Ludwig² and Mirjam A. Kabel¹

¹Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

²Biocatalysis and Biosensing Laboratory, Department of Food Science and Technology, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

³Institute of Molecular Modeling and Simulation, Department of Material Sciences and Process Engineering, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

[†]These authors contributed equally to this work

Abstract

Lytic polysaccharide monooxygenases (LPMOs) are powerful enzymes that oxidatively cleave plant cell wall polysaccharides. LPMOs classified as fungal Auxiliary Activities family 9 (AA9) have been mainly studied for their activity towards cellulose, however, various members of this AA9 family have been also shown to oxidatively cleave hemicelluloses, in particularly xyloglucan (XG). So far, it has not been studied in detail how various AA9 LPMOs act in XG degradation, and in particular, how the mode-of-action relates to the structural configuration of these LPMOs. Two *Neurospora crassa* (Nc) LPMOs were found to represent different mode-of-action towards XG. Interestingly, the configuration of active site segments of these LPMOs differed as well, with a shorter Segment 1 (\cdot Seg1) and a longer Segment 2 ($^{+}$ Seg2) present in NcLPMO9C and the opposite for NcLPMO9M ($^{+}$ Seg1 \cdot Seg2). We confirmed that NcLPMO9C cleaved the non-reducing end of unbranched glucosyl residues within XG via the oxidation of the C4-carbon. In contrast, we found that the oxidative cleavage of the XG backbone by NcLPMO9M occurred next to both unbranched and substituted glucosyl residues. The latter are decorated with xylosyl, xylosyl-galactosyl and xylosyl-galactosyl-fucosyl units. The relationship between active site segments and the mode-of-action of these NcLPMOs was rationalized by a structure-based phylogenetic analysis of fungal AA9 LPMOs. LPMOs with a \cdot Seg1 $^{+}$ Seg2 configuration clustered together and appear to have a similar XG substitution-intolerant cleavage pattern. LPMOs with the $^{+}$ Seg1 \cdot Seg2 configuration also clustered together and are reported to display a XG substitution-tolerant cleavage pattern. A third cluster contained LPMOs with a \cdot Seg1 \cdot Seg2 configuration and no oxidative XG activity. The detailed characterization of XG degradation products released by LPMOs reveal a correlation between the configuration of active site segments and mode-of-action of LPMOs. In particular, oxidative XG-active LPMOs, which are tolerant and intolerant to XG substitutions are structurally and phylogenetically distinguished from XG-inactive LPMOs. This study contributes to a better understanding of the structure-function relationship of AA9 LPMOs.

Keywords











Plant cell wall; Lignocellulose; Biomass; Biorefinery; Hemicellulose; Xyloglucan; AA9 LPMO; *Neurospora crassa*; Active site segments; Phylogenetic tree

1. Introduction

For establishing sustainable processes and a circular economy, plant biomass is an essential source for the production of fuels and chemicals, in particular, to replace fossil-based resources (1). Plant biomass dry matter is mainly composed of plant cell wall polymers, which are present in the middle lamella, primary and secondary cell wall layers (2). The primary cell wall is mainly built of pectin, cellulose and hemicellulosic xyloglucan (XG), while the secondary cell wall is mostly composed of cellulose, hemicellulosic xylan or mannan and the aromatic polymer lignin (3-6). While in general dicotyledonous plant biomass dry matter is majorly composed of primary cell wall components, other species such as grasses and wood plant biomass dry matter are majorly composed of secondary cell wall components (3-6). An important step in biomass-based processes is the release of fermentable carbohydrates. In the last decade, monocopper-dependent lytic polysaccharide monooxygenases (LPMOs) have been shown to assist glycosyl hydrolases as green and effective tools for biomass polysaccharide degradation (7-10). In this research, we aimed to understand how LPMOs oxidatively cleave XG, and hypothesized that the mode-of-action of LPMOs towards XG correlates with their active site configuration.

Cellulose is a linear macromolecule composed of β -(1 \rightarrow 4) linked D-Glcp units (4). Due to the absence of side chains, cellulose forms crystalline (ordered) microfibrils via van der Waals force and hydrogen bonds (2, 5). Like cellulose, XG has a backbone of β -(1 \rightarrow 4) linked D-Glcp residues, which is further substituted via the C6 position by α -(1 \rightarrow 6) linked D-Xylp residues (2, 11). The xylosyl units can be further substituted with β -(1 \rightarrow 2) linked D-Galp and, rarely, α -(1 \rightarrow 3) linked L-Araf residues. In addition, an L-Fucp unit via an α -(1 \rightarrow 2) linkage to D-Galp also occurs (11-14). The type and amount of XG substituents highly vary and, for example, depends on the plant species or tissue (13). To simplify the complicated trivial names of individual XG oligosaccharides, Fry and coworkers developed an unambiguous nomenclature using one-letter codes to represent the XG structure (**Table 1**) (15). Most XG structures have been defined as block-wise "XXXG"- and "XXGG"-types (16). For example, XG from tamarind seed (TXG) and black currant (BCXG) have been shown to consist of "XXXG" repeating units, with partially β -D-Galp substituted blocks (XXLG, XLXG and XLLG) (15, 17). Furthermore, additional substitutions with fucosyl units were defined for BCXG (XXFG and XLFG) (17). Fucosylated XG has been found in many plant sources from campanulids (i.e., carrot), while XG from grass species is not composed of fucosylated residues (11). Other modifications have been found as well, for instance acetylation on galactosyl residues (13, 18), but are not further discussed in this study.

Table 1. One-letter codes according to Fry et al. (15) and symbols used in this study for representing XG oligosaccharide structures

Structure represented	Symbols ^a	One-letter code
-β-D-Glcp-		G
-β-D-Glcp- α-D-Xylp-(1→6) ↓	 	X
-β-D-Glcp- β-D-Galp-(1→2)-α-D-Xylp-(1→6) ↓	  	L
-β-D-Glcp- α-L-Fucp-(1→2)-β-D-Galp-(1→2)-α-D-Xylp-(1→6) ↓	   	F

^ablue circle: glucosyl unit; star: xylosyl unit; yellow circle: galactosyl unit; red triangle: fucosyl unit

LPMOs classified in the Carbohydrate-Active enZymes database (CAZy; www.cazy.org (19)) Auxiliary Activities family 9 (AA9) originate from fungi and have been shown to oxidatively cleave cellulose. Although less studied, for some AA9 members additional oxidative XG cleavage has been reported (20–32). So far, oxidative XG cleavage of AA9 LPMOs has been mainly determined by the detection of formed oligosaccharides by using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) and matrix assisted laser desorption/ionization–time of flight–mass spectrometry (MALDI-TOF-MS). Although these analytical techniques indicate an oxidative cleavage of XG, exact structures of released XG degradation products by AA9 LPMOs were not identified (20–32). Nevertheless, from both studies it could be concluded that among the XG-active AA9 LPMOs generally two HPAEC-PAD patterns of TXG oligosaccharides have been shown: i) generation of (block-wise) “XXXG”-type (oxidized) TXG oligosaccharides, e.g., *Mt*LPMO9J and *Nc*LPMO9C (22, 32); ii) generation of many different, non-“XXXG”-type (oxidized) TXG oligosaccharides, e.g., *Fg*LPMO9A and *Gt*LPMO9A-2 (27, 28). To better understand these two suggested routes a detailed product characterization is essential.

Whether these different XG-cleavage pathways result from distinct XG binding sites neighboring the catalytic site of the LPMOs has yet to be defined. Active site structures of LPMOs interacting with cellulosic substrates have already been reported (24, 33–

37), but information about relevant binding sites of XG is still scarcely available in the literature. Courtade and coworkers have shown through NMR analysis that a so-called L3 loop around the active site of a XG-active *NcLPMO9C* strongly interacted with XG (38). This L3 loop also has been shown to be present in other XG-active LPMOs like *PaLPMO9H* (31) and *MtLPMO9J* (22). However, in another XG-active *GtLPMO9A-2*, the L3 loop is absent. Instead, *GtLPMO9A-2* has an extended L2 loop (28). This difference might indicate that the configuration of segments around the AA9 LPMO active site influences their catalytic behavior on XG. The definition of the loops (L2, L3, LS and LC) around the active site has previously been suggested (37, 39, 40), and further redefined as segments in our previous study due to the presence of secondary structure elements (41). Briefly, five segments (Seg1–Seg5) were defined, of which Seg1, Seg2, Seg3 and Seg5 are comparable, but slightly different, to the previously defined L2, L3, LS and LC regions, respectively (see also **Fig. 1**). Seg4 was newly defined and has not been described before.

In this work, two distinct product profiles of two different XG-active LPMOs from *Neurospora crassa* (*NcLPMO9C* and *NcLPMO9M*) were characterized by identification of the formed non-oxidized and oxidized XG oligosaccharides. In addition to various other chromatographic techniques, hydrophilic interaction chromatography coupled with electrospray ionization–collision induced dissociation–mass spectrometry (HILIC-ESI-CID-MS/MS²) was used. To test our hypothesis that the mode-of-action of LPMOs towards XG is a result of their specific structural configuration around the active site, a structure-based sequence analysis of AA9 LPMOs was performed. The resulting phylogenetic tree shows three distinct groups, which not only differ in structural active site segments, but also seemingly correlate to the oxidative XG cleavage being either tolerant or intolerant to substitutions, and to XG-inactive LPMOs.

2. Materials and methods

2.1. XG substrates, carbohydrate standards and other chemicals

XG from tamarind (*Tamarindus indica*, TXG) seed, TXG oligosaccharide standards (xyloglucan hepta + octa + nona saccharides) and XEG (GH5) from *Paenibacillus* sp. were purchased from Megazyme (Bray, Ireland). XG from black currants (*Ribes nigrum* L., BCXG) was available in our laboratory (fraction CASS) extracted by Hilz and coworkers (17). Glucose was purchased from Sigma-Aldrich (St. Louis, Missouri, USA) and Asc was purchased from VWR International (Radnor, PA, USA). Cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose were used as standards and purchased from Megazyme. Water used in all experiments was generated by a Milli-Q system (Millipore, Molsheim, France), unless mentioned otherwise.

2.2. Catalytic performance of XEG, *NcLPMO9C* and *NcLPMO9M* on XG

Expression, production and purification of *NcLPMO9C* and *NcLPMO9M* were described previously (41). XG substrates (TXG or BCXG, 2 mg/mL) were dissolved in 50 mM ammonium acetate buffer (pH 5.0) with the addition of Asc (1 mM final concentration).

Subsequently, XEG, *NcLPMO9C* and *NcLPMO9M* were added to a concentration of 1.25 μM . Control reactions were performed without the addition of Asc. Single 200 μL reactions were incubated in an Eppendorf ThermoMixer® comfort at 800 rpm (in an almost vertical orientation) and reactions used to produce the time curves were incubated in a head-over-tail rotator at 20 rpm (5 mL total volume). *NcLPMO9C* and *NcLPMO9M* reactions were incubated at 30°C while XEG reaction was at 50°C. All reactions were performed in duplicate. To create a time curve for *NcLPMO9C* and *NcLPMO9M*, a larger reaction volume of 500 μL was sampled at 0, 1, 2, 4, 8 and 24 h after enzyme addition. The reactions were stopped while incubating for 10 min at 97 °C in an Eppendorf ThermoMixer® C. Subsequently, the supernatant was recovered after centrifugation in a Hermle Z 233 MK-2 centrifuge at $22000 \times g$ (Rotor: 220.87 VO5/6) for 20 min and stored at -20 °C until further usage. Parts of XEG- and *NcLPMO9C*-TXG-digests were further treated with β -galactosidase (GH35 from *Aspergillus niger*, Megazyme), which is further described in Supplementary information.

2.3. Analytical methods

2.3.1. HPSEC analysis for molecular weight distribution of (degraded) TXG

TXG and corresponding digests were analyzed by high performance size exclusion chromatography coupled to a refractive index detector (HPSEC-RI) for their molecular weight distribution. Instrument settings, column and elution program were the same as described previously (41). Pullulans (Associated Polymer Labs Inc., New York, USA) in the MW range of 0.4–708 kDa were used for calibration.

2.3.2. HPAEC analysis for profiling oligosaccharides

TXG and corresponding digests were analyzed by HPAEC-PAD on an ICS5000 (Dionex) system equipped with a CarboPac PA-1 column (2 mm ID \times 250 mm) in combination with a CarboPac PA guard column (2 mm ID \times 50 mm). Mobile phases were (A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH. The column temperature was 20 °C. The elution program applied has been described previously (30). Samples were diluted five times before analysis. Commercial TXG oligosaccharide mixture (50 $\mu\text{g}/\text{mL}$), glucose (2.5 $\mu\text{g}/\text{mL}$) and cello-oligosaccharides (DP2–6, 2.5 $\mu\text{g}/\text{mL}$) were used as standards.

2.3.3. MALDI-TOF-MS analysis of oligosaccharides

To analyze the mass of formed XG oligosaccharides, MALDI-TOF-MS (Bruker Daltonics, Billerica, Massachusetts, USA) was used as previously described (42). The mass spectrometer was calibrated using maltodextrins (Avebe, Veendam, The Netherlands) in a mass range (m/z) of 500–3000 and a total of 300 spectra were collected for each measurement. Prior to analysis, samples were desalted using Dowex AG 50W-X8 Resin (Bio-Rad Laboratories, Hempel Hempstead, UK). The desalted supernatants were dried under nitrogen and re-dissolved in water containing 20 mM LiCl to obtain lithium (Li)-adducts. 1 μL of each lithium-rich sample was mixed with 1 μL matrix solution (50%

(v/v) acetonitrile in H₂O containing 12 mg/mL 2,5-dihydroxy-benzoic acid (Bruker Daltonics)) and dried under nitrogen.

2.3.4. HILIC-ESI-CID-MS/MS² for structural elucidation of (degraded) XG

The LPMO-TXG- and -BCXG-digests were separated and analyzed using HILIC coupled to ESI-MS. To separate the TXG oligosaccharides, a Vanquish UHPLC system (Thermo Fisher Scientific, San Jose, CA, USA) equipped with an Acquity UPLC BEH Amide column (1.7 μ m, 2.1 mm ID \times 150 mm) and a VanGuard pre-column (1.7 μ m, 2.1 mm ID \times 5 mm) was used. Supernatants from LPMO-TXG- and LPMO-BCXG-digests were concentrated five times and then subjected (2 μ L) to the column. The column temperature was set at 35°C using the still air mode and the flow rate was 0.45 mL/min. Water (A) and acetonitrile (B) both containing 0.1% formic acid (all were UHPLC-grade; Biosolve, Valkenswaard, The Netherlands) were used as mobile phases. The elution profile was: 0–2 min at 82% B (isocratic), 2–62 min from 82% to 60% B (linear gradient), 62–62.5 min from 60% to 42% B (linear gradient), 62.5–69 min at 42% B (isocratic), 69–70 min from 42% to 82% B (linear gradient) and 70–80 min at 82% B (isocratic). The MS settings have been described previously (43). The full MS (m/z) range was set to 300–2000. To improve the fragmentation, MS² was performed using dependent scan followed by a parent mass list. The mass list used is displayed in Table S1. For MS², the CID with a normalized collision energy was set at 35%, the minimum signal threshold was 20000 counts, activation Q was 0.15 and activation time was 10 ms. Mass spectrometric data were processed using Xcalibur 2.2 (Thermo Fisher Scientific).

2.3.5. Crystal structures and homology models

Structural data of LPMOs were derived from the RCSB protein data bank (<https://www.rcsb.org>). Homology models of LPMOs without published three-dimensional structures were generated using SWISS-MODEL (<https://swissmodel.expasy.org>) (44–48). Template search with BLAST (49) and HHBlites (50) were performed against the SWISS-MODEL template library (SMTL). The target sequences were searched with BLAST against the primary amino acid sequence contained in the SMTL. The PyMOL Molecular Graphics System (Version 1.7.2.1 Schrödinger, LLC) was used for visualization and structural alignments.

2.3.6. Sequence mining, structure-based multiple sequence alignment and phylogenetic analysis

In order to obtain an unbiased set of amino acid sequences, which covers the whole range of the large variety within AA9 LPMOs, sequences were selected randomly from the 498 available eukaryotic AA9 LPMO sequences in the CAZy database. This set was completed by addition of all AA9 LPMO sequences labeled as “characterized” in the CAZy database, all AA9 LPMO sequences with a resolved structure, and those with known XG (in)activity, if not already present in the set. The amino acid sequences were aligned (full-length; FL) using the MUSCLE algorithm (51) in MEGA7 (52) and

fine-tuned by cutting out the signal peptide, the linker- and the CBM-region, as well as sequences not fitting to the alignment. The amino acid sequences were then realigned using the structure-based MAFFT-DASH algorithm (53). The resulting structure-based alignment was then cut down to the regions of interest termed "Segments 1 to 5" (Seg1–Seg5) (segment-only; SO).

Phylogenetic analysis of both the FL and SO structure-based multiple sequence alignment was done using RAXML-NG (54). Firstly, the alignments were tested for the most applicable substitution model using ModelTest-NG (55). The tree was inferred using the BLOSUM62 model (56) (number of discrete gamma categories:4; with frequencies and invariant sites) for the FL alignment, and the Probability Matrix from Blocks (PMB) (57) model (number of discrete gamma categories:4; with frequencies and invariant sites) for the SO alignment and 20 starting trees were calculated. Bootstrap analysis was then carried out until convergence criteria (cut-off: 0.03) based on the bootstopping test (58) were reached (800 and 1120 bootstraps for the FL and SO alignment, respectively). The resulting phylogenetic trees were prepared for publication using MEGA7.

3. Results

3.1. *NcLPMO9C* and *NcLPMO9M* and their oxidative XG cleavage patterns

Two LPMOs from *N. crassa* with different active site segment configuration (*NcLPMO9C* and *NcLPMO9M*; **Fig. 1**) were tested for their mode-of-action towards XG. As presented in **Fig. 1**, *NcLPMO9C* holds a short (–) Seg1 and a long (+) Seg2, whereas *NcLPMO9M* has a +Seg1–Seg2 configuration (41).

We first monitored the mode-of-action of two *NcLPMOs* on TXG by profiling the molecular weight (MW) distribution of *NcLPMO9M*- and *NcLPMO9C*-TXG-digests during incubation using HPSEC-RI (**Fig. 2** and Fig. S1). The MW distribution of both *NcLPMO*-TXG-digests after 24 h incubation showed only little change in the absence of ascorbic acid (Asc) (Fig. S1), which showed that these enzyme preparations were almost free of hydrolytic side activities. However, upon addition of Asc an autooxidation of the TXG could be observed, resulting in a visible decrease in the MW distribution after 24 h (Fig. S1). Therefore, the MW distributions of the *NcLPMO* digests (with Asc) were compared to the ones of TXG without enzyme but with Asc (24 h; **Fig. 2**).

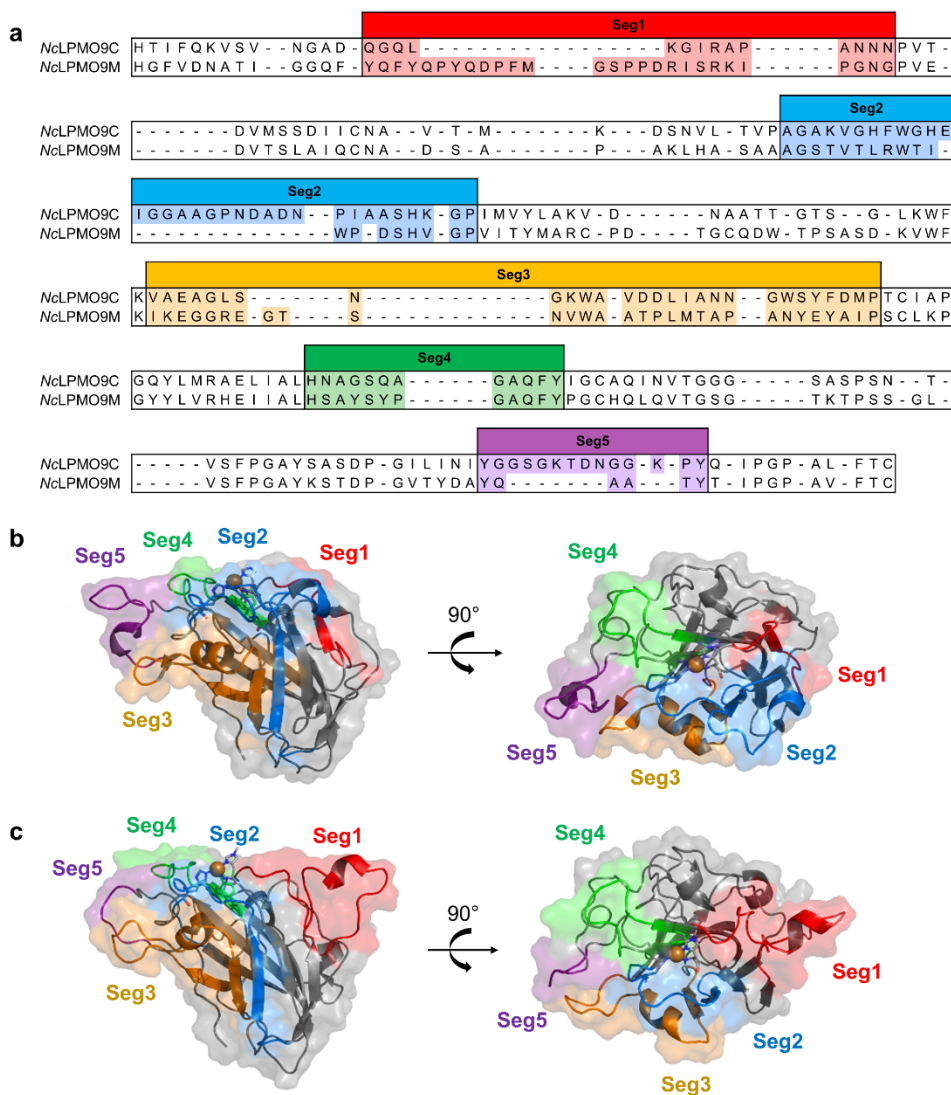


Fig. 1. (a) Amino acid sequence alignments of NcLPMO9C (PDB entry: 4D7U) and NcLPMO9M (PDB entry: 4EIS) and (b and c) their respective crystal structures. The active site segments are indicated as Seg1 (red), Seg2 (blue), Seg3 (yellow), Seg4 (green) and Seg5 (purple).

Already after 2 h, the products formed by NcLPMO9M had a lower MW-range compared to NcLPMO9C indicating that both LPMOs show distinct mode-of-action on TXG (**Fig. 2**). To be more precise, NcLPMO9M formed two rather broad populations (**Fig. 2a**), one ranging from 30–200 kDa and another ranging from 1–30 kDa, while NcLPMO9C formed two larger MW populations (80–700 kDa and 1–80 kDa, **Fig. 2b**). Notably, oxidative XG cleavage of NcLPMO9M has not been reported previously and neither have the MW distributions of XG digests of these NcLPMOs. Further, seen from the MW profiles (**Fig. 2a**), after 8 h the NcLPMO9M TXG degradation was complete, no high

MW population (30–200 kDa) of XG remained, and final products ranged from 0.4 to 3 kDa (**Fig. 2a**). In contrast, for the *NcLPMO9C*-TXG-digest the high MW XG population (80–700 kDa) remained and a decrease in MW of the products was observed even between 8 h and 24 h of incubation (**Fig. 2b**). The final digest was composed of products ranging from 0.4 to 3 kDa and showed a different MW distribution profile as the 24 h *NcLPMO9M*-TXG-digest (**Fig. 2b**).

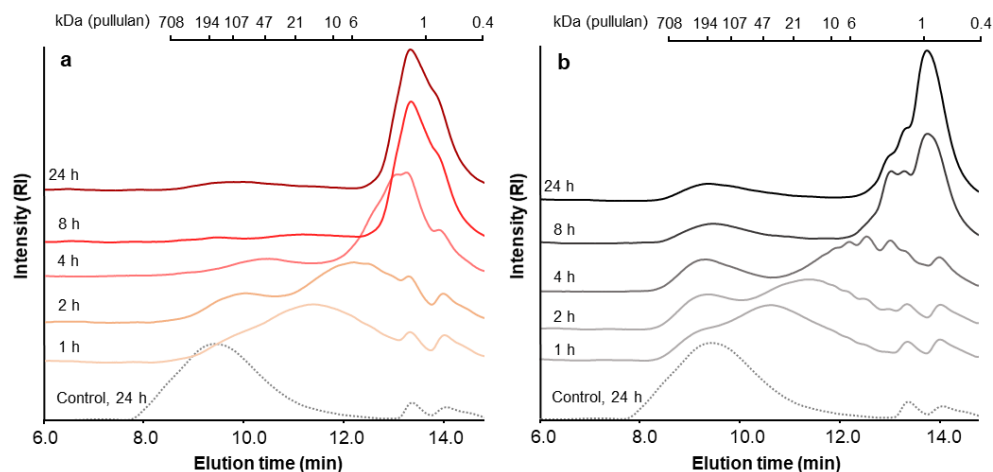


Fig. 2. Molecular weight (MW) distribution of tamarind seed xyloglucan (TXG) digests in presence of ascorbic acid (Asc) from (a) 1.25 μ M *NcLPMO9M* and (b) 1.25 μ M *NcLPMO9C*, analyzed by HPSEC-RI. Dotted lines are control reactions containing only TXG with Asc. MW distributions of additional control digests are shown in Fig. S1. MWs are indicated based on calibration with a series of pullulan standards (0.4–708 kDa).

To learn more about the exact cleavage sites in the TXG for both *NcLPMOs*, the formed TXG oligosaccharides were characterized in detail. First, the digests were analyzed by HPAEC-PAD and the corresponding chromatograms are shown in **Fig. 3**. For comparison, the commercial xyloglucanase (XEG)-TXG-digest (**Fig. 3g**) and commercial non-oxidized TXG oligosaccharide (XXXG, XLXG, XXLXG and XLXG) standards (**Fig. 3h**) were analyzed, of which the annotation of HPAEC-peaks has been well defined in literature (59–61). The control reactions (**Fig. 3b and 3d**) did not show the formation of (detectable) oligosaccharides, which confirms the absence of hydrolytic xyloglucanase (side-)activities. In the presence of Asc, both *NcLPMOs* released noticeably different types of TXG oligosaccharides (**Fig. 3a and 3c**), underlining the differences in the above-described MW distributions (**Fig. 2**). The TXG-digest of *NcLPMO9C* has been described previously and our HPAEC profile corresponds with the published one (32). However, the annotation, of in particular the non-oxidized products, seems to be different compared to the previous research. Based on our results, the common non-oxidized “XXXG”-type products were not present in *NcLPMO9C*-TXG-digest (**Fig. 3a**). Our annotation was based on i) comparison with the XEG-TXG-digest and standards of a mixture of XXXG, XLXG, XXLXG and XLXG (**Fig. 3g and 3h**), and ii) β -galactosidase treatment of the *NcLPMO9C*- and XEG-TXG-digest to confirm that L units were degraded to X units (Fig. S2). Indeed, β -galactosidase

treatment of the XEG-TXG-digest (Fig. S2b) resulted in removal of XLXG, XXLG and XLLG, and only XXXG remained.

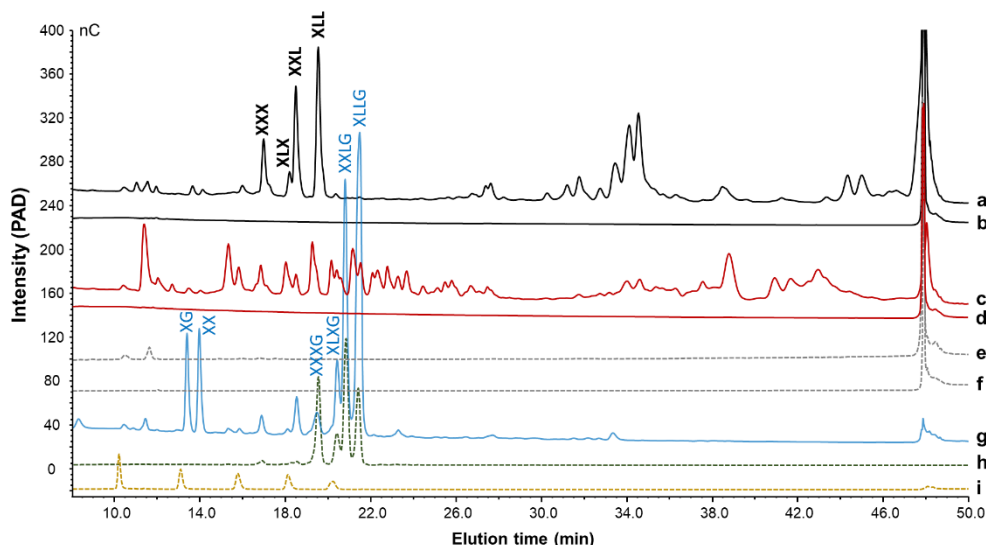


Fig. 3. HPAEC elution patterns of oligosaccharide products after the incubation of tamarind seed xyloglucan (TXG) with *NcLPMO9C* (1.25 μ M; a, b) and *NcLPMO9M* (1.25 μ M; c, d) in the presence (1 mM; a, c) and absence (b, d) of ascorbic acid (Asc). TXG oligosaccharides released by xyloglucanase (XEG, 1.25 μ M) in the presence of Asc (g) was added as the reference. In addition, TXG only (e), TXG with 1 mM Asc (f), TXG oligosaccharide standards (xyloglucan hepta + octa + nona saccharides; h) and a standard (i) containing a mixture of cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose (from left to right in chromatogram) are shown.

In addition, XXG was formed, confirmed by MALDI-TOF-MS (Fig. S3a; m/z 775.3 (lithium (Li)-adduct, $[M+Li]^+$)), due to the presence of isoprimeverase in the commercial β -galactosidase (62, 63), which was further substantiated by the formation of isoprimeverose (X unit) (Fig. S2b). In contrast, β -galactosidase treated *NcLPMO9C*-TXG-digest, majorly resulted in XXX (Fig. S2d), which was confirmed by MALDI-TOF-MS (Fig. S3b; m/z 907.3 ($[M+Li]^+$)), and no other main non-oxidized compounds remained. Again minor isoprimeverase side-activity was seen, resulting in formation of X and XX. The peak representing XXX was also present in the *NcLPMO9C*-TXG-digest, without β -galactosidase treatment, in addition to three peaks now defined as XLX, XXL and XLL. These last three peaks were removed by the β -galactosidase treatment, which confirmed the presence of L unit. It should be noted that in previous research studying LPMO activity towards TXG, the non-oxidized oligosaccharides now annotated as XXX, XLX, XXL and XLL, were incorrectly suggested to be XXXG, in addition to XLXG, XXLG and XLLG (22, 24, 29, 31, 32).

The HPAEC pattern of the *NcLPMO9M*-TXG-digest showed considerably more oligosaccharides peaks compared to the TXG-digest of *NcLPMO9C*. As with HPAEC the type of oligosaccharides (especially the oxidized ones) formed cannot be identified without standards, further characterization of degraded TXG oligosaccharides was carried out by MALDI-TOF-MS and HILIC-ESI-CID-MS/MS².

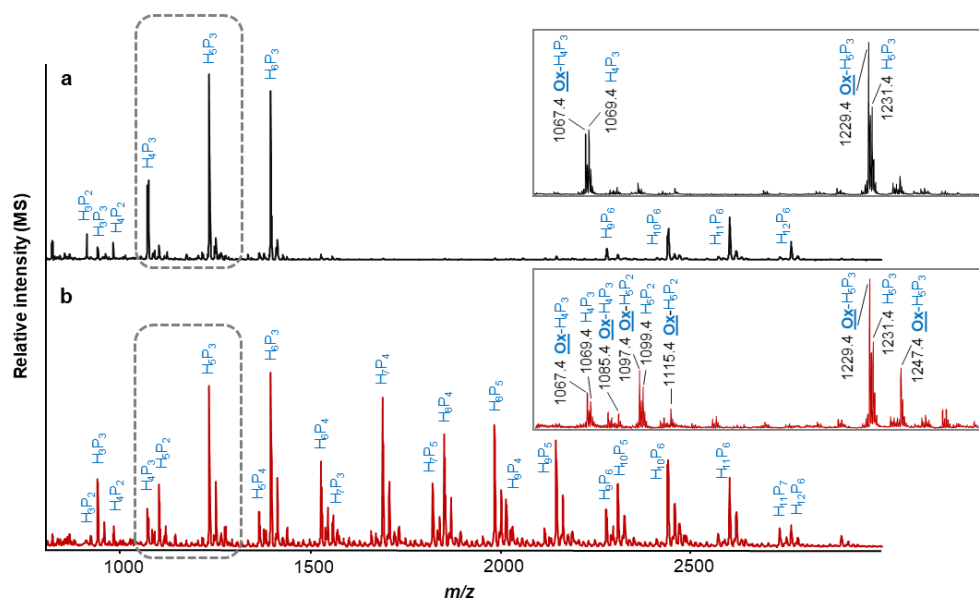


Fig. 4. MALDI-TOF mass spectra of tamarind seed xyloglucan (TXG), after incubation with (a) 1.25 μ M NcLPMO9C and (b) 1.25 μ M NcLPMO9M, both in the presence of 1 mM of ascorbic acid (Asc). TXG oligosaccharide clusters included non-oxidized oligosaccharides (H_nP_m) and C4-oxidized oligosaccharides ($Ox-H_nP_m$). The clusters of H_4P_3 and H_5P_3 are enlarged in the inserts. Abbreviations: H, hexaose (glucose or galactose, 162 Da); P, pentaoxe (xylose, 132 Da); n, number of hexaoses; m, number of pentaoses; Ox, oxidized. M/z shown correspond to lithium (Li)-adducts.

The MALDI-TOF mass spectrum of the NcLPMO9C-TXG-digest (**Fig. 4a**) clearly indicated masses ($[M+Li]^+$) corresponding to blocks of TXG, which has also been shown in the previous research (32, 63). These blocks were present as non-oxidized (i.e., H_4P_3 , H_5P_3 and $H_{10}P_6$) and C4-oxidized oligosaccharides (i.e., $Ox-H_4P_3$ and $Ox-H_5P_3$), where “H” and “P” represented as hexaose and pentaoxe, respectively. Taking the above described HPAEC results into account, it could be concluded that, for instance, H_4P_3 represents XXL/XLX, and H_5P_3 represents XLL. The annotation of C4-oxidized XG oligosaccharides (i.e., $Ox-H_4P_3$ (m/z 1067.4)) was based on the -2 m/z difference compared to the m/z -value of the corresponding non-oxidized block (i.e., H_4P_3 (m/z 1069.4; **Fig. 4a**) and is comparable to previous annotations of C4-oxidized LPMO products (23, 30, 42). The C4-oxidized TXG oligosaccharides were determined to be of “XXXG”-type, as in the β -galactosidase treated NcLPMO9C-TXG-digest analyzed by MALDI-TOF-MS, a major peak with m/z 1067.4 ($[M+Li]^+$), representing C4-oxidized XXXG, remained (Fig. S2b). The C4-selectivity of NcLPMO9C towards TXG previously has been reported by Agger and coworkers (32).

The NcLPMO9M-TXG-digest showed m/z -values ($[M+Li]^+$) corresponding to many different types of TXG oligosaccharides (i.e., H_7P_5 ; **Fig. 4b**). The NcLPMO9M-TXG-digest was again composed of both non-oxidized (i.e., H_5P_3 (m/z 1231.4) and oxidized oligosaccharides (i.e., $Ox-H_5P_3$; m/z 1229.4 and 1247.4; **Fig. 4b**). The m/z difference of +16 suggested the occurrence of C1-oxidation and can be explained by the

spontaneous hydrolysis of the unstable δ -lactone form (-2 Da) into the aldonic acid form ($+16$ Da) (23, 30, 42). Although some studies have shown that m/z of $+16$ could also attribute to the gem-diol form of the C4-oxidized products, other studies, i.e., in our laboratory, by using the same MALDI-TOF-MS settings as in the current work, did not observe m/z of $+16$ for C4-oxidized products (30, 42). Therefore, we suggest that TXG, most likely, was oxidatively cleaved by NcLPMO9M at C1 position. Still, occurrence of C4-oxidation could not be excluded, because of the presence of oxidized oligosaccharides with the m/z difference of -2 . These masses ($M-2$) not only represent the unstable δ -lactone form, but also the keto-form of C4-oxidized oligosaccharides (30, 32, 64).

3.2. Unambiguous structural characterization of XG degradation products generated by NcLPMO9C and NcLPMO9M

To further identify the exact TXG cleavage sites of the two NcLPMOs, digests were subjected to negative ion mode HILIC-ESI-CID-MS/MS². Similar to the data discussed above (**Fig. 2–4**), the HILIC-ESI-MS patterns of the two LPMO-TXG-digests were different (**Fig. 5**). Firstly, the masses in the base peak chromatograms of both digests showed that non-oxidized TXG oligosaccharides were present as single charged deprotonated ($[M-H]^-$) and double charged deprotonated ($[M-2H]^{2-}$) products (data not shown, M indicates the m/z of non-oxidized oligosaccharides). The same was observed for oxidized oligosaccharides represented by the $m/z -2$ products ($[M-2-H]^-$, $[M-2-2H]^{2-}$) and the $m/z +16$ products ($[M+16-H]^-$, $[M+16-2H]^{2-}$) compared to the same degree of polymerization (DP) of non-oxidized oligosaccharides. Secondly, masses that could be either C1-oxidized products or, based on their mass, formic acid adducts of non-oxidized products were observed (**Fig. 5**). For instance, m/z 1107 could represent the C1-oxidized H₅P₂, but also the formic acid adduct of non-oxidized H₄P₃ (Table S1). Nevertheless, corresponding MS² data easily distinguished formic acid adducts as these products showed a clear fragment of $m/z -46$ (formic acid; data not shown).

Due to the complexity of multiple charges and formic acid adducts, the intensity of the MS² spectra was too poor for structural elucidation. The spectral quality improved considerably after having established MS and MS² analysis via a defined mass list (Table S1; $[M-H]^-$, $[M-2-H]^-$, $[M+16-H]^-$). The chromatograms and spectra obtained via the mass list allowed structural characterization of the non-oxidized and oxidized TXG oligosaccharides released by NcLPMO9C and NcLPMO9M (**Fig. 5–7 and Table 2–3**).

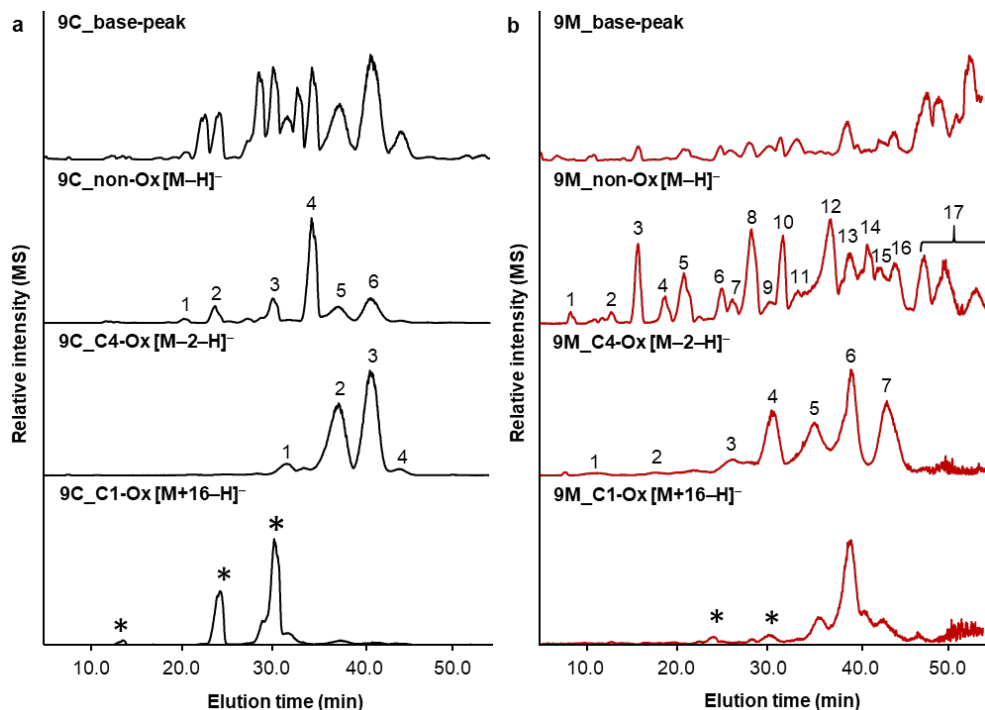


Fig. 5. HILIC-ESI-MS base-peak and extracted ion chromatograms of tamarind seed xyloglucan (TXG) digests from (a) NcLPMO9C and (b) NcLPMO9M. Extracted ion chromatograms are made for non-, C4- and C1-oxidized products (non-Ox, C4-Ox and C1-Ox) released by NcLPMO9C and NcLPMO9M (in the presence of 1 mM ascorbic acid). Determination of C4- and C1-oxidized TXG oligosaccharides is based on the m/z difference of -2 and $+16$, respectively, compared to m/z -values of corresponding non-oxidized oligosaccharides. The identification (numbered peaks) of C4-oxidized TXG oligosaccharides in NcLPMO9C- and NcLPMO9M-TXG-digest, based on the MS² fragmentation patterns, is shown in **Table 2** and **3**, respectively. Asterisks indicate formic acid adducts of non-oxidized TXG oligosaccharides having the same m/z as C1-oxidized products.

3.2.1. Characterization of non-oxidized TXG oligosaccharide products

Multiple non-oxidized TXG oligosaccharides released by the two NcLPMOs were identified (see Fig. S4 and S5 for examples). A summary of all MS² fragments and structural annotations can be found in Table S2 (for NcLPMO9C) and Table S3 (for NcLPMO9M). MS² fragments of non-oxidized products were annotated following the principle of predominance of C-/Z-type and A-type fragments of neutral oligosaccharides in negative MS-mode (65, 66). In addition, a double C-/Z-type cleavage on three linked sugar residues was observed and annotated as D-type (Fig. S4 and S5), which has previously been reported for TXG oligosaccharides (66). Overall, non-oxidized XXX (m/z 899.3, Fig. S4a), XXL (m/z 1061.4), XLX (m/z 1061.4), XLL (m/z 1223.5, Fig. S4b) and GXLL (m/z 1385.7) were formed in the NcLPMO9C-TXG-digest (Table S2). These non-oxidized "XXX"-type TXG oligosaccharides reflected cleavage at the non-reducing end of an unbranched glucosyl unit in TXG (see below). In summary, nineteen different non-oxidized TXG oligosaccharides released by NcLPMO9M were identified (Fig. S5 and Table S3).

3.2.2. Characterization of C4-oxidized TXG oligosaccharide products

Based on our previous study on CID-MS² fragmentation patterns of C4-oxidized cello-oligosaccharides (43), we identified multiple structures of C4-oxidized TXG oligosaccharides, which are shown in **Table 2 and 3**, for NcLPMO9C and NcLPMO9M, respectively. In the NcLPMO9C-TXG-digest, we found several "XXXG"-type C4-oxidized products such as ${}_{\text{O}=\text{G}}\text{GXXX}$ (m/z 1059.4, ${}_{\text{O}=\text{G}}$ indicates the C4-oxidized glucosyl unit), ${}_{\text{O}=\text{G}}\text{GXLX}$ (m/z 1221.5), ${}_{\text{O}=\text{G}}\text{GXXL}$ (m/z 1221.5), ${}_{\text{O}=\text{G}}\text{GLXX}$ (m/z 1221.5) and ${}_{\text{O}=\text{G}}\text{G}(\text{H}_5\text{P}_3)$ (m/z 1383.7) (**Table 2**). To explain the identification of these compounds, for instance through annotation of MS² fragments of ${}_{\text{O}=\text{G}}\text{GXXX}$ (m/z 1059.4, **Fig. 6a**) and ${}_{\text{O}=\text{G}}\text{GXLX}$ (m/z 1221.5, **Fig. 6b**), a fragment (Y_4) was observed having the terminal oxidized unbranched glucosyl residue removed via B-/Y-cleavage (m/z difference of 160 compared to the parent m/z). In addition, the diagnostic cross-ring fragment ${}^{2,4}\text{X}_4$ confirmed the single C4-oxidation on an unbranched glucosyl unit. This diagnostic cleavage fragment has been shown for C4-oxidized cello-oligosaccharides as well (43). Additionally, to a much lesser extent, oligosaccharides with a C4-oxidized terminal X unit were determined, such as in ${}_{\text{O}=\text{G}}\text{XXXG}$ (m/z 1059.4) and ${}_{\text{O}=\text{G}}\text{X}(\text{H}_4\text{P}_2)$ (m/z 1221.5). Again, fragments resulting from B-/Y-cleavage of the glycosidic linkage between the glucosyl units next to the C4-oxidized glucosyl unit were observed in the MS² spectra. Fragments of (m/z) 767 and 929 showed a 292 m/z difference compared to the parent m/z of 1059 and 1221, respectively. The 292 m/z difference indicated the loss of the oxidized glucosyl unit (m/z 160) substituted with a xylosyl residue (m/z 132).

Table 2. List of C4-oxidized XG oligosaccharides identified based on fragmentation patterns in CID-MS² present in the NcLPMO9C-TXG-digest. Chromatograms, including peak numbers, are shown in **Fig. 5**. Nomenclature (annotation) according to Fry et al. (15).

9C_C4-Ox Peak Nr.	Elution time (min)	m/z ([M-H] ⁻)	MS ² fragments (m/z) ^a	Annotation
1	30.7–32.7	1059.4	293 (4), 453 (1), 527 (3), 605 (13), 765 (1), 821 (16), 839 (7), 899 (100), 999 (2), 1041 (14) 455 (1), 473 (1), 767 (4), 999 (2), 1041 (14)	${}_{\text{O}=\text{G}}\text{GXXX}$ ${}_{\text{O}=\text{G}}\text{XXXG}$
2	36.4–37.5	1221.5	311 (2), 767 (4), 1001 (5), 1061 (100), 1161 (2), 1203 (34) 311 (2), 605 (3), 1001 (5), 1061 (100), 1161 (2), 1203 (34) 455 (8), 929 (12), 1161 (2), 1203 (39)	${}_{\text{O}=\text{G}}\text{GXLX}$ ${}_{\text{O}=\text{G}}\text{GXXL}$ ${}_{\text{O}=\text{G}}\text{GLXX}$ ${}_{\text{O}=\text{G}}\text{X}(\text{H}_4\text{P}_2)$
3	39.7–41.4	1383.7	311 (3), 1141 (1), 1163 (5), 1223 (100), 1365 (7)	${}_{\text{O}=\text{G}}\text{G}(\text{H}_5\text{P}_3)$
4	43.0–45.0	1515.5	n.d.	${}_{\text{O}=\text{G}}\text{H}_6\text{P}_4$

^aRelative intensities of MS² fragments are shown between brackets and informative MS² fragments are indicated in bold

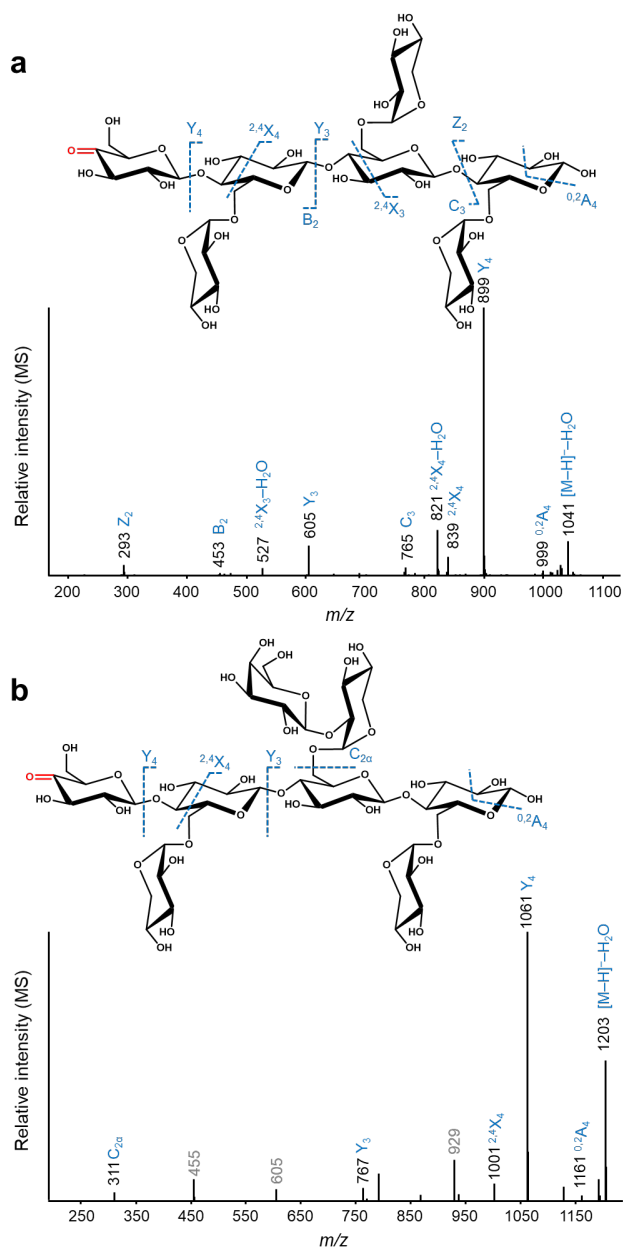


Fig. 6. Negative ion mode CID-MS² fragmentation patterns of C4-oxidized tamarind seed xyloglucan (TXG) oligosaccharides present in the NcLPMO9C-TXG-digest annotated as $o=GXXX$ (m/z 1059.4, a) and $o=GXLX$ (m/z 1221.5, b). $o=G$ indicates that the oxidation is on the glucosyl unit in keto-form. Oxidation of the C4-carbon position is indicated in red. The fragments are annotated according to the nomenclature proposed by Domon & Costello (65). The m/z-values shown in grey in b are from the other co-eluted isomers.

C4-oxidized TXG oligosaccharides released by NcLPMO9M were different from the ones formed by NcLPMO9C, which is summarized in **Table 3**. First, two small motifs, $o=GGX$ and $o=GXG$ (both m/z 471.2), were identified. The single C4-oxidation on these G and X units was confirmed by MS/MS fragments of m/z 159 and 291, respectively. In addition, C4-oxidized oligosaccharides not having "XXXG"-type structure were detected mainly including $o=GXXG$ (m/z 765.4), $o=GXGX$ (m/z 765.4), $o=GXGL$ (m/z

927.3), ${}_{\text{O}=\text{G}}\text{GLX}$ (m/z 927.3), ${}_{\text{O}=\text{G}}\text{LGX}$ (m/z 927.3). Other structures such as ${}_{\text{O}=\text{G}}\text{LLG}$ (m/z 1089.1) and ${}_{\text{O}=\text{G}}\text{LGL}$ (m/z 1089.1) were also identified (**Table 3**). Among these structures, the single C4-oxidation of G and X units was elucidated by MS² fragments

Table 3. List of C4-oxidized XG-oligosaccharides identified based on fragmentation patterns in CID-MS2 present in the NcLPMO9M-TXG-digest. Chromatograms, including peak numbers, are shown in **Fig. 5**. Nomenclature (annotation) according to Fry et al. (15).

9M_C4-Ox Peak Nr.	Elution time (min)	m/z ([M-H] ⁻)	MS ² fragments (m/z) ^a	Annotation
1	10.7–11.7	471.2	159 (59), 293 (15), 311 (15), 353 (6), 395 (14), 453 (20)	${}_{\text{O}=\text{G}}\text{GX}$
			179 (20), 291 (2), 351 (3), 221 (3), 395 (14), 453 (20)	${}_{\text{O}=\text{G}}\text{XG}$
2	17.5–18.7	765.4	291 (100), 473 (3), 543 (1), 603 (1), 747 (9)	${}_{\text{O}=\text{G}}\text{XXG}$
			291 (100), 293 (1), 353 (1), 453 (1), 455 (1), 471 (1), 473 (3), 747 (9)	${}_{\text{O}=\text{G}}\text{XGX}$
3	24.1–27.9	927.3	291 (1), 353 (1), 411 (1), 453 (100), 455 (1), 473 (18), 635 (4), 747 (1), 765 (1), 867 (2), 909 (3)	${}_{\text{O}=\text{G}}\text{XGL}$
			293 (8), 353 (1), 707 (1), 767 (1), 867 (2), 909 (3)	${}_{\text{O}=\text{G}}\text{GLX}$
4	27.9–32.0	927.3	293 (20), 311 (1), 413 (2), 453 (100), 473 (42), 615 (3), 747 (2), 849 (1), 867 (3), 909 (13)	${}_{\text{O}=\text{G}}\text{LGX}$
	30.5–31.5	1059.4	291 (37), 293 (3), 353 (4), 455 (33), 473 (2), 707 (18), 765 (4), 767 (66), 999 (1), 1041 (100)	${}_{\text{O}=\text{G}}\text{XXGX}$
			291 (37), 293 (3), 353 (4), 605 (7), 526 (3), 765 (4), 767 (66), 999 (1), 1041 (100)	${}_{\text{O}=\text{G}}\text{XGXG}$
			1041 (100)	
5	32.5–34.1	1089.1	453 (100), 557 (3), 635 (28), 867 (5), 927 (7), 1029 (6), 1071 (33)	${}_{\text{O}=\text{G}}\text{LLG}$
	33.9–34.9	1221.5	453 (100), 473 (12), 557 (5), 635 (28), 1029 (6), 1071 (33)	${}_{\text{O}=\text{G}}\text{LGL}$
			453 (100), 455 (4), 689 (2), 767 (32), 869 (4), 929 (15), 1161 (5), 1203 (5)	${}_{\text{O}=\text{G}}\text{XGXL}$
			293 (3), 689 (2), 767 (32), 869 (4), 929 (15), 1161 (5), 1203 (5)	${}_{\text{O}=\text{G}}\text{XGLX}$
6	38.1–38.9	1221.5	293 (3), 453 (100), 473 (8), 689 (2), 767 (32), 1161 (5), 1203 (5)	${}_{\text{O}=\text{G}}\text{LXGX}$
			353 (3), 453 (100), 455 (18), 689 (7), 707 (13), 767 (56), 929 (52), 1161 (2), 1203 (92)	${}_{\text{O}=\text{G}}\text{XGXL}$
	38.6–39.6	1383.7	453 (100), 473 (7), 689 (7), 707 (13), 767 (56), 1161 (2), 1203 (92)	${}_{\text{O}=\text{G}}\text{LXGX}$
			411 (2), 453 (96), 515 (2), 851 (2), 867 (2), 869 (2), 929 (100), 1013 (11), 1031 (8), 1091 (71), 1221 (10), 1305 (6), 1323 (8), 1365 (9)	${}_{\text{O}=\text{G}}\text{XGLL}$
7	41.7–44.4	1383.7	453 (13), 455 (4), 635 (3), 767 (1), 851 (3), 869 (3), 929 (100), 1091 (2), 1223 (2), 1365 (17)	${}_{\text{O}=\text{G}}\text{H}_6\text{P}_3$
		1515.5	n.d.	${}_{\text{O}=\text{G}}\text{H}_6\text{P}_4$

^aRelative intensities of MS² fragments are shown between brackets and informative MS² fragments are indicated in bold

having m/z differences of 160 and 292 from their parent m/z , respectively, as described previously. An example for the identification of $o\text{-}_G\text{L}$ units in MS^2 is shown in **Fig. 7a**, where the B_3 (m/z 453) indicated the oxidation on the H_2P_1 structure ($o\text{-}_G\text{H}_2\text{P}_1$). However, $o\text{-}_G\text{H}_2\text{P}_1$ has three isomeric structures: $o\text{-}_G\text{L}$, $o\text{-}_G\text{XG}$ and $o\text{-}_G\text{GX}$. These three structures were further distinguished by the ion B_4 ($o\text{-}_G\text{H}_3\text{P}_1$, m/z 615) and the cross-ring fragment ${}^{2,4}\text{X}_3$ (an X unit and a cross-ring cleaved G unit, m/z 413). Altogether, including the m/z of the parent oligosaccharide ($o\text{-}_G\text{H}_4\text{P}_2$, m/z 927.3), it is concluded that $o\text{-}_G\text{LGX}$ represented m/z 927.3.

All above-mentioned motifs were generated by the oxidative XG cleavage of *NcLPMO9M* at the non-reducing end of substituted glucosyl units from "XXXXG"-type building block of TXG. Furthermore, the C4-oxidized oligosaccharides having an m/z -value of 1059.5 ($o\text{-}_G\text{H}_4\text{P}_3$) in *NcLPMO9M*-TXG-digest were composed of mainly $o\text{-}_G\text{XXGX}$ and $o\text{-}_G\text{XGXX}$ instead of compounds having terminal G units (for example $o\text{-}_G\text{GXXX}$ and $o\text{-}_G\text{XXXG}$ in the *NcLPMO9C*-TXG-digest). Similarly, an m/z -value of 1221.5 was also annotated as mainly $o\text{-}_G\text{XGXL}$, $o\text{-}_G\text{XGLX}$ and $o\text{-}_G\text{LXGX}$ and an m/z -value of 1383.7 was $o\text{-}_G\text{XGLL}$ (only one was identified, **Fig. 7b**) in the *NcLPMO9M*-TXG-digest.

3.2.3. Characterization of C1-oxidized TXG oligosaccharide products

C1-oxidized products were only detected in the *NcLPMO9M*-TXG-digest. However, due to the poor signal intensity and heavy co-elution of all C1-oxidized products in HILIC-ESI-MS, these products could not be structurally identified. Nevertheless, the presence of the parent masses of C1-oxidized products confirmed that *NcLPMO9M* resulted in both C1- and C4-oxidized XG oligosaccharides.

3.2.4. Characterization of (oxidized) BCXG oligosaccharide products

We further analyzed the cleavage patterns of *NcLPMO9C*- and *NcLPMO9M*-digests towards BCXG which is a XG having additional F units (glucosyl-xylosyl-galactosyl-fucosyl residue; **Table 1**), again by using HILIC-ESI-CID-MS/ MS^2 (Fig. S6). The HILIC-ESI-MS base-peak chromatograms of two *NcLPMO*-BCXG-digests showed once more the striking difference between the patterns (Fig. S6a and S6b). Due to the high complexity, not all (oxidized) released BCXG degradation products by LPMOs were fully elucidated. Nevertheless, in the *NcLPMO9C*-BCXG-digest, we were able to identify BCXG oligosaccharides with a C4-oxidized terminal G unit (e.g. $o\text{-}_G\text{GXXF}$, m/z 1367.7, Fig. S6c), which is absent in the *NcLPMO9M*-BCXG-digest. Interestingly, a diagnostic C4-oxidized F unit ($o\text{-}_G\text{F}(\text{H}_3\text{P}_2)$, m/z 1367.7, Fig. S6d) was identified in the *NcLPMO9M*-BCXG-digest, which was absent in the *NcLPMO9C*-BCXG-digest. The identified C4-oxidized F unit indicated that oxidative cleavage of BCXG by *NcLPMO9M* also occurred next to the extensively substituted glucosyl units.

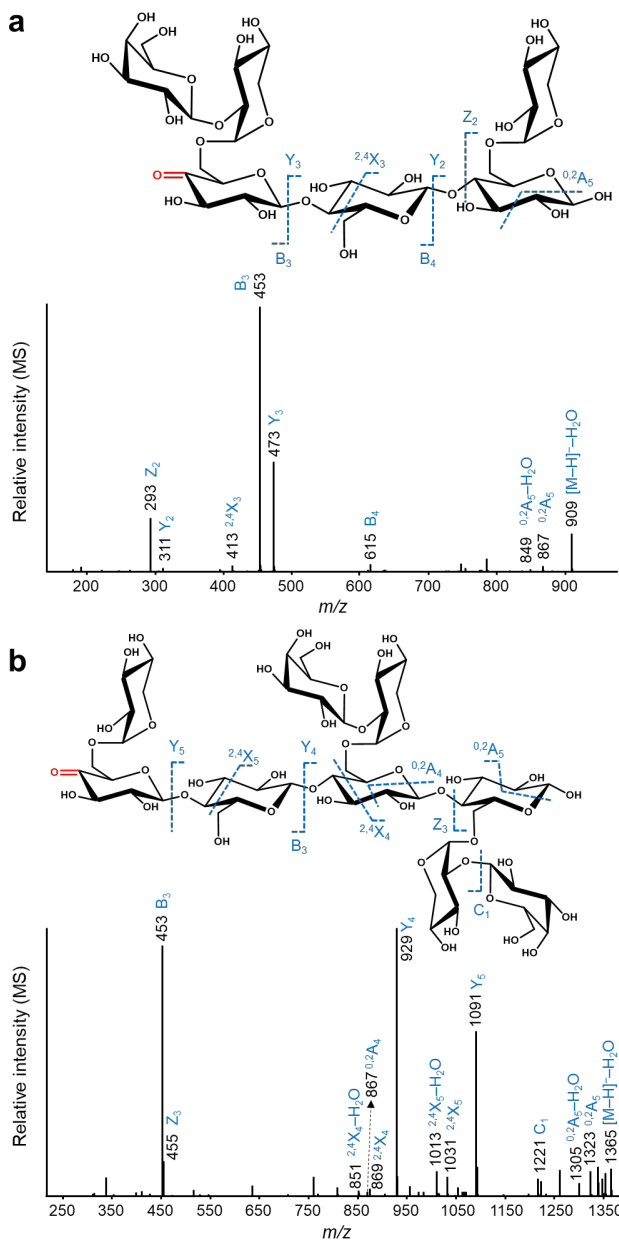


Fig. 7. Negative ion mode CID-MS² fragmentation patterns of C4-oxidized tamarind seed xyloglucan (TXG) oligosaccharides present in the NcLPMO9M-TXG-digest annotated as o=GLGX (m/z 927.3, a) and o=GXLGL (m/z 1383.7, b). o=G indicates that the oxidation is on the glucosyl unit in keto-form. Oxidation of the C4-carbon position is indicated in red. The fragments are annotated according to the nomenclature proposed by Domon & Costello (65).

3.2.5. Distinct mode-of-action of NcLPMO9C and NcLPMO9M towards XG

In this study, the structures of oxidized TXG oligosaccharides generated by two NcLPMO9C (Table 2) and NcLPMO9M (Table 3) from XG were unambiguously elucidated. In the NcLPMO9C-TXG-digest, TXG oligosaccharides were found mostly to be typical “XXXG”-type block units, but with C4-oxidized unbranched G units (e.g.

$o=GXXX$, $o=GXLX$, $o=GXXL$, $o=GLXX$ and $o=G(H_5P_3)$). Another C4-oxidized “XXXG”-type product ($o=GXXF$) was identified in the *NcLPMO9C*-BCXG-digest. In contrast, non-“XXXG”-type of C4-oxidized TXG oligosaccharides were identified in the *NcLPMO9M*-TXG-digest. The C4-oxidation of TXG oligosaccharides by *NcLPMO9M* on X and L units confirmed that *NcLPMO9M* can oxidize substituted glucosyl units at the C4-carbon. In addition, the oxidation predominately found on X and L units in HILIC-ESI-CID-MS/MS² characterized TXG oligosaccharides, instead of on unbranched G units, may reflect that *NcLPMO9M* has the preference in cleaving the substituted glucosyl backbone. The identified C4-oxidized F unit from *NcLPMO9M*-BCXG-digest further indicated that the oxidative cleavage of XG by *NcLPMO9M* is independent of the type and length of the branches. Based on these determined XG cleavage sites, it was defined that *NcLPMO9C* oxidatively cleaves XG predominantly at the non-reducing end of single unbranched glucosyl units (32), further referred to as a substitution-intolerant mode-of-action towards XG (in brief “Substitution-intolerant”) (Fig. 8). In contrast, the oxidative cleavage of XG by *NcLPMO9M* was shown to be more tolerant to substitutions with even a preference next to substituted glucosyl units and referred to as “Substitution-tolerant” (Fig. 8).

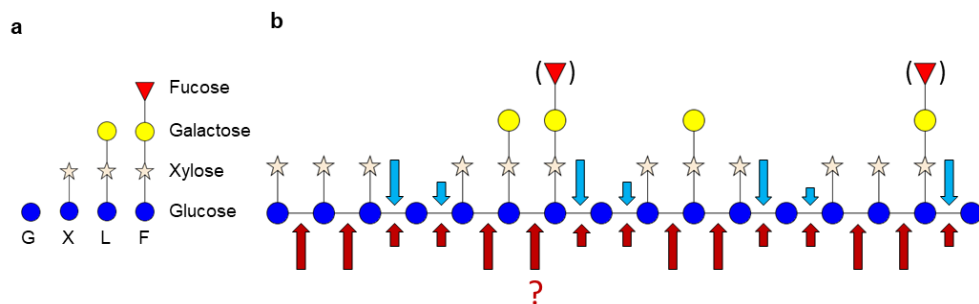


Fig. 8. (a) Illustration of structural units in xyloglucan (XG) and (b) schematic representation of XG cleavage patterns by *NcLPMO9C* (blue arrows) and *NcLPMO9M* (red arrows), respectively. G unit, glucosyl residues only (blue circle); X unit, glucosyl-xylosyl residues (xylose, star); L unit, glucosyl-xylosyl-galactosyl residues (galactose, yellow circle) and F unit, glucosyl-xylosyl-galactosyl-fucosyl residues (fucose, red triangle). Positions of galactosyl units may vary and fucosyl units are present in black currant XG, but not in tamarind seed XG. *NcLPMO9C* showed substitution-intolerant mode-of-action meaning that its oxidative cleavage towards XG was (predominately) at the non-reducing end of unbranched G units, while *NcLPMO9M* oxidatively cleaved XG regardless of substitution (substitution-tolerant) with seemingly preference on substituted glucosyl units. Whether *NcLPMO9M* can cleave between two L units remains to be studied and is shown as red question mark. The size of the arrows is indicative for more pronounced cleavage sites, which was based on (the number of) structures found of identifiable (oxidized) oligosaccharides by using HILIC-ESI-MS.

3.3. Phylogenetic and structural analysis of LPMOs with XG activity

To test our hypothesis whether the mode-of-action of AA9 LPMOs towards XG is dependent on the type of active site segments, as showcased by *NcLPMO9C* and *NcLPMO9M* (Fig. 8), amino acid sequence alignment and phylogenetic analysis were conducted. Here, all characterized fungal AA9 LPMOs (cellulose-active and XG-(plus

cellulose)active LPMOs) and a number of randomly selected uncharacterized AA9 LPMOs from the CAZy database were compared. We first aligned the mature amino acid sequences (Supplementary information), which revealed three main clusters, and generated an unrooted FL phylogenetic tree (Fig. S7). The clustering of AA9 LPMOs into three groups has already been described in literature (35, 40, 41, 67–69), however, never been used for comparisons of active site segments and XG catalytic behavior. Next, only the amino acids of the five active site segments (Seg1–Seg5, based on the definition described in our previous study (41)) were aligned (Supplementary information) and subjected to a phylogenetic analysis. The resulting structure-based SO phylogenetic tree (Fig. S8 and **Fig. 9**) shows three main clusters: one with the structural features $^{+}\text{Seg1}\text{-Seg2}$ (red area), the second defined as $^{-}\text{Seg1}^{+}\text{Seg2}$ (light blue area) and the third defined as $^{-}\text{Seg1}\text{-Seg2}$ (yellow area). A sub-cluster with a $^{-}\text{Seg1}^{+}\text{Seg2}$ feature was found (dark blue area in **Fig. 9**), but mostly with an extended Seg3 ($^{-}\text{Seg1}^{+}\text{Seg2}^{+}\text{Seg3}$).

4. Discussion

4.1. Homology of active site segments of XG-active and XG-inactive LPMOs

As previously described, *NcLPMO9C* and *NcLPMO9M* have a different catalytic site configuration in terms of neighboring segments, in particular for Seg1 and Seg2 (**Fig. 1**). In this research, we characterized *NcLPMO9C* as “Substitution-intolerant” and *NcLPMO9M* as “Substitution-tolerant”. From this, we hypothesized that the long/short Seg1 and Seg2 is a generic feature amongst AA9 LPMOs altering their interaction with XG, which further steers their mode-of-action in degrading XG. Indeed, the characterized *NcLPMOs* belong to different clusters of the structure-based SO phylogenetic tree of AA9 LPMOs (**Fig. 9**). Whether other characterized AA9 LPMOs, shown in the three clusters, have been reported to represent “Substitution-intolerant” or “Substitution-tolerant” oxidative cleavage activities is discussed here. Note that all discussed AA9 LPMOs are able to oxidatively cleave cellulose. For ease of structural comparison, published three-dimensional structures or homology models of selected characterized AA9 LPMOs from each of the three main phylogenetic clusters are shown in Fig. S9.

Based on published HPAEC and MS data, multiple LPMOs clustering in the $^{-}\text{Seg1}^{+}\text{Seg2}$ area (light blue area in **Fig. 9**) were reported to show “Substitution-intolerant” oxidative cleavage activities, like *NcLPMO9C* [(blue dots, No. 43 in **Fig. 9**; Fig. S9a), e.g., *MtLPMO9J* (No. 41 in **Fig. 9**) (22), *NcLPMO9D* (No.51 in **Fig. 9**; Fig. S9b) (20), *CvAA9A* (No. 54 in **Fig. 9**; Fig. S9c) and *LsAA9A* (No. 56 in **Fig. 9**; Fig. S9d) (24)]. Again, mainly based on published HPAEC and MS data, LPMOs clustering in the $^{+}\text{Seg1}\text{-Seg2}$ area (red area in **Fig. 9**) were reported to show seemingly “Substitution-intolerant” behavior (red dots), as observed for *GtLPMO9A-2* (No. 7 in **Fig. 9**; Fig. S9f) (28), *FgLPMO9A* (No. 15 in **Fig. 9**; Fig. S9g) (27), *TaLPMO9A* (No. 31 in **Fig. 9**; Fig. S9h) (21), *GcLPMO9B* (No. 38 in **Fig. 9**) and *GcLPMO9A* (No. 39 in **Fig. 9**) (25). The

latter enzymes all have an active site segment configuration comparable to *NcLPMO9M* (Fig. S9e).

It should be stressed that several other LPMOs have been reported to be active on XG (26, 29-31), however, the corresponding published product profiles were not conclusive due to poor chromatographic or mass spectrometric representations (**Fig. 9**, green dots). In addition, several AA9 LPMOs have been shown to be inactive on XG alone (in brief "Inactive"; see references (20, 70) and Fig. S10). Generally, we found that the $\bar{\text{Seg1}}\text{Seg2}$ configuration (**Fig. 9**, yellow area) promotes "Inactive" (e.g. *NcLPMO9F*, No. 78 in **Fig. 9**; Figs. S9i, S10).

Interestingly, we found that LPMOs having the structural feature of $\bar{\text{Seg1}}^+\text{Seg2}^+\text{Seg3}$ (dark blue area in **Fig. 9**) also showed no activity on XG alone, e.g. *MtLPMO9B* (No. 63 in **Fig. 9**; Fig. S9j) (30) and *MtLPMO9I* (No. 66 in **Fig. 9**; Fig. S11). In the $\bar{\text{Seg1}}^+\text{Seg2}^+\text{Seg3}$ cluster, although sharing the similar structural $\bar{\text{Seg1}}^+\text{Seg2}$ feature, the Seg3 segment is more extended (Supplementary information) compared to other LPMOs in the $\bar{\text{Seg1}}^+\text{Seg2}$ cluster [e.g. *NcLPMO9C* (No.43 in **Fig. 9**; Fig. S9a), *NcLPMO9D* (No.51 in **Fig. 9**; Fig. S9b), *CvAA9A* (No. 54 in **Fig. 9**; Fig. S9c) and *LsAA9A* (No. 56 in **Fig. 9**; Fig. S9d)]. In addition, our previous study demonstrated that LPMOs in the dark blue sub-cluster possess a cysteine in Seg2, which forms a disulfide bridge to a second cysteine in Seg3 (41). This disulfide bridge may pull the Seg2 away from the active site, which could mimic the effect of a short Seg2 (41). This "shortening" effect might also affect the catalytic performance towards XG as Seg2 cannot properly align to the XG alone. This could result in the described XG inactivity of LPMOs in this sub-cluster.

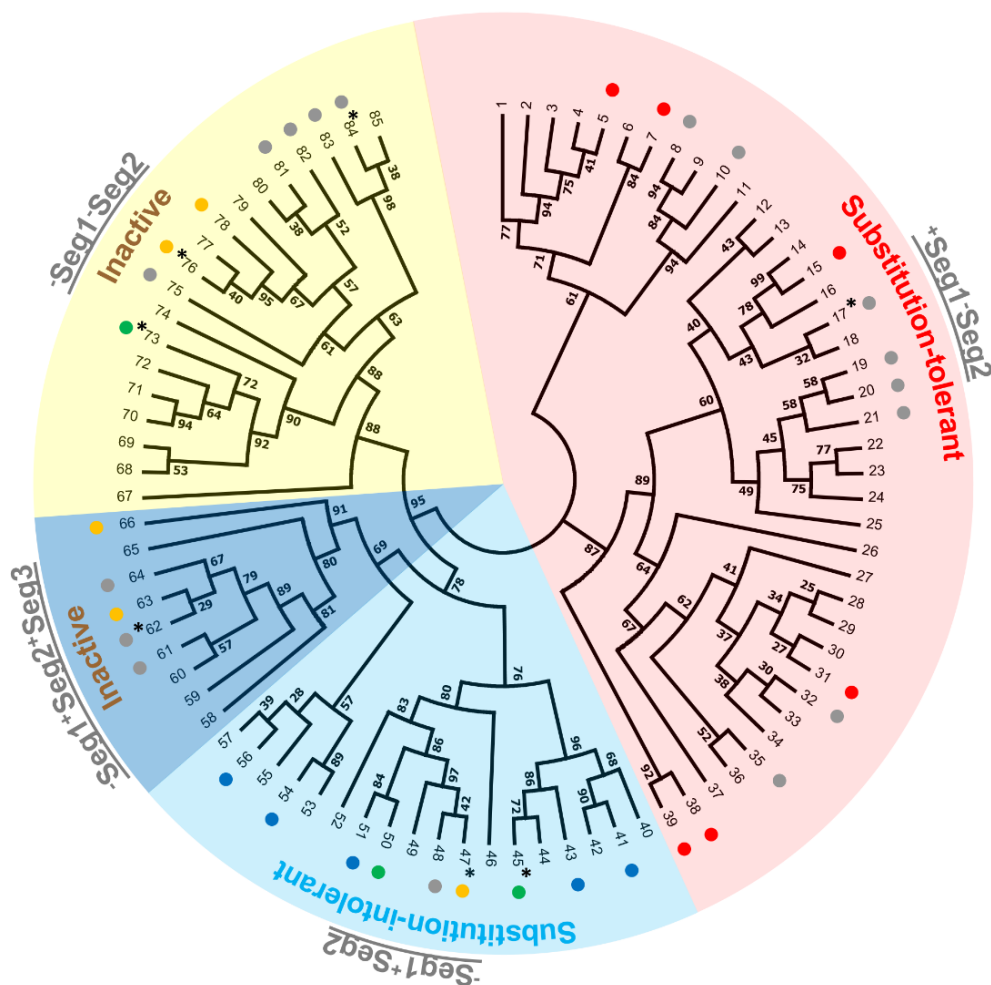


Fig. 9. Unrooted topology tree based on active site segments only of AA9 LPMOs with numbering on the terminal nodes. Corresponding organism name, Genbank accession number, LPMO name (if characterized) and PDB entry (if applicable) of each number in the terminal node is listed in Supplementary information, Table S4. Background colors in the topology tree show the long/short (+/-) of Seg1 and Seg2 segments (red, +Seg1-Seg2; light blue, -Seg1-Seg2; dark blue, -Seg1-Seg2 but extended Seg3 (+Seg3); yellow, -Seg1-Seg2). Reported characterized LPMOs are indicated by colored dots; XG product patterns alike that of NcLPMO9C (blue, Substitution-intolerant), alike that of NcLPMO9M (red, Substitution-tolerant) or not XG-active (yellow, Inactive). Grey dots indicate that the LPMOs were reported for 1) their activity on cellulose only and not tested on XG (e.g. No. 10, *HiLPMO9B*) or 2) their cellulolytic enhancing activity and not tested on XG (e.g. No. 35, *AfAA9_B*) or 3) tested with XG but the data were not conclusive (e.g. No. 20, *PaLPMO9D*). Green dots indicate that oxidative XG cleavage has been shown, but reported data are inconclusive to be determined as substitution-intolerant or -tolerant. LPMOs with red dots: 5. *NcLPMO9M*; 7. *GtLPMO9A-2*; 15. *FgLPMO9A*; 31. *TaLPMO9A*; 38. *GcLPMO9B*; 39. *GcLPMO9A*. LPMOs with blue dots: 41. *MtLPMO9J*; 43. *NcLPMO9C*; 51. *NcLPMO9D*; 54. *CvAA9A*; 56. *LsAA9A*. LPMOs with yellow dots: 47. *NcLPMO9A*; 63. *MtLPMO9B*; 66. *MtLPMO9I*; 76. *TtLPMO9E*; 78. *NcLPMO9F*.

*17. *PaLPMO9D*: XG catalytic performance was determined based on a colorimetric H₂O₂-production assay (31). *45. *PaLPMO9H*: XG product profiles were shown to be either *NcLPMO9C*-like ("Substitution-intolerant") via HPAEC (31) and *NcLPMO9M*-like ("Substitution-tolerant") via direct infusion ESI-MS/MS (26), hence inconclusive. *47. *NcLPMO9A*: "Inactive" on XG alone but "Substitution-intolerant" in combination with cellulose (20). *62. *PaLPMO9B*: XG catalytic performance was determined based on a colorimetric H₂O₂-production assay (31). *73. AN3046: Only one XG product profile was shown (MALDI-TOF mass spectrum), hence inconclusive (29). *76. *TtLPMO9E*: Inactive towards XG using ascorbic acid, but XG-active when using photosynthetic pigments with light (33). *84. *PaLPMO9E*: XG catalytic performance was determined based on a colorimetric H₂O₂-production assay (31).

4.2. Challenges in linking the mode-of-action of AA9 LPMOs to their active site segments

The AA9 structure-based phylogenetic tree (**Fig. 9**) showed three main clusters: i) LPMOs with a $\bar{\text{Seg1}}^+\text{Seg2}$ configuration and following a "Substitution-intolerant" mode-of-action; ii) LPMOs with a $^+\text{Seg1}\bar{\text{Seg2}}$ configuration and following a "Substitution-tolerant" one and iii) LPMOs with a $\bar{\text{Seg1}}^+\text{Seg2}$ and a $\bar{\text{Seg1}}^+\text{Seg2}^+\text{Seg3}$ configuration and showing only activity towards cellulose and no activity towards XG alone (**Fig. 9**). These correlations between protein structure and XG activity further reflect that AA9 LPMOs may require at least one long Seg1 or Seg2 to enable oxidative XG cleavage. Still, in each (sub-)cluster (**Fig. 9**) of the SO phylogenetic tree, beyond the challenge that a large number of reported LPMOs have not yet been tested for XG activity (**Fig. 9**, grey dots; (33, 36, 40, 67, 71-78)), irregularities seem to occur.

In the $^+\text{Seg1}\bar{\text{Seg2}}$ cluster (red area in **Fig. 9**), a "Substitution-tolerant" mode-of-action was found for all characterized LPMOs, except for *PaLPMO9D* (No. 17 in **Fig. 9**), which was determined to be "Inactive", although only based on a colorimetric H₂O₂-production assay (31). A similar conclusion of "Inactive" for *PaLPMO9B* (No. 62 in **Fig. 9**) and *PaLPMO9E* (No. 84 in **Fig. 9**) was drawn also based on the H₂O₂-production assay (31). As only a repression of the H₂O₂ production of the LPMOs is measured with this peroxidase assay, it cannot be concluded whether these LPMO show really no oxidative cleavage of XG. Hence, to confirm their (non-) XG activity, a more detailed chromatography- and mass spectrometry-based analysis is required.

In the cluster of $\bar{\text{Seg1}}^+\text{Seg2}$ (light blue area in **Fig. 9**), *NcLPMO9A* (No. 47 in **Fig. 9**), having a high structural similarity to *NcLPMO9C* (No. 43 in **Fig. 9**) and *NcLPMO9D* (No. 51 in **Fig. 9**), displayed no activity on XG alone (20). *NcLPMO9A* showed the "Substitution-intolerant" degradation only when cellulose was present (20) and apparently is, an exception in this cluster. From the same cluster, *PaLPMO9H* (No. 45 in **Fig. 9**) was reported as "Substitution-tolerant" LPMO by using direct infusion mass spectrometry (26). But, in another research, the HPAEC chromatogram of a *PaLPMO9H*-TXG-digest showed a more "Substitution-intolerant" behavior (31). Again, a more detailed chromatography- and mass spectrometry-based analysis is required to unambiguously define the mode-of-action of *PaLPMO9H* towards XG. Nevertheless, taking a closer look at the *PaLPMO9H* structure, it appeared that this enzyme has a higher content of hydrophobic amino acid residues (F, W, Y) in Seg1, less charged residues but a higher negative net charge in Seg3, and one additional positively charged residue in Seg4 (Supplementary information), compared to *NcLPMO9C*.

Also, in the $\text{Seg1}^+\text{Seg2}$ cluster (yellow area in **Fig. 9**) some exceptions were annotated. For example, AN3046 (No. 73 in **Fig. 9**) was reported to be active towards XG based on MALDI-TOF-MS data (29). However, these data remain to be verified with other analytical techniques, as the reported MALDI-TOF mass spectra only showed aldonic acid forms, while m/z -values of δ -lactone forms were absent. Detection of aldonic acids without δ -lactones in MALDI-TOF-MS analysis of LPMO-digests has not been observed in other studies. In addition, only XXLG^{ox} and XLLG^{ox} were detected in the LPMO-TXG digest, while the more common XXXG^{ox} block was not found (29). Another still difficult to classify candidate in the $\text{Seg1}^+\text{Seg2}$ cluster (yellow area in **Fig. 9**) is *Tt*LPMO9E, which has been reported as “Inactive” when using Asc as electron donor, but as active when reduced by photosynthetic pigments with light (70). The above special cases, together with LPMOs not yet tested on XG, further exemplify the difficulties and pitfalls in understanding LPMO mode-of-action towards XG based on their active site segment configuration. The latter can only be properly understood if not only experimental conditions and assays used are carefully considered, but also detailed characterization of LPMO-XG degradation products is performed, which further reflects the importance of our research. Hence, careful characterization of more LPMO mode-of-actions towards XG is highly recommended to further understand how active site segments steer the XG degradation by AA9 LPMOs.

5. Conclusions

In this study, we described two distinct XG degradation patterns generated by two AA9 *Nc*LPMOs representing different configuration of active site segments. The oxidative cleavage of XG by *Nc*LPMO9C predominantly occurred at the non-reducing end of single unbranched glucosyl units (“Substitution-intolerant”), while *Nc*LPMO9M displayed a more substitution-tolerant cleavage behavior (“Substitution-tolerant”). Based on active site segment phylogeny of AA9 LPMOs, “Substitution-intolerant” was found to correlate to the configuration $\text{Seg1}^+\text{Seg2}$, while “Substitution-tolerant” correlated to $\text{Seg1}^-\text{Seg2}$. These findings support the hypothesis that the mode-of-action of AA9 LPMOs towards XG is based on the distinct structural features of their active site segments.

6. References

1. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, et al. The path forward for biofuels and biomaterials. *Science*. 2006;311:484-9.
2. Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr Res*. 2009;344:1879-900.
3. Vogel J. Unique aspects of the grass cell wall. *Curr Opin Plant Biol*. 2008;11:301-7.
4. Chen H. Chemical composition and structure of natural lignocellulose. *Biotechnology of lignocellulose*. Dordrecht: Springer; 2014. p. 25-71.
5. Cosgrove DJ. Growth of the plant cell wall. *Nat Rev Mol Cell Biol*. 2005;6:850-61.
6. McNeil M, Darvill AG, Fry SC, Albersheim P. Structure and function of the primary cell walls of plants. *Annu Rev Biochem*. 1984;53:625-63.
7. Eijsink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnol Biofuels*. 2019;12:58.
8. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33:747-61.

9. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, et al. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219-22.
10. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.
11. Hoffman M, Jia Z, Peña MJ, Cash M, Harper A, Blackburn II AR, et al. Structural analysis of xyloglucans in the primary cell walls of plants in the subclass *Asteridae*. *Carbohydr Res*. 2005;340:1826-40.
12. Bauer WD, Talmadge KW, Keegstra K, Albersheim P. The structure of plant cell walls: II. The hemicellulose of the walls of suspension-cultured sycamore cells. *Plant Physiol*. 1973;51:174-87.
13. Zabortina OA. Xyloglucan and its biosynthesis. *Front Plant Sci*. 2012;3:134.
14. Fry SC. The structure and functions of xyloglucan. *J Exp Bot*. 1989;40:1-11.
15. Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, et al. An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plant*. 1993;89:1-3.
16. Vincken JP, York WS, Beldman G, Voragen AG. Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiol*. 1997;114:9-13.
17. Hilz H, de Jong LE, Kabel MA, Schols HA, Voragen AG. A comparison of liquid chromatography, capillary electrophoresis, and mass spectrometry methods to determine xyloglucan structures in black currants. *J Chromatogr*. 2006;1133:275-86.
18. Kiefer LL, York WS, Darvill AG, Albersheim P. Xyloglucan isolated from suspension-cultured sycamore well-walls is O-acetylated. *Phytochemistry*. 1989;28:2105-7.
19. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res*. 2014;42:D490-D495.
20. Petrovic DM, Varnai A, Dimarogona M, Mathiesen G, Sandgren M, Westereng B, et al. Comparison of three seemingly similar lytic polysaccharide monooxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J Biol Chem*. 2019;294:15068-81.
21. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sørlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci*. 2018;27:1636-50.
22. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. *PLOS ONE*. 2018;13:e0202148.
23. Frommhagen M, Westphal AH, Van Berkel WJ, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol*. 2018;9:1080.
24. Simmons TJ, Frandsen KEH, Ciano L, Tryfona T, Lenfant N, Poulsen JC, et al. Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat Commun*. 2017;8:1064.
25. Ladeveze S, Haon M, Villares A, Cathala B, Grisel S, Herpoel-Gimbert I, et al. The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2017;10:215.
26. Fanuel M, Garajova S, Ropartz D, McGregor N, Brumer H, Rogniaux H, et al. The *Podospira anserina* lytic polysaccharide monooxygenase PaLPMO9H catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnol Biofuels*. 2017;10:63.
27. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Varnai A, et al. FgLPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Lett*. 2016;590:3346-56.
28. Kojima Y, Varnai A, Ishida T, Sunagawa N, Petrovic DM, Igarashi K, et al. A lytic polysaccharide monooxygenase with broad xyloglucan specificity from the brown-rot fungus *Gloeophyllum trabeum* and its action on cellulose-xyloglucan complexes. *Appl Environ Microbiol*. 2016;82:6557-72.
29. Jagadeeswaran G, Gainey L, Prade R, Mort AJ. A family of AA9 lytic polysaccharide monooxygenases in *Aspergillus nidulans* is differentially regulated by multiple substrates and at least one is active on cellulose and xyloglucan. *Appl Microbiol Biotechnol*. 2016;100:4535-47.
30. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol biofuels*. 2016;9:186.

31. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels*. 2015;8:90.
32. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A*. 2014;111:6287-92.
33. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, et al. Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem*. 2013;288:12828-39.
34. Courtade G, Wimmer R, Rohr AK, Preims M, Felice AKG, Dimarogona M, et al. Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc Natl Acad Sci U S A*. 2016;113:5922-7.
35. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.
36. Liu B, Kognole AA, Wu M, Westereng B, Crowley MF, Kim S, et al. Structural and molecular dynamics studies of a C1-oxidizing lytic polysaccharide monooxygenase from *Heterobasidion irregulare* reveal amino acids important for substrate recognition. *FEBS J*. 2018;285:2225-42.
37. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure*. 2012;20:1051-61.
38. Courtade G, Wimmer R, Rohr AK, Preims M, Felice AKG, Dimarogona M, et al. Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc Natl Acad Sci U S A*. 2016;113:5922-7.
39. Danneels B, Tanghe M, Desmet T. Structural features on the substrate-binding surface of fungal lytic polysaccharide monooxygenases determine their oxidative regioselectivity. *Biotechnol J*. 2019;14:1800211.
40. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136:562-5.
41. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
42. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
43. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
44. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 2011;27:343-50.
45. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep*. 2017;7.
46. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, et al. The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res*. 2017;45:D313-D9.
47. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 2009;30:S162-S73.
48. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296-W303.
49. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST plus: architecture and applications. *BMC Bioinformatics*. 2009;10:421.
50. Remmert M, Biegert A, Hauser A, Soding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods*. 2012;9:173-5.
51. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792-7.
52. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870-4.

53. Rozewicki J, Li SL, Amada KM, Standley DM, Katoh K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* 2019;47:W5-W10.
54. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAXML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics.* 2019;35:4453-5.
55. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol.* 2019;29:1-4.
56. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A.* 1992;89:10915-9.
57. Veerassamy S, Smith A, Tillier ERM. A transition probability model for amino acid substitutions from blocks. *J Comput Biol.* 2003;10:997-1010.
58. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol.* 2010;17:337-54.
59. Vincken J-P, de Keizer A, Beldman G, Voragen AGJ. Fractionation of xyloglucan fragments and their interaction with cellulose. *Plant Physiol.* 1995;108:1579-85.
60. Desmet T, Cantaert T, Gualfetti P, Nerinckx W, Gross L, Mitchinson C, et al. An investigation of the substrate specificity of the xyloglucanase Cel74A from *Hypocrea jecorina*. *FEBS J.* 2007;274:356-63.
61. Ichinose H, Araki Y, Michikawa M, Harazono K, Yaoi K, Karita S, et al. Characterization of an endo-processive-type xyloglucanase having a β -(1 \rightarrow 4)-glucan-binding module and an endo-type xyloglucanase from *Streptomyces avermitilis*. *Appl Environ Microbiol.* 2012;78:7939-45.
62. Kato Y, Matsushita J, Kubodera T, Matsuda K. A novel enzyme producing isoprimeverose from oligoxyloglucans of *Aspergillus oryzae*. *J Bio chem.* 1985;97:801-10.
63. Martinez-Fleites C, Guerreiro CI, Baumann MJ, Taylor EJ, Prates JA, Ferreira LM, et al. Crystal structures of *Clostridium thermocellum* xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. *J Biol Chem.* 2006;281:24922-33.
64. Forsberg Z, Mackenzie AK, Sorlie M, Rohr AK, Helland R, Arvai AS, et al. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci U S A.* 2014;111:8446-51.
65. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 1988;5:397-409.
66. Quémener B, Vigouroux J, Rathahao E, Tabet JC, Dimitrijevic A, Lahaye M. Negative electrospray ionization mass spectrometry: a method for sequencing and determining linkage position in oligosaccharides from branched hemicelluloses. *J Mass Spectrom.* 2015;50:247-64.
67. Patel I, Kracher D, Ma S, Garajova S, Haon M, Faulds CB, et al. Salt-responsive lytic polysaccharide monooxygenases from the mangrove fungus *Pestalotiopsis sp* NCi6. *Biotechnol Biofuels.* 2016;9:108.
68. Lenfant N, Hainaut M, Terrapon N, Drula E, Lombard V, Henrissat B. A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res.* 2017;448:166-74.
69. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc.* 2012;134:890-2.
70. Cannella D, Mollers KB, Frigaard NU, Jensen PE, Bjerrum MJ, Johansen KS, et al. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. *Nat Commun.* 2016;7:11134.
71. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol.* 2017;102:1281-95.
72. Span EA, Suess DLM, Deller MC, Britt RD, Marletta MA. The role of the secondary coordination sphere in a fungal polysaccharide monooxygenase. *ACS Chem Bio.* 2017;12:1095-103.
73. Gusakov AV, Bulakhov AG, Demin IN, Sinitsyn AP. Monitoring of reactions catalyzed by lytic polysaccharide monooxygenases using highly-sensitive fluorimetric assay of the oxygen consumption rate. *Carbohydr Res.* 2017;452:156-61.
74. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, et al. High-resolution structure of a lytic polysaccharide monooxygenase from *Hypocrea jecorina* reveals a predicted linker as an integral part of the catalytic domain. *J Biol Chem.* 2017;292:19099-109.

75. Pierce BC, Agger JW, Zhang ZH, Wichmann J, Meyer AS. A comparative study on the activity of fungal lytic polysaccharide monooxygenases for the depolymerization of cellulose in soybean spent flakes. *Carbohydr Res.* 2017;449:85-94.
76. Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 angstrom resolution. *J Mol Biol.* 2008;383:144-54.
77. Lo Leggio L, Weihe CD, Poulsen JCN, Sweeney M, Rasmussen F, Lin J, et al. Structure of a lytic polysaccharide monooxygenase from *Aspergillus fumigatus* and an engineered thermostable variant. *Carbohydr Res.* 2018;469:55-9.
78. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels.* 2012;5:79.

7. Supplementary information

7.1. Supplementary materials and methods

7.1.1. Materials

Isoprimeverose (X unit) and β -galactosidase (GH35) from *Aspergillus niger* were purchased from Megazyme (Bray, Ireland). Other carbohydrate substrates, standards and chemicals were described in the section 2.

7.1.2. Removal of galactosyl residues in TXG oligosaccharides by β -galactosidase

Stock β -galactosidase (4000 U/mL) was diluted with 100 mM sodium acetate buffer (pH 4.0) to 4 U/mL. Afterwards, 60 μ L of diluted β -galactosidase solution was added to 180 μ L of XEG- and NcLPMO9C-TXG-digests (prepared as described in the section 5.2), respectively. The reactions were incubated at 30 °C for 24 h under shaking at 800 rpm in an Eppendorf ThermoMixer® C. After 24 h incubation, the reactions were stopped at 97 °C for 10 min in an Eppendorf ThermoMixer® C. Subsequently, the supernatant was recovered after centrifugation in a Hermile Z 233 MK-2 centrifuge at 22000 $\times g$ (Rotor: 220.87 VO5/6) for 10 min. The β -galactosidase treated XEG- and NcLPMO9C-TXG-digests were analyzed by HPAEC-PAD and MALDI-TOF-MS with the same settings described in the section 2. TXG oligosaccharide standards (100 μ g/mL) and isoprimeverose (X unit, 50 μ g/mL) were included in the analysis as standards.

7.2. Supplementary data

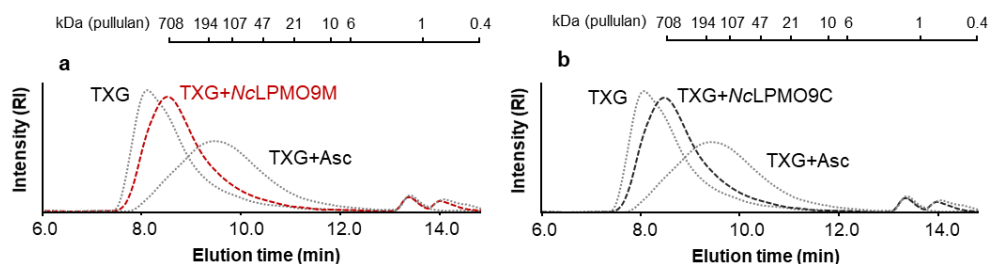


Fig. S1. Elution patterns that show the molecular weight (MW) distributions of tamarind seed xyloglucan (TXG) digests from (a) NcLPMO9M (1.25 μ M, TXG + NcLPMO9M) and (b) NcLPMO9C (1.25 μ M, TXG + NcLPMO9C) in the absence of ascorbic acid (Asc), analyzed by HPSEC-RI. For comparison, TXG without enzyme-addition, but with Asc (24 h; TXG + Asc) is also shown.

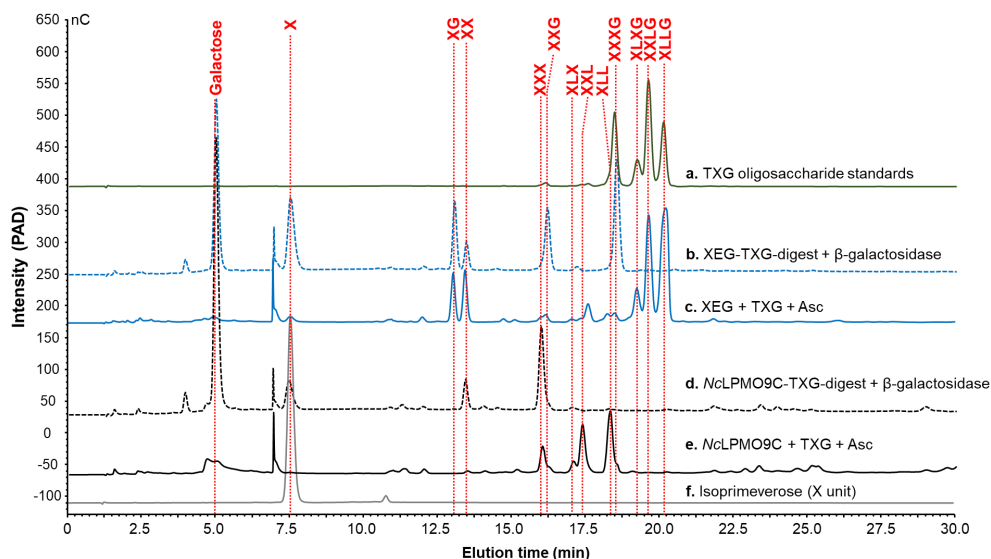


Fig. S2. HPAEC elution patterns of products after β -galactosidase treatment of xyloglucanase (XEG; c) and NcLPMO9C (d) digested tamarind seed xyloglucan (TXG) in the presence of ascorbic acid (Asc). XEG- (b) and NcLPMO9C-TXG-digests (d) before β -galactosidase treatment are included for comparison. Likewise, TXG oligosaccharide standards (a) and isoprimeverose (X unit; f) are also shown.

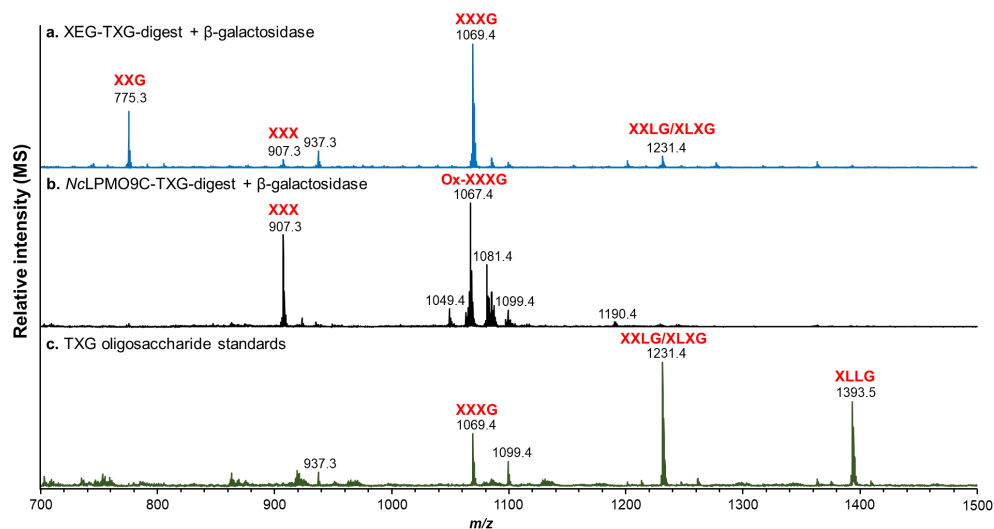


Fig. S3. MALDI-TOF mass spectra of β -galactosidase-treated xyloglucanase (XEG; a) and NcLPMO9C (b) digested tamarind seed xyloglucan (TXG) in the presence of ascorbic acid (Asc). The TXG oligosaccharide standards (c) are shown for comparison. Ox, oxidized. M/z -value corresponded to lithium (Li)-adducts.

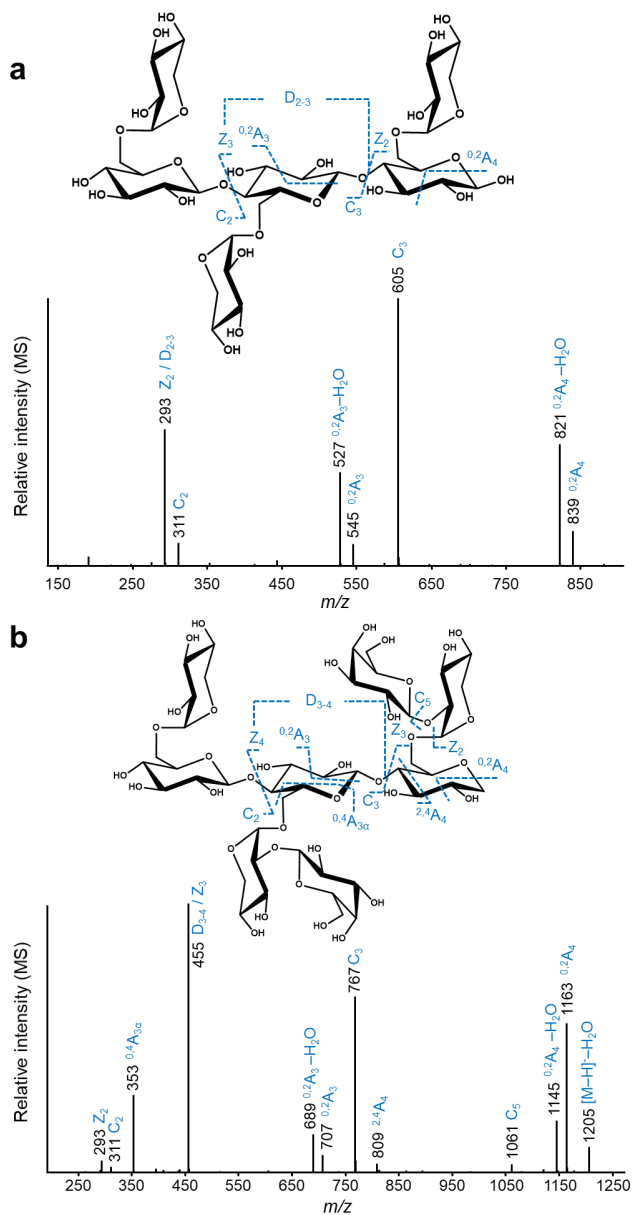


Fig. S4. Negative ion mode CID-MS² fragmentation patterns of non-oxidized tamarind seed XG oligosaccharides present in the NcLPMO9C-TXG-digest annotated as XXX (m/z 899.3, a) and XLL (m/z 1223.5, b). The fragments are annotated according to the nomenclature proposed by Domon & Costello (1).

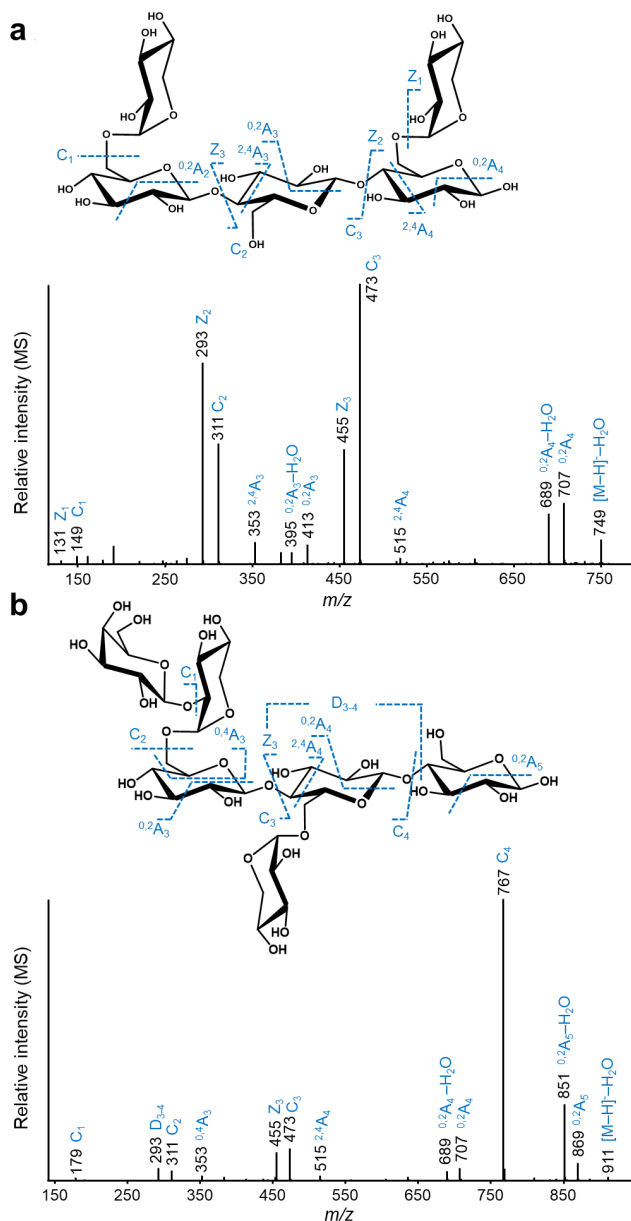


Fig. S5. Negative ion mode CID-MS² fragmentation patterns of non-oxidized tamarind seed XG oligosaccharides present in the NcLPMO9M-TXG-digest annotated as XGX (m/z 767.4, a) and LXG (m/z 929.3, b). The fragments are annotated according to the nomenclature proposed by Domon & Costello (1).

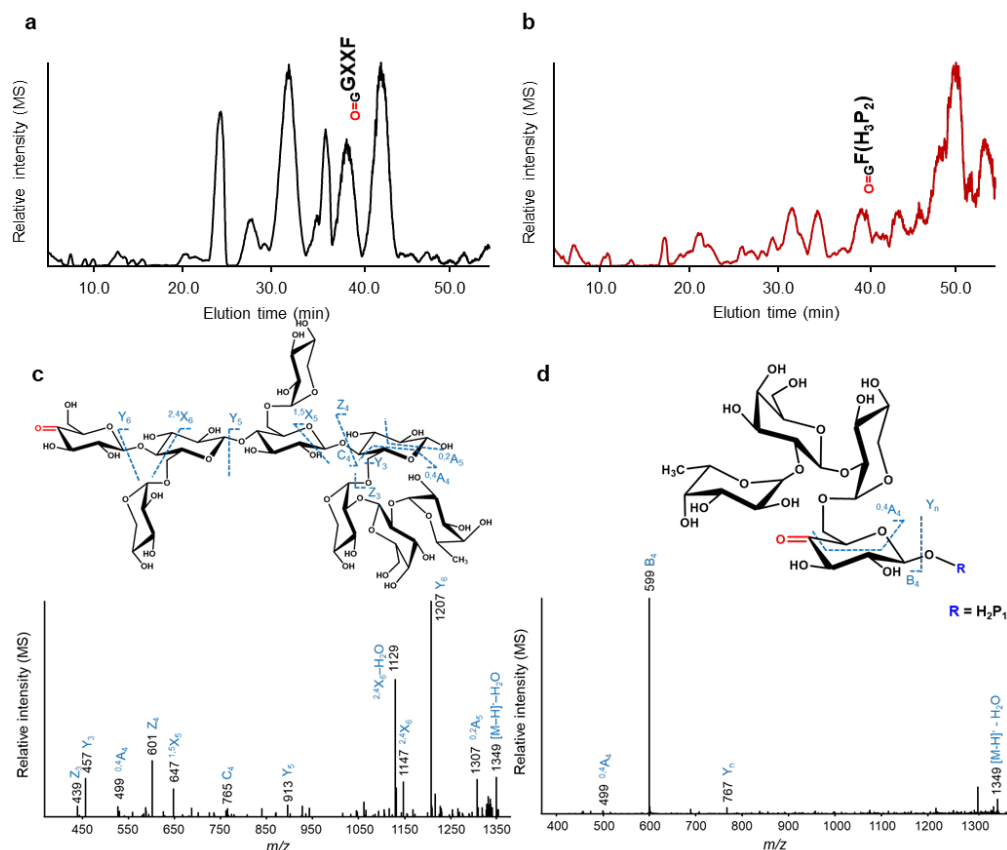


Fig. S6. HILIC-ESI-MS base-peak chromatograms of black currant xyloglucan (BCXG) digested by (a) *NcLPMO9C* and (b) *NcLPMO9M*. Negative ion mode CID-MS/MS fragmentation patterns of C4-oxidized BCXG oligosaccharides present in *NcLPMO9C*-BCXG-digest annotated as $\text{o}=\text{G} \text{GXXF}$ (m/z 1367.7, c) and in *NcLPMO9M*-BCXG-digest annotated as $\text{o}=\text{G} \text{F}(\text{H}_3\text{P}_2)$ (m/z 1367.7, d). $\text{o}=\text{G}$ indicates that the oxidation is on the glucosyl unit in keto-form. Oxidation of the C4-carbon position is indicated in red. Products formed indicated that C4-oxidative cleavage occurred next to the F unit by *NcLPMO9M*, but not by *NcLPMO9C*. The fragments are annotated according to the nomenclature proposed by Domon & Costello (1).

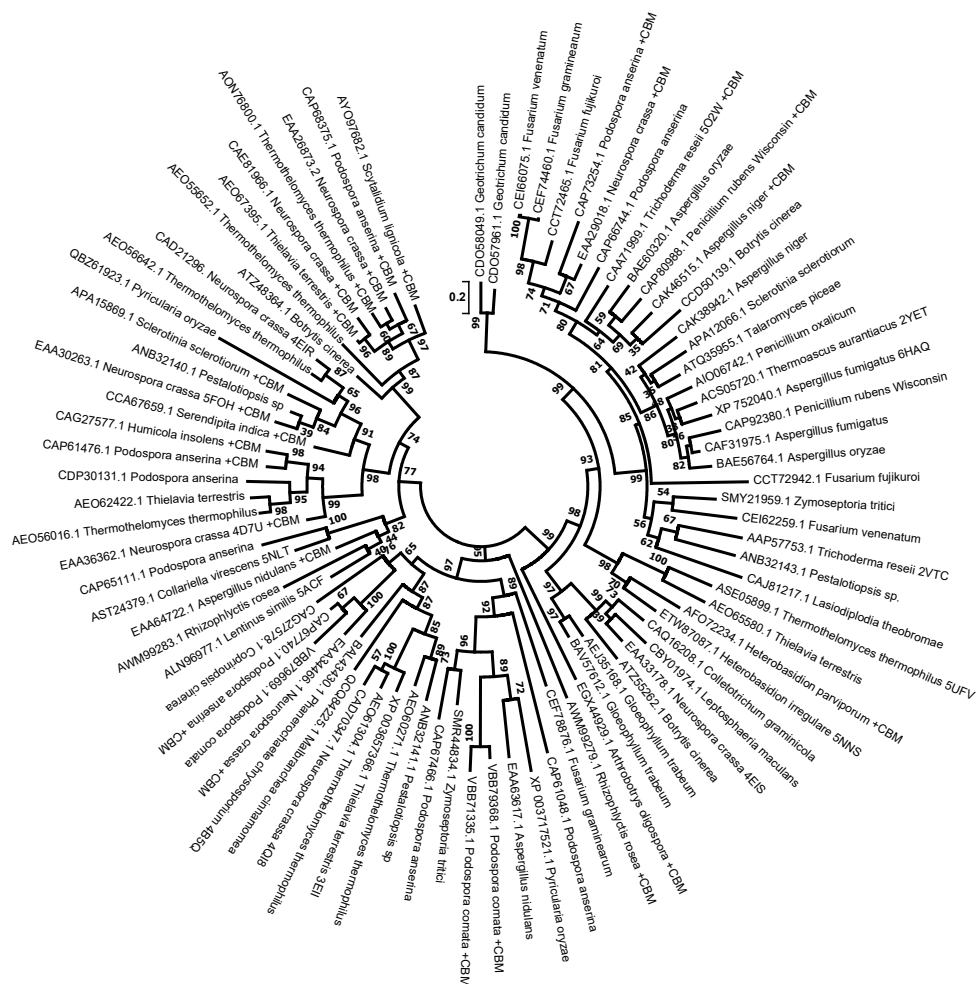


Fig. S7. Phylogenetic tree (RaxML, 800 bootstraps) in circle style of "full length" (FL) (without signal peptide) AA9 LPMOs specified by their GenBank accession number. The Protein Data Bank (PDB) accession code as well as the possession of a carbohydrate binding module (+CBM) are indicated where it applies. Bootstrap values and evolutionary distance scale are indicated. All amino acid "full sequence" used for alignment of AA9 LPMOs can be found in the online file (<https://doi.org/10.1186/s13068-020-01731-x>).

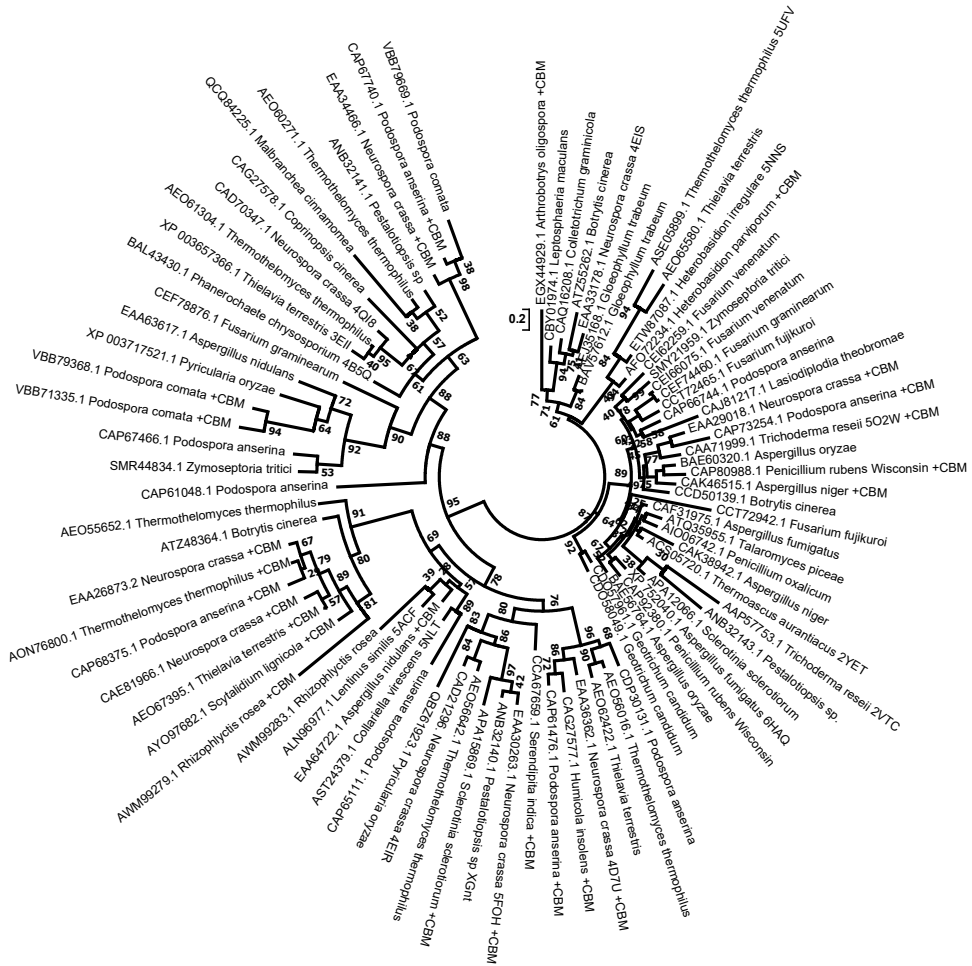


Fig. S8. Phylogenetic tree (RaxML, 1120 bootstraps) in circle style of "segments only" (SO) AA9 LPMOs specified by their GenBank accession number. The Protein Data Bank (PDB) accession code as well as the possession of a carbohydrate binding module (+CBM) are indicated where it applies. Bootstrap values and evolutionary distance scale are indicated. All amino acid "segments only" used for alignment of AA9 LPMOs can be found in the online file (<https://doi.org/10.1186/s13068-020-01731-x>).

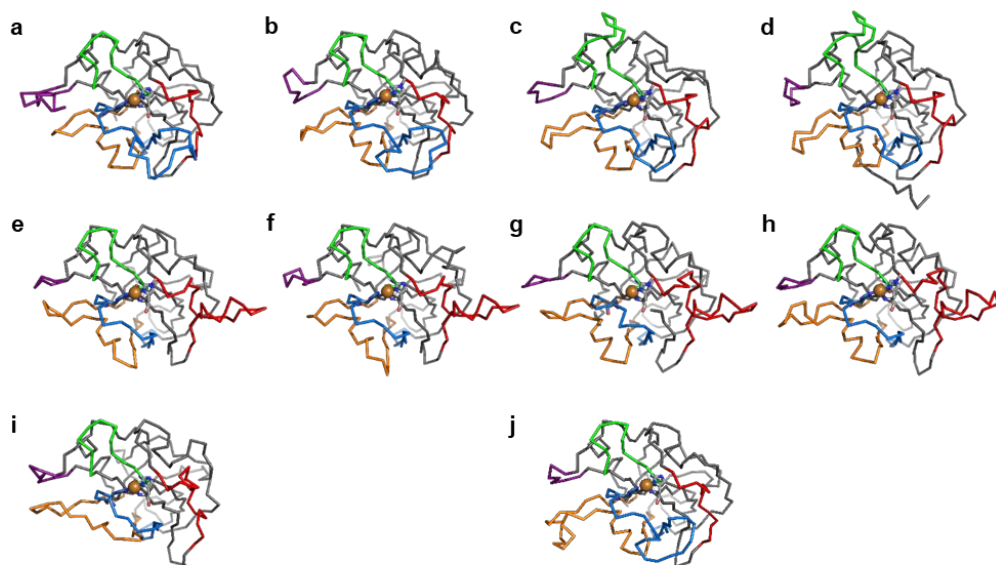


Fig. S9. Crystal structures and homology models of LPMOs tested on xyloglucan. Seg1 to Seg5 are indicated in red, blue, yellow, green and purple, respectively. (a) *NcLPMO9C* (EAA36362.1, "Substitution-intolerant", PDB entry: 4D7U), (b) *NcLPMO9D* (CAD21296.1, "Substitution-intolerant", PDB entry: 4EIR), (c) *CvAA9A* (AST24379.1, "Substitution-intolerant", PDB entry: 5NLT), (d) *LsAA9A* (ALN96977.1, "Substitution-intolerant", PDB entry: 5ACF), (e) *NcLPMO9M* (EAA33178.1, "Substitution-tolerant", PDB entry: 4EIS), (f) *GtLPMO9A-2* (BAV57612.1, "Substitution-tolerant", SwissModel SegID 59.09% with *NcLPMO9M*), (g) *FgLPMO9A* (CEF74660.1, "Substitution-tolerant", SwissModel SegID 48.89% with *TrLPMO9A*, CAA71999.1, PDB entry: 5O2W), (h) *TaLPMO9A* (ACS5720.1, "Substitution-tolerant", PDB entry: 2YET), (i) *NcLPMO9F* (CAD70347.1, "Inactive", PDB entry: 4QI8). (j) *MtLPMO9B* (AON76800.1, "Inactive", SegID 41.86% with *NcLPMO9D*, EAA32426.1, PDB entry: 5TKF).

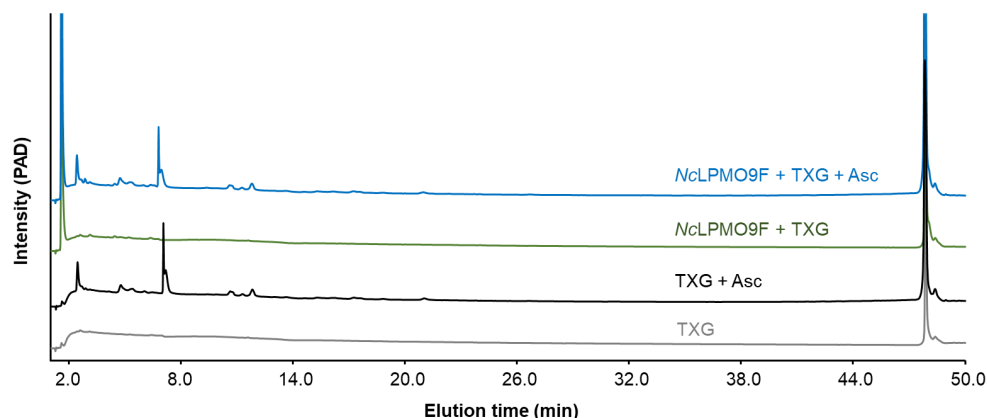


Fig. S10. HPAEC elution patterns of tamarind seed xyloglucan (TXG), and of TXG digested with *NcLPMO9F* (1.25 μ M) with addition of 1 mM ascorbic acid (Asc). Control reactions are TXG with *NcLPMO9F* in the absence of Asc and TXG with only Asc. The chromatograms show no release of oxidized products by *NcLPMO9F*. The expression, production and purification of *NcLPMO9F* were described by Kittl et al. (2).

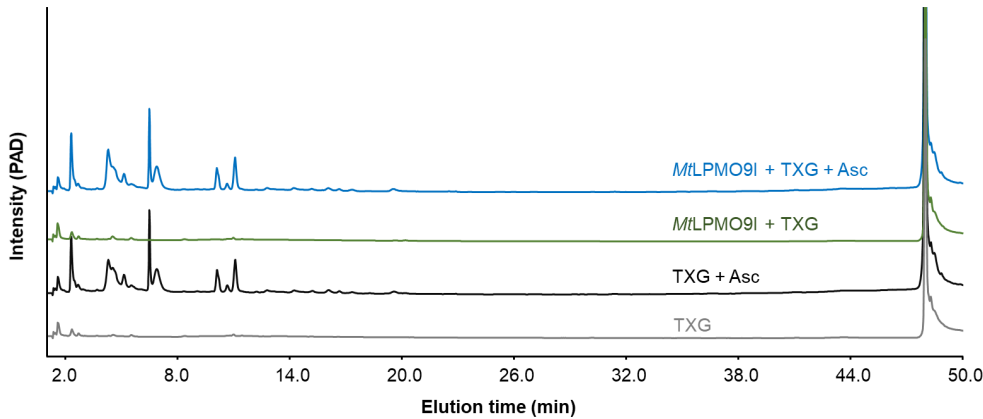


Fig. S11. HPAEC elution patterns of tamarind seed xyloglucan (TXG), and of TXG digested with *MtLPMO9I* (1.25 μ M) with addition of 1 mM ascorbic acid (Asc). Control reactions are TXG with *MtLPMO9I* in the absence of Asc and TXG with only Asc. The chromatograms show no release of oxidized products by *MtLPMO9I*. The expression, production and purification of *MtLPMO9I* were described by Sun et al. (3).

Table S1. Selected mass (m/z) list for non-oxidized (non-Ox), C4-oxidized (C4-Ox) and C1-oxidized (C1-Ox) TXG oligosaccharides (CID-MS² fragmentation). FA: formic acid. *Same m/z values for non-Ox-TXG oligosaccharides ($[M+FA-H]^-$) and C1-Ox-TXG oligosaccharides ($[M-H]^-$). Hex = hexaose, Pen = pentaose.

	Non-Ox-TXG		C4-Ox-TXG		C1-Ox-TXG	Hex	Pen
	$[M-H]^-$	$[M+FA-H]^-$	$[M-2-H]^-$	$[M-2+FA-H]^-$	$[M+16-H]^-$		
m/z	311.1	357.1	309.1	355.1	327.1	1	1
m/z	473.2	519.2	471.2	517.2	489.2	2	1
m/z	605.2	*651.2	603.2	649.2	621.2	2	2
m/z	635.3	681.3	633.3	679.3	*651.2	3	1
m/z	767.4	813.4	765.4	811.4	783.4	3	2
m/z	899.3	*945.4	897.3	943.4	915.3	3	3
m/z	929.3	975.3	927.3	973.3	*945.4	4	2
m/z	1061.4	*1107.3	1059.4	1105.3	1077.4	4	3
m/z	1091.3	1137.3	1089.1	1135.3	*1107.3	5	2
m/z	1223.5	1269.4	1221.5	1267.4	1239.5	5	3
m/z	1355.6	*1401.6	1353.6	1399.6	1371.6	5	4
m/z	1385.7	1431.6	1383.7	1429.6	*1401.6	6	3
m/z	1517.6	1563.6	1515.5	1561.4	1531.5	6	4
m/z	1679.6	1725.6	1677.6	1723.6	1695.6	7	4
m/z	1843.6	1889.6	1841.6	1887.6	1859.6	8	4

Table S2. List of non-oxidized XG oligosaccharides identified based on fragmentation patterns in CID-MS² present in the NcLPMO9C-TXG-digest. Chromatograms, including peak numbers, are shown in **Fig. 5**. Nomenclature (annotation) according to Fry et al. (4).

9C_non-Ox Peak Nr.	Elution time (min)	<i>m/z</i> ([M-H] ⁻)	MS ² fragments (<i>m/z</i>) ^a	Annotation
1	20.1–20.7	767.4	161 (1), 293 (37), 311 (6), 353 (1), 413 (1), 527 (30), 545 (9), 605 (100), 689 (15), 707 (3), 749 (11)	XXG
2	23.4–24.5	899.3	293 (51), 311 (8), 527 (34), 545 (8), 605 (100), 821 (44), 839 (13)	XXX
3	28.8–30.5	1061.4	311 (6), 455 (100), 527 (20), 545 (3), 605 (40), 899 (1), 983 (24), 1001 (21), 1043 (5)	XXL
			293 (25), 353 (23), 689 (2), 707 (2), 767 (17), 983 (24), 1001 (21), 1043 (5)	XLX
4	33.9–35.2	1223.5	293 (4), 311 (2), 353 (25), 455 (100), 689 (14), 707 (8), 767 (60), 1061 (2), 1145 (18), 1163 (58), 1205 (6)	XLL
5	36.9–37.7	1385.7	455 (9), 707 (2), 767 (12), 1145 (6), 1163 (8), 1223 (100), 1307 (12), 1325 (50), 1367 (10)	(H ₅ P ₃)G
6	39.6–41.0	1385.7	293 (2), 353 (9), 455 (46), 473 (2), 851 (10), 869 (12), 929 (100), 1163 (2), 1223 (16), 1307 (20), 1325 (73), 1367 (3)	GXLL

^aRelative intensities of MS² fragments are shown between brackets

Chapter 6

Table S3. List of non-oxidized XG oligosaccharides identified based on fragmentation patterns in CID-MS² present in the NcLPMO9M-TXG-digest. Chromatograms, including peak numbers, are shown in **Fig. 5**. Nomenclature (annotation) according to Fry et al.

9M_non-Ox Peak Nr.	Elution time (min)	<i>m/z</i> ([M-H] ⁻)	MS ² fragments (<i>m/z</i>) ^a	Annotation
1	8.3–8.7	473.2	149 (7), 161 (9), 221 (3), 311 (100), 395 (9), 413 (1), 455 (25)	XG
			149 (7), 179 (1), 293 (10), 353 (10), 395 (9), 413 (1), 455 (25)	GX
2	12.8–13.1	605.2	149 (6), 293 (100), 311 (14), 353 (1), 527 (20), 545 (7), 587 (5)	XX
3	15.3–16.2	635.3	161 (1), 179 (3), 221 (1), 311 (41), 353 (8), 413 (5), 455 (11), 557 (5), 575 (6), 617 (1)	LG
			161 (1), 179 (3), 293 (9), 311 (41), 353 (8), 473 (100), 515 (2), 557 (5), 575 (6), 617 (1)	GL
4	18.1–19.3	767.4	149 (1), 161 (1), 293 (30), 311 (7), 413 (1), 455 (1), 527 (26), 545 (9), 605 (100), 689 (13), 707 (3), 749 (3)	XXG
5	20.3–21.2	767.4	131 (1), 149 (3), 293 (73), 311 (45), 353 (9), 395 (5), 413 (7), 455 (48), 473 (100), 513 (1), 689 (18), 707 (22), 749 (8)	XGX
6	24.2–25.1	929.3	179 (1), 293 (1), 353 (6), 455 (84), 635 (3), 575 (1), 851 (20), 869 (13), 911 (2)	GLX
			161 (1), 221 (2), 353 (6), 455 (84), 689 (6), 707 (8), 767 (100), 851 (20), 869 (13), 911 (2)	XLG
7	25.6–26.7	929.3	179 (1), 293 (4), 311 (3), 353 (1), 455 (10), 473 (11), 515 (1), 689 (3), 707 (4), 767 (100), 851 (27), 869 (6), 911 (2)	LXG
8	27.5–29.0	929.3	179 (3), 293 (8), 311 (15), 353 (12), 455 (46), 473 (40), 515 (2), 557 (4), 575 (5), 635 (100), 851 (22), 869 (25), 911 (2)	LGX
9	30.0–31.1	1061.4	311 (3), 455 (6), 527 (7), 545 (2), 605 (47), 899 (1), 983 (19), 1001 (6), 1043 (5)	XXXG
			293 (13), 353 (1), 689 (25), 707 (6), 767 (100), 983 (19), 1001 (6), 1043 (5)	GXXX
10	31.2–32.2	1091.3	353 (2), 413 (1), 455 (13), 473 (2), 851 (1), 869 (5), 929 (100), 1013 (12), 1031 (12), 1073 (1)	LLG
11	34.9–35.7	1223.5	293 (1), 311 (1), 353 (2), 455 (21), 689 (3), 707 (3), 767 (20), 851 (25), 869 (12), 929 (100), 1061 (15), 1091 (8), 1145 (22), 1163 (17), 1205 (3)	XLGX
12	36.6–37.2	1223.5	293 (3), 311 (3), 353 (2), 455 (3), 473 (6), 689 (3), 707 (1), 767 (19), 851 (20), 869 (8), 929 (100), 1061 (1), 1091 (1), 1145 (29), 1163 (30), 1205 (1)	LXGX
13	37.5–38.2	1385.7	353 (2), 455 (9), 689 (1), 707 (1), 767 (8), 1145 (5), 1163 (13), 1223 (100), 1307 (20), 1325 (79), 1367 (3)	XLLG
14	40.8–41.2	1385.7	353 (1), 455 (4), 473 (1), 851 (2), 869 (1), 929 (26), 1013 (13), 1031 (11), 1091 (100), 1307 (17), 1325 (14), 1367 (1)	LLGX
15	41.9–42.3	1385.7	293 (3), 311 (5), 353 (4), 455 (2), 515 (1), 557 (2), 575 (2), 635 (24), 851 (42), 869 (10), 929 (81), 1091 (8), 1223 (3), 1307 (25), 1325 (100), 1367 (1)	LGXL
16	43.2–44.7	1517.6	n.d.	H ₆ P ₄
17	> 46.0	n.d.	n.d.	DP > 10

^aRelative intensities of MS² fragments are shown between brackets

Table S4. Characterized LPMOs; organism, GenBank accession number, LPMO name (if applicable), Protein Data Bank (PDB) entry (if applicable) and reference (if applicable). Numbers correspond to the numbers shown in the SO phylogenetic tree in **Fig. 9**. “#” Sequence contains a carbohydrate binding module (CBM).

Number	Organism	GenBank accession number	LPMO name	PDB entry	Reference
1	<i>Arthrobotrys oligospora</i>	EGX44929.1#			
2	<i>Leptosphaeria maculans</i>	CBY01974.1			
3	<i>Colletotrichum graminicola</i>	CAQ16208.1			
4	<i>Botrytis cinerea</i>	ATZ55262.1			
5	<i>Neurospora crassa</i>	EAA33178.1	NcLPMO9M	4EIS	(5, 6)
6	<i>Gloeophyllum trabeum</i>	AEJ35168.1			
7	<i>Gloeophyllum trabeum</i>	BAV57612.1	GtLPMO9A-2		(7)
8	<i>Thermothelomyces thermophilus</i> (<i>Myceliophthora thermophila</i>)	ASE05899.1	MtLPMO9D	5UFV	(8, 9)
9	<i>Thielavia terrestris</i>	AEO65580.1			
10	<i>Heterobasidion irregulare</i>	ETW87087.1	HtLPMO9B	5NNS	(10)
11	<i>Heterobasidion parviporum</i>	AFO72234.1#			
12	<i>Fusarium venenatum</i>	CEI62259.1			
13	<i>Zymoseptoria tritici</i>	SMY21959.1			
14	<i>Fusarium venenatum</i>	CEI66075.1			
15	<i>Fusarium graminearum</i>	CEF74460.1	FgLPMO9A		(11)
16	<i>Fusarium fujikuroi</i>	CCT72465.1			
17	<i>Podospora anserina</i>	CAP66744.1	PaLPMO9D		(12)
18	<i>Lasiodiplodia theobromae</i>	CAJ81217.1			
19	<i>Neurospora crassa</i>	EAA29018.1#	NCU07760		(6)
20	<i>Podospora anserina</i>	CAP73254.1#	PaLPMO9A		(12)
21	<i>Trichoderma reesei</i> (<i>Hypocrea jecorina</i>)	CAA71999.1#	TrCel61A (HjLPMO9A)	5O2W	(13-15)
22	<i>Aspergillus oryzae</i>	BAE60320.1			
23	<i>Penicillium rubens Wisconsin</i>	CAP80988.1#			
24	<i>Aspergillus niger</i>	CAK46515.1#			

25	<i>Botrytis cinerea</i>	CCD50139.1			
26	<i>Fusarium fujikuroi</i>	CCT72942.1			
27	<i>Aspergillus fumigatus</i>	CAF31975.1			
28	<i>Talaromyces piceae</i>	ATQ35955.1			
29	<i>Penicillium oxalicum</i>	AIO06742.1			
30	<i>Aspergillus niger</i>	CAK38942.1			
31	<i>Thermoascus aurantiacus</i>	ACS05720.1	TaLPMO9A	2YET	(16)
32	<i>Trichoderma reesei (Hypocrea jecorina)</i>	AAP57753.1	TrCel61B	2VTC	(17)
33	<i>Pestalotiopsis sp.</i>	ANB32143.1			
34	<i>Sclerotinia sclerotiorum</i>	APA12066.1			
35	<i>Aspergillus fumigatus</i>	XP_752040.1	AfAA9_B	6HAQ	(18)
36	<i>Penicillium rubens Wisconsin</i>	CAP92380.1			
37	<i>Aspergillus oryzae</i>	BAE56764.1			
38	<i>Geotrichum candidum</i>	CDO57961.1	GcLPMO9B		(19)
39	<i>Geotrichum candidum</i>	CDO58049.1	GcLPMO9A		(19)
40	<i>Podospora anserina</i>	CDP30131.1			
41	<i>Thermothelomyces thermophilus (Myceliophthora thermophila)</i>	AEO56016.1	MtLPMO9J		(20)
42	<i>Thielavia terrestris</i>	AEO62422.1			
43	<i>Neurospora crassa</i>	EAA36362.1#	NcLPMO9C	4D7U	(5, 6, 21)
44	<i>Humicola insolens</i>	CAG27577.1#			
45	<i>Podospora anserina</i>	CAP61476.1	PaLPMO9H		(12, 22)
46	<i>Serendipita indica</i>	CCA67659.1#			
47	<i>Neurospora crassa</i>	EAA30263.1	NcLPMO9A	5FOH	(6, 23, 24)
48	<i>Pestalotiopsis sp</i>	ANB32140.1	PsLPMO9B		(25)
49	<i>Sclerotinia sclerotiorum</i>	APA15869.1#			
50	<i>Thermothelomyces thermophilus (Myceliophthora thermophila)</i>	AEO56642.1	MtLPMO9C		(26)
51	<i>Neurospora crassa</i>	CAD21296.1	NcLPMO9D	4EIR	(6, 23, 27)

AA9 LPMO active site segment configuration steers xyloglucan degradation

52	<i>Pyricularia oryzae</i>	QBZ61923.1			
53	<i>Podospora anserina</i>	CAP65111.1			
54	<i>Collariella virescens</i>	AST24379.1	CvAA9A	5NLT	(28)
55	<i>Aspergillus nidulans</i>	EAA64722.1#			
56	<i>Lentinus similis</i>	ALN96977.1	LsAA9A	5ACF	(28)
57	<i>Rhizophlyctis rosea</i>	AWM99283.1			
58	<i>Rhizophlyctis rosea</i>	AWM99279.1#			
59	<i>Scytalidium lignicola</i>	AYO97682.1#			
60	<i>Thielavia terrestris</i>	AEO67395.1#			
61	<i>Neurospora crassa</i>	CAE81966.1#	NcLPMO9J		(2)
62	<i>Podospora anserina</i>	CAP68375.1#	PaLPMO9B		(12)
63	<i>Thermothelomyces thermophilus</i> (<i>Myceliophthora thermophila</i>)	AON76800.1#	MtLPMO9B		(26)
64	<i>Neurospora crassa</i>	EAA26873.2#	NcLPMO9E		(6)
65	<i>Botrytis cinerea</i>	ATZ48364.1			
66	<i>Thermothelomyces thermophilus</i> (<i>Myceliophthora thermophila</i>)	AEO55652.1	MtLPMO9I		(3)
67	<i>Podospora anserina</i>	CAP61048.1			
68	<i>Zymoseptoria tritici</i>	SMR44834.1			
69	<i>Podospora anserina</i>	CAP67466.1			
70	<i>Podospora comata</i>	VBB71335.1#			
71	<i>Podospora comata</i>	VBB79368.1#			
72	<i>Pyricularia oryzae</i>	XP_003717521.1			
73	<i>Aspergillus nidulans</i>	EAA63617.1	AN3046		(29)
74	<i>Fusarium graminearum</i>	CEF78876.1			
75	<i>Phanerochaete chrysosporium</i>	BAL43430.1	PcLPMO9D	4B5Q	(30)
76	<i>Thielavia terrestris</i>	XP_003657366.1	TtLPMO9E	3EII	(31)
77	<i>Thermothelomyces thermophilus</i> (<i>Myceliophthora thermophila</i>)	AEO61304.1			

78	<i>Neurospora crassa</i>	CAD70347.1	NcLPMO9F	4QI8	(5)
79	<i>Coprinopsis cinerea</i>	CAG27578.1			
80	<i>Malbranchea cinnamomea</i>	QCQ84225.1			
81	<i>Thermothelomyces thermophilus</i>	AEO60271.1	MYCTH_112089		(6)
82	<i>Pestalotiopsis sp</i>	ANB32141.1	PsLPMO9A		(25)
83	<i>Neurospora crassa</i>	EAA34466.1#	NCU00836		(6)
84	<i>Podospira anserina</i>	CAP67740.1#	PaLPMO9E		(12)
85	<i>Podospira comata</i>	VBB79669.1			

7.3. Supplementary references

1. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 1988;5:397-409.
2. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels.* 2012;5:79.
3. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym.* 2020;234:115917.
4. Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, et al. An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plant.* 1993;89:1-3.
5. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure.* 2012;20:1051-61.
6. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc.* 2014;136:562-5.
7. Kojima Y, Varnai A, Ishida T, Sunagawa N, Petrovic DM, Igarashi K, et al. A lytic polysaccharide monooxygenase with broad xyloglucan specificity from the brown-rot fungus *Gloeophyllum trabeum* and its action on cellulose-xyloglucan complexes. *Appl Environ Microbiol.* 2016;82:6557-72.
8. Span EA, Suess DLM, Deller MC, Britt RD, Marletta MA. The role of the secondary coordination sphere in a fungal polysaccharide monooxygenase. *ACS Chem Biol.* 2017;12:1095-103.
9. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol.* 2017;102:1281-95.
10. Liu B, Kognole AA, Wu M, Westereng B, Crowley MF, Kim S, et al. Structural and molecular dynamics studies of a C1-oxidizing lytic polysaccharide monooxygenase from *Heterobasidion irregulare* reveal amino acids important for substrate recognition. *FEBS J.* 2018;285:2225-42.
11. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Varnai A, et al. FgLPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Lett.* 2016;590:3346-56.
12. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospira anserina*. *Biotechnol Biofuels.* 2015;8:90.
13. Pierce BC, Agger JW, Zhang ZH, Wichmann J, Meyer AS. A comparative study on the activity of fungal lytic polysaccharide monooxygenases for the depolymerization of cellulose in soybean spent flakes. *Carbohydr Res.* 2017;449:85-94.

14. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, et al. High-resolution structure of a lytic polysaccharide monooxygenase from *Hypocrea jecorina* reveals a predicted linker as an integral part of the catalytic domain. *J Biol Chem*. 2017;292:19099-109.
15. Gusakov AV, Bulakhov AG, Demin IN, Sinitsyn AP. Monitoring of reactions catalyzed by lytic polysaccharide monooxygenases using highly-sensitive fluorimetric assay of the oxygen consumption rate. *Carbohydr Res*. 2017;452:156-61.
16. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sorlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci*. 2018;27:1636-50.
17. Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 angstrom resolution. *J Mol Biol*. 2008;383:144-54.
18. Lo Leggio L, Weihe CD, Poulsen JCN, Sweeney M, Rasmussen F, Lin J, et al. Structure of a lytic polysaccharide monooxygenase from *Aspergillus fumigatus* and an engineered thermostable variant. *Carbohydr Res*. 2018;469:55-9.
19. Ladeveze S, Haon M, Villares A, Cathala B, Grisel S, Herpoel-Gimbert I, et al. The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2017;10:215.
20. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. *PLOS ONE*. 2018;13:e0202148.
21. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A*. 2014;111:6287-92.
22. Fanuel M, Garajova S, Ropartz D, McGregor N, Brumer H, Rogniaux H, et al. The *Podospora anserina* lytic polysaccharide monooxygenase PaLPMO9H catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnol Biofuels*. 2017;10:63.
23. Petrovic DM, Varnai A, Dimarogona M, Mathiesen G, Sandgren M, Westereng B, et al. Comparison of three seemingly similar lytic polysaccharide monooxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J Biol Chem*. 2019;294:15068-81.
24. Westereng B, Arntzen MO, Aachmann FL, Varnai A, Eijsink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monooxygenases acting on cellulose. *J Chromatogr A*. 2016;1445:46-54.
25. Patel I, Kracher D, Ma S, Garajova S, Haon M, Faulds CB, et al. Salt-responsive lytic polysaccharide monooxygenases from the mangrove fungus *Pestalotiopsis sp* NCi6. *Biotechnol Biofuels*. 2016;9.
26. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol biofuels*. 2016;9:186.
27. Bodenheimer AM, O'Dell WB, Stanley CB, Meilleur F. Structural studies of *Neurospora crassa* LPMO9D and redox partner CDHIIA using neutron crystallography and small-angle scattering. *Carbohydr Res*. 2017;448:200-4.
28. Simmons TJ, Frandsen KEH, Ciano L, Tryfona T, Lenfant N, Poulsen JC, et al. Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat Commun*. 2017;8:1064.
29. Jagadeeswaran G, Gainey L, Prade R, Mort AJ. A family of AA9 lytic polysaccharide monooxygenases in *Aspergillus nidulans* is differentially regulated by multiple substrates and at least one is active on cellulose and xyloglucan. *Appl Microbiol Biotechnol*. 2016;100:4535-47.
30. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, et al. Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem*. 2013;288:12828-39.
31. Cannella D, Mollers KB, Frigaard NU, Jensen PE, Bjerrum MJ, Johansen KS, et al. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. *Nat Commun*. 2016;7:11134.

Chapter 7

Extending the diversity of *Myceliophthora thermophila* lytic polysaccharide monooxygenases: two different xyloglucan cleavage profiles

Peicheng Sun, Melanie de Munnik, Willem J.H. van Berkel and Mirjam A. Kabel

Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

Abstract

Lytic polysaccharide monooxygenases (LPMOs) play a key role in enzymatic conversion of plant cell wall polysaccharides. Continuous discovery and functional characterization of LPMOs highly contribute to the tailor-made design and improvement of hydrolytic-activity based enzyme cocktails. In this context, a new *Mt*LPMO9F was characterized for its substrate (xyloglucan) specificity, and *Mt*LPMO9H was further delineated. Aided by sodium borodeuteride reduction and hydrophilic interaction chromatography coupled to mass spectrometric analysis, we found that both *Mt*LPMOs released predominately C4-oxidized, and C4/C6-double oxidized xylogluco-oligosaccharides. Further characterization showed that *Mt*LPMO9F, having a short active site segment 1 and a long active site segment 2 (-Seg1+Seg2), followed a “substitution-intolerant” xyloglucan cleavage profile, while for *Mt*LPMO9H (+Seg1-Seg2) a “substitution-tolerant” profile was found. The here characterized xyloglucan specificity and substitution (in)tolerance of *Mt*LPMO9F and *Mt*LPMO9H were as predicted according to our previously published phylogenetic grouping of AA9 LPMOs based on structural active site segment configurations.

Keywords

Biomass; Lignocellulose; LPMO; Active site segments; Cellulose; Xyloglucan; HILIC-ESI-CID-MS/MS²; Mass spectrometric fragmentation; Reduction

1. Introduction

Lignocellulose-based biorefineries have lately attracted interest to replace fossil based refineries (1, 2). An important process step in these biorefineries is the enzymatic release of (fermentable) monosaccharides from lignocellulosic (hemi-)celluloses (3-5). Traditionally, only hydrolytic enzymes were considered relevant for (hemi-)cellulose degradation activity, and are, therefore, the basis of commercial enzyme cocktails (6, 7). In the last decade, the composition of these cocktails benefit from the discovery of lytic polysaccharide monooxygenases (LPMOs), which have been shown to greatly enhance hydrolytic conversion of (hemi-)celluloses (8-10). Continuous discovery and functional characterization of novel LPMOs is expected to highly contribute to future application-tailored hydrolytic-activity based enzyme cocktails. In this context, in our research, we aim to understand the role of all LPMOs discovered in the genome of the thermophilic fungus *Myceliophthora thermophila* C1 (*Mt*) (11, 12).

LPMOs are mono-copper dependent redox enzymes and currently classified into sequence-based "Auxiliary Activity" families (AA) 9–11 and 13–17 in the Carbohydrate-Active enZymes (CAZy) database (<http://www.cazy.org>) (13, 14). The fungal AA9 family constitutes the largest LPMO family (11). AA9 members catalyze the regioselective (C1- and/or C4-) oxidative cleavage of cellulose using molecular oxygen (O₂) and/or hydrogen peroxide (H₂O₂) and an external electron donor as co-substrates (15, 16). C1-oxidative cleavage results in δ -lactones, which convert to aldonic acids in aqueous solutions, C4-oxidative cleavage forms 4-ketaldoses. These C4-ketones are in equilibrium with their geminal diol form in aqueous solutions, although the equilibrium will majorly be on the ketone side (17, 18). Recently, we showed that C4 cellulose oxidation can be accompanied by C6-oxidation, based on identified double, C4 and C6, oxidized cello-oligosaccharides (19). Although the regioselectivity of LPMOs is not fully understood, it has been proposed that it may reflect how LPMOs bind to their substrates (20-23). The latter might also reflect their substrate specificity, as was concluded from structure-based (e.g., active site segment based) multiple sequence alignment of AA9 LPMOs (24, 25). This analysis indicated three (main) groups: i) cellulose specific LPMOs (\sim Seg1 \sim Seg2 and \sim Seg1⁺Seg2⁺Seg3), ii) cellulose and xyloglucan (substitution-intolerant) active LPMOs (\sim Seg1⁺Seg2), iii) cellulose and xyloglucan (substitution-tolerant) active LPMOs (⁺Seg1 \sim Seg2). Although in that work, a number of candidates were shown to have the named specificities, only one *Mt*LPMO has been studied for its xyloglucan specificity (**Table 1**) (25). For the other eight (partially) characterized AA9 *Mt*LPMOs, out of twenty-two present in the genome, and for yet uncharacterized *Mt*LPMOs, xyloglucan specificity needs to be unraveled.

Table 1. (Partially) characterized AA9 LPMO from *M. thermophila*^a.

LPMO name	UniProt ID	Active site segment configuration	Substrate specificity	Xyloglucan specificity
<i>MtLPMO9A</i>	G2QNT0	-Seg1-Seg2 ^b	Cellulose (RAC), xylan associated to RAC, xyloglucan ^c , mixed β -(1→3, 1→4)-linked glucan	Active ^c
<i>MtLPMO9L</i>	Unknown	-Seg1-Seg2 ^b	Cellulose (PASC, Avicel)	Inactive
MYCTH_112089	G2Q9F7	-Seg1-Seg2 ^b	Cellulose (PASC) ^d	n.d.
<i>MtLPMO9B</i>	G2QCJ3	-Seg1+Seg2+Seg3	Cellulose (RAC, Avicel, BC)	Inactive
<i>MtLPMO9I</i>	G2Q774	-Seg1+Seg2+Seg3	Cellulose (RAC, Avicel)	Inactive
<i>MtLPMO9C</i>	G2QA92	-Seg1+Seg2	Cellulose (RAC), xyloglucan ^e , mixed β -(1→3, 1→4)-linked glucan	n.d.
<i>MtLPMO9E</i> (<i>MtLPMO9J</i>)	G2Q7A5	-Seg1+Seg2	Cellulose (RAC, Avicel), xyloglucan, cello-oligosaccharides (DP≥5)	Substitution-intolerant
<i>MtLPMO9D</i>	G2QAB5	+Seg1-Seg2	Cellulose (RAC) ^d	n.d.
<i>MtLPMO9H</i>	G2Q9T3	n.d.	Cellulose (RAC, Avicel) ^d	n.d.

^aAbbreviations: RAC, regenerated amorphous cellulose; PASC, phosphoric acid swollen cellulose; BC, bacterial cellulose; n.d., not determined

^bBased on the reported short L3 loop and L2 loop

^cTrace of activity towards xyloglucan, too low to determine the xyloglucan specificity

^dOnly cellulose was tested

^eData was not conclusive to determine xyloglucan specificity

Reference: *MtLPMO9A* (26, 27); *MtLPMO9L* (28); MYCTH_112089 (29); *MtLPMO9B* (26, 30); *MtLPMO9I* (31); *MtLPMO9C* (32); *MtLPMO9E* (31, 33); *MtLPMO9D* (34); *MtLPMO9H* (19, 30, 35, 36)

Xyloglucan is a heteropolysaccharide composed of a cellulose-like β -(1→4) linked-D-glucosyl backbone. The glucosyl residues can be substituted by a D-xylosyl residue via α -(1→6) linkages (37-39). The unsubstituted and D-xylosyl substituted glucosyl units are coded as "G" and "X" based on the one-letter nomenclature developed by Fry et al (40). The D-xylosyl residues can be even further substituted with β -(1→2) linked D-galactosyl residues (coded "L"). Other substitutions are less common and described elsewhere (40). The most common xyloglucan structure is built by so-called "XXXG-" and "XXGG-type" block-wise units (41). For instance, tamarind xyloglucan (TXG) is constructed by the repeated "XXXG-type" units with partially substituted galactosyl residues (XLXG, XXLG and XLLG) (40).

In this work, a new *MtLPMO9F* and a partially characterized *MtLPMO9H* were studied for their active site configuration, and produced for characterization of their regioselectivity and substrate specificity with a focus on oxidative cleavage patterns of xyloglucan. *MtLPMO9F*- and *MtLPMO9H*-generated C4-oxidized xylogluco-oligosaccharides, and double oxidized ones (C4/C6), were identified in detail by using

sodium borodeuteride (NaBD₄) reduction and hydrophilic interaction chromatography–mass spectrometry (HILIC-ESI-CID-MS/MS²).

2. Materials and methods

2.1. Carbohydrate and chemicals

NaBD₄ and ammonium acetate were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). Xyloglucan from tamarind (*Tamarindus indica*, TXG), TXG oligosaccharide standard (xyloglucan hepta-, octa- and nona-saccharides), cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose were purchased from Megazyme (Bray, Ireland). Regenerated amorphous cellulose (RAC) was prepared from Avicel® PH-101 (Sigma-Aldrich) as described previously (27). Ascorbic acid (Asc) was purchased from VWR International (Radnor, Pennsylvania, USA). Water used in all experiments was produced by a Milli-Q system (Millipore, Molsheim, France). Other carbohydrates used for substrate screening were purchased from either Sigma-Aldrich or Megazyme.

2.2. Structure-based multiple sequence alignment

Amino acid sequences of *MtLPMO9F* (MYCTH_111088, UniProt ID: G2Q9F7) and *MtLPMO9H* (MYCTH_46583, UniProt ID: G2Q9T3), together with previously studied *NcLPMO9C* (24, 25), *MtLPMO9E* (31), *NcLPMO9M* (25) and *FgLPMO9A* (42) were fine-tuned by removing the signal peptide, the linker- and the CBM-domain as described previously (25). Subsequently, a structure-based multiple sequence alignment (25) was performed with these six AA9 LPMOs. The resulting structure-based alignment was further divided into regions “Segments 1 to 5” (Seg1–Seg5) as described previously (25), which was used to determine the short and/or long segments.

2.3. Expression, production and purification of *MtLPMO9F* and *MtLPMO9H*

The gene encoding *MtLPMO9F* was homologously expressed in a low protease/low (hemi-) cellulose producing *Myceliophthora thermophila* C1 strain by IFF Nutrition & Biosciences (Leiden, The Netherlands), essentially as described elsewhere (43, 44). The expression, production and purification of *MtLPMO9H* and *MtAA16A* have been described previously (30, 45).

MtLPMO9F was purified by four subsequent chromatographic steps. Crude *MtLPMO9F*-rich fermentation broth was filtrated and dialyzed against 10 mM potassium phosphate buffer pH 7.6 before chromatographic purification. The dialyzed *MtLPMO9F* was purified by step-wise anion exchange chromatography (AEC), size exclusion chromatography (SEC) and cation exchange chromatography (CEC). Columns used, purification settings and elution program of AEC, SEC and CEC have been described previously (30). The purest (based on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)) third step CEC-purified *MtLPMO9F*-containing fractions were further purified by the fourth-step CEC on an ÄKTA-Micro preparative chromatography system (GE Healthcare). Settings and elution program used in this last CEC purification step of

MtLPMO9F was the same as described for the last CEC purification step of *MtAA16A* (45). All fractions were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to an approximate concentration of 2 mg/mL determined by BCA assay and analyzed by SDS-PAGE, as described previously (31) to determine their purity. *MtLPMO9F* fractions with the highest purity based on SDS-PAGE were frozen in liquid nitrogen and stored at -80 °C.

2.4. Generation of carbohydrate digests by *MtLPMO9F* and *MtLPMO9H*

Carbohydrates (listed in **Table 2**) were mixed with 50 mM ammonium acetate buffer (pH 5.0) to a concentration of 2 mg/mL (for RAC and hemicellulose combination, each type was 2 mg/mL). Subsequently, 2 μ M (corrected by impurities based on SDS-PAGE results in Fig. S1) of *MtLPMO9F* or *MtLPMO9H* was added to the corresponding carbohydrate mixture containing 1 mM Asc (final concentration). For RAC and TXG incubations, in order to increase the yield of oxidative products, 2 μ M of *MtAA16A* (previously reported as an AA9 *MtLPMO* boosting enzyme) was added. Control reactions were performed in the absence of Asc (with and without *MtAA16A* for RAC and TXG incubation). All reactions (500 μ L total volume) were incubated at 30 °C by using an Eppendorf Thermomixer® comfort, placed in an almost vertical direction, at 800 rpm for 24 h in duplicate. The reactions were stopped by immediately separating supernatants after centrifugation at 22000 $\times g$ for 10 mins at 4 °C in a table centrifuge. The resulting supernatants were collected and diluted five times for high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) analysis.

Part of supernatants from RAC and TXG digests were reduced by NaBD₄ followed by solid phase extraction (SPE) as described in section 2.5. A cello-oligosaccharide standard mixture (containing cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose, 1 μ g/mL each) and a TXG oligosaccharide standard (20 μ g/mL) were also analyzed by HPAEC-PAD.

2.5. Reduction of generated (oxidized) cello- and xylogluco-oligosaccharides with NaBD₄ and clean-up with SPE

Reduction was performed by adding 200 μ L freshly prepared 0.5 M NaBD₄ to 200 μ L of i) the cello-oligosaccharide standard mixture (50 μ g/mL of each degree of polymerization (DP)), ii) 100 μ g/mL of TXG oligosaccharide standard and iii) supernatants obtained from the *MtLPMO9F*- and *MtLPMO9H*-RAC or TXG digests (with and without *MtAA16A*) at 20 °C for 20 h. A clean-up procedure for reduced digests was carried out by using SPE with Supelclean™ ENVI-Carb™ columns (3 mL, Sigma-Aldrich) as described previously (31). The dried reduced RAC and TXG digests were dissolved in 60% (v/v) acetonitrile in water. The reduced RAC and TXG digests were directly used for HILIC-ESI-CID-MS/MS² analysis or diluted twenty times for HPAEC-PAD analysis.

2.6. Analytic methods

2.6.1. HPAEC-PAD analysis for profiling oligosaccharides

All samples, including the non-reduced and reduced cello-oligosaccharide standard mixture, the TXG oligosaccharide standard, RAC and TXG digests of *MtLPMO9F* or *MtLPMO9H*, were analyzed by HPAEC-PAD with an ICS-5000 system (Dionex, Sunnyvale, California, USA) equipped with a CarboPac PA-1 column (2 mm ID × 250 mm; Dionex) in combination with a CarboPac PA guard column (2 mm ID × 50 mm; Dionex). The two mobile phases were (A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH and the elution profile used has been described previously (26, 31). HPAEC data was processed by using Chromeleon 7.2.10 software (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

2.6.2. HILIC-ESI-CID-MS/MS² for elucidating the reduced oligosaccharide structures

Reduced forms of the cello-oligosaccharide standard mixture, the TXG oligosaccharide standard and digests of RAC and TXG were analyzed by using HILIC-ESI-CID-MS/MS². A Vanquish UHPLC system (Thermo Fisher Scientific) equipped with an Acquity UPLC BEH Amide column (Waters, Millford, Massachusetts, USA; 1.7 μm, 2.1 mm ID × 150 mm) and a VanGuard pre-column (Waters; 1.7 μm, 2.1 mm ID × 150 mm) was used. The UHPLC settings, mobile phases and elution program have been described previously (25). The MS and MS² (negative ion mode) settings have been described previously (31), with the minor changes of MS²: minimum signal threshold of 10000 counts at an activation Q of 0.25. Mass spectrometric data were processed by using Xcalibur 4.3.73.11 software (Thermo Fisher Scientific).

3. Results and discussion

3.1. *MtLPMO9H* and *MtLPMO9F*: active site segment configuration, substrate screening and cellulose regioselectivity

To visualize the configuration of active site segments (Seg1 to Seg5) of *MtLPMO9H* and *MtLPMO9F*, their amino acid sequences were (structurally-based) aligned with four AA9 LPMOs with known configuration of active site segments described in the previously published phylogeny (24, 25). Based on the alignment shown in Fig. S2, it was concluded that *MtLPMO9H* has a long Seg1 and a short Seg2 (+Seg1-Seg2), similar to the configuration of *NcLPMO9M* and *FgLPMO9A*. In contrast, *MtLPMO9F* holds a short Seg1 and a long Seg2 (-Seg1+Seg2), which is comparable to *NcLPMO9C* and *MtLPMO9E*. In our previous work, as outlined in the introduction, AA9 LPMOs having "+Seg1-Seg2" structural elements have been shown to oxidatively cleave cellulose in addition to xyloglucan via a "substitution-tolerant" cleavage behavior. AA9 LPMOs having "-Seg1+Seg2" structural elements have been shown to oxidatively cleave cellulose in addition to xyloglucan via a "substitution-intolerant" cleavage behavior. To test these, structure based, predicted xyloglucan cleavage behaviors of *MtLPMO9F* and *MtLPMO9H*, first, an extensive substrate screening was performed (Table 2), with the two,

homologously produced, and purified, *Mt*LPMOs (Fig. S2). In an attempt to increase the yield of oxidized products, a previously reported *Mt*LPMO-boosting enzyme *Mt*AA16A (45) (**Chapter 8**) was added to the corresponding cellulose and xyloglucan incubation. Although *Mt*LPMO9H and *Mt*LPMO9F still contained a trace of cellulase impurity (Fig. S3), while this was absent for the *Mt*AA16 used (Fig. S3), LPMO-generated oxidized cello-oligosaccharides dominated (Fig. S3).

Overall, in the presence of Asc, *Mt*LPMO9H and *Mt*LPMO9F showed detectable oxidative cleavage of all four celluloses (**Table 2**). The oxidative cleavage of cellulose by two *Mt*LPMOs was indeed boosted by *Mt*AA16A (Fig. S3). Based on the previously published, well-characterized LPMO-RAC profiles (19, 26, 31), *Mt*LPMO9F released predominantly C4-oxidized cello-oligosaccharides in addition to (lesser) C1-oxidized ones from RAC (Fig. S3). Note that after NaBD₄ reduction and HILIC-ESI-CID-MS/MS² analysis, it was confirmed that *Mt*LPMO9F also generated a series of (reduced) C4/C6-double oxidized cello-oligosaccharides (RD-C4C6ox) (Fig. S4) as has been shown for other AA9 LPMOs (**Chapter 4**). These MS² fragmentation patterns (DP4 as example) are shown in Fig. S5. C1oxG_n were barely detected in the (*Mt*AA16A-)*Mt*LPMO9F-RAC digest.

In addition to cellulosic substrates, *Mt*LPMO9F also showed oxidative cleavage towards xyloglucan, mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked β -glucan, glucomannan, cellopentaose and celohexaose. Interestingly, the substrate specificity of *Mt*LPMO9F is comparable to that of *Nc*LPMO9C (18, 46) and *Mt*LPMO9E (25, 33), all having a similar active site configuration (\sim Seg1+Seg2). *Mt*LPMO9H, on the other hand, only showed cleavage towards xyloglucan next to oxidative cleavage of cellulosic substrates.

Next, we studied the xyloglucan cleavage by *Mt*LPMO9H and *Mt*LPMO9F in detail. As predicted, oxidative cleavage of xyloglucan, and distinct product profiles were observed (Fig. S6). Both enzymes were free of xyloglucanase impurity (Fig. S6). Remarkably, in contrast to the boosting effect mentioned above for cellulose, *Mt*AA16A did not boost oxidative cleavage of xyloglucan by *Mt*LPMO9H and *Mt*LPMO9F (Fig. S6). The non-boosting effect might be explained by early LPMO inactivation when xyloglucan was used as the substrate, which is reflected by the incomplete conversion of TXG by these two LPMOs. Interestingly, early inactivation towards xyloglucan has not been observed for *Nc*LPMO9C and *Nc*LPMO9M, both having comparable active site segment configurations as *Mt*LPMO9H and *Mt*LPMO9F (25).

Table 2. Carbohydrate substrate specificity screening of *MtLPMO9F* and *MtLPMO9H*.

Occurrence of oxidative cleavage (+) or not (-) (in the presence of 1 mM Asc)		
Substrate	<i>MtLPMO9F</i>	<i>MtLPMO9H</i>
Cellulose		
RAC	+	+
BC	+	+
AVI	+	+
CMC	+	+
Hemicellulose		
Xyloglucan (tamarind)	+	+
Mixed β -(1→3, 1→4)-linked glucan (barley)	+	-
Mixed β -(1→3, 1→4)-linked glucan (oat spelt)	+	-
Xylan (oat spelt)	-	-
Xylan (birchwood)	-	-
Arabinoxylan (wheat)	-	-
Mannan (acacia)	-	-
Galactan (potato)	-	-
Glucomannan (konjac)	+	-
Arabinan (sugar beet)	-	-
Laminarin (<i>Laminaria digitata</i>)	-	-
RAC/Hemicellulose combination		
RAC + Xyloglucan (tamarind)	+ ^a	+ ^a
RAC + Xylan (birchwood)	+ ^b	+ ^b
Oligosaccharides		
Cellopentaose	+	-
Cellohexaose	+	-
Xylo-oligosaccharides (DP2–6)	-	-
Others		
Chitin (shrimp shells)	-	-
Starch (maize)	-	-

^aOxidative cleavage towards both RAC and xyloglucan^bOxidative cleavage only towards RAC

3.2. Xyloglucan cleavage profiles of *MtLPMO9H* and *MtLPMO9F* correlate to their active site segment configuration

To further map detailed xyloglucan product profiles generated by *MtLPMO9H* and *MtLPMO9F*, the corresponding digests were reduced by using NaBD₄ and subjected to HPAEC (**Fig. 1**) and HILIC-ESI-CID-MS/MS² (**Fig. 2**). In HPAEC chromatograms of the reduced TXG digests, again different TXG oligosaccharide profiles were observed for *MtLPMO9H* (**Fig. 1b**), and *MtLPMO9F* (**Fig. 1c**), respectively.

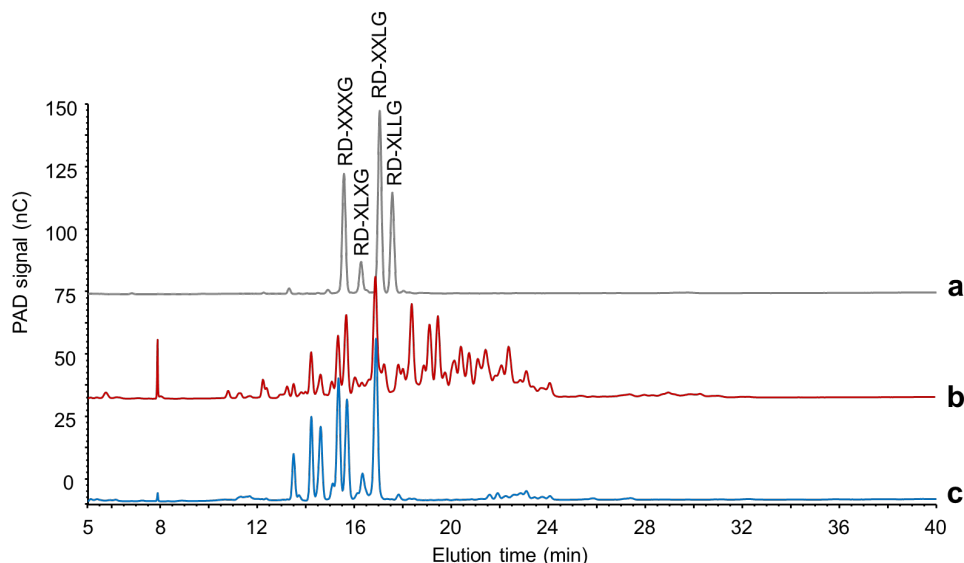


Fig. 1. HPAEC chromatograms of *MtLPMO9H*- and *MtLPMO9F*-TXG digests after NaBD_4 -reduction. (a) Reduced TXG oligosaccharide standard mixture (RD-XXXG, RD-XLXG, RD-XXLG and RD-XLLG); (b) *MtLPMO9H*-TXG digest in the presence of Asc and (c) *MtLPMO9F*-TXG digest in the presence of Asc.

Similar to the HPAEC data, the HILIC-ESI-MS base-peak chromatograms of the two digests were different (**Fig. 2**). The reduction significantly improved the separation, especially of C4-oxidized TXG oligosaccharides in HILIC, compared to the previously reported non-reduced ones (25). Nevertheless, a drawback of reduction of C4-oxidized oligosaccharides (RD-C4ox) is that both glucosyl and galactosyl non-reducing ends are formed (31), depending whether the hydroxyl group adds in axial or equatorial position to the C4 of the non-reducing end. Unfortunately, these reduced “corresponding” couples (e.g., reduced C4-oxidized TXG-products) were not well separated in HILIC, and comprise the same m/z . Therefore, in the further characterization, glucosyl or galactosyl non-reducing ends were not further distinguished. On the basis of m/z values and corresponding MS^2 fragmentation patterns, multiple (reduced) C4-oxidized TXG oligosaccharides were identified. In particular, for the *MtLPMO9H*-TXG digest, originally C4-oxidized TXG oligomers having the C4-oxidation at their (non-reducing) X unit (e.g., RD-C4ox-XG (m/z 477.3), RD-C4ox-XX (m/z 609.3), RD-C4ox-XXL (m/z 1065.5; **Fig. 3a**)), and at their (non-reducing) L unit (e.g., RD-C4ox-LGX (m/z 933.4; **Fig. 3b**)) were identified. All these products are evident for TXG “substitution-tolerant” cleavage, and these products were absent in the (reduced) *MtLPMO9F*-TXG digest.

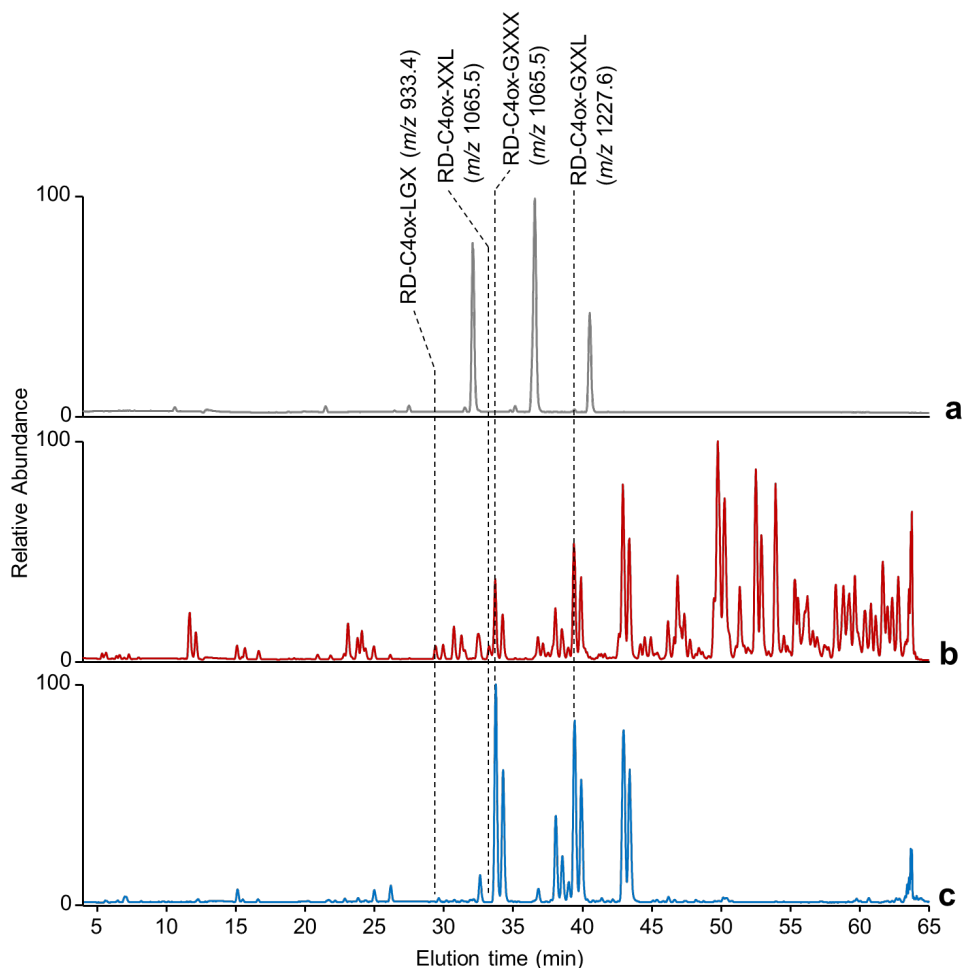


Fig. 2. HILIC-ESI-MS base-peak chromatograms. (a) Reduced TXG oligosaccharide standard mixture, (b) *MtLPMO9H*-TXG digest in the presence of Asc and (c) *MtLPMO9F*-TXG digest in the presence of Asc.

To briefly explain the structural identification shown in **Fig. 3a**, for RD-C4ox-XXL (m/z 1065, $[M-H]^-$), predominantly Y- and Z-type fragments were found, including Y_2 (m/z 770), Z_2 (m/z 752), Y_1 (m/z 476) and Z_1 (m/z 458). These fragments, especially Y_2 and Z_2 , indicated a loss of a reduced C4-oxidized X unit. The m/z difference of 294, from Y_2 to Y_1 and Z_2 to Z_1 , indicated an internal X unit, situated directly next to the (reduced) C4-oxidized X unit. The reduced C4-oxidized X unit was further confirmed by the fragments from cross-ring cleavage including ${}^{2,4}A_3$ (m/z 354), ${}^{0,2}A_3 (-H_2O)$ (m/z 546 and 528) and ${}^{2,4}A_4$ (m/z 648). The m/z difference of 162 and 294 compared Y_{10} (m/z 903) and Y_{20} (m/z 771) to the parent m/z (1065.5), respectively, and confirmed the final structure of RD-C4ox-XXL, but not the isomeric RD-C4ox-XXXG. In addition, Y_{10} (m/z 903) and Y_{20} (m/z 771) ions were diagnostic fragments represented the loss

of non-deuterium-added hexaosyl (H_1P_0) and hexaosyl+pentaosyl (H_1P_1) units, respectively, which cannot be generated from the RD-C4ox-XXXG structure, and thus again confirmed the RD-C4ox-XXL structure.

a) RD-C4ox-XXL (m/z 1065.5)

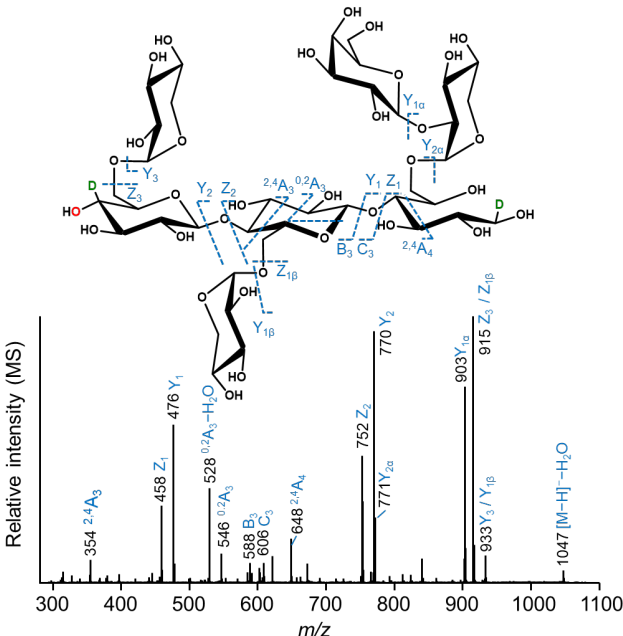
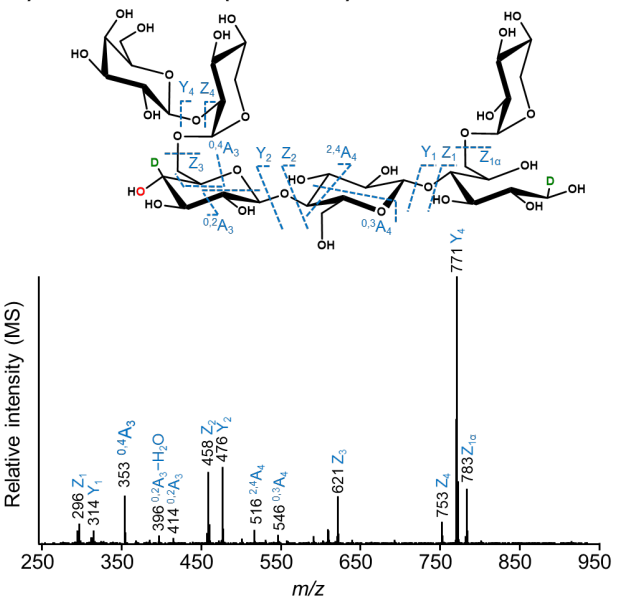


Fig. 3. Negative ion mode CID-MS² fragmentation patterns of reduced C4-oxidized TXG oligosaccharide. (a) RD-C4ox-XXL (m/z 1065.6) and (b) RD-C4ox-LGX (m/z 933.4). Only the structures with glucosyl non-reducing end were used to demonstrate their structural elucidation.

b) RD-C4ox-LGX (m/z 933.4)



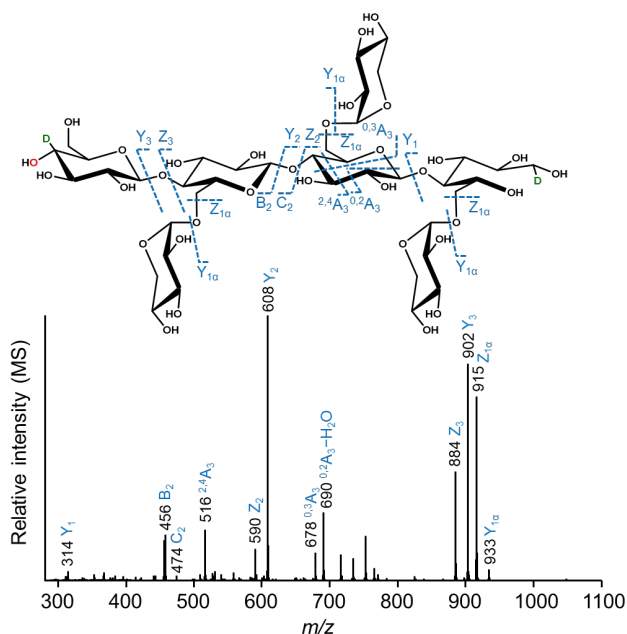
For RD-C4ox-LGX (m/z 933.4, $[M-H]^-$), having a reduced C4-oxidized L unit, identification was similar as described above (**Fig. 3b**). Fragmentation of RD-C4ox-LGX MS² showed as main fragments Y₂ (m/z 476), Z₂ (m/z 458), ^{0,2}A₃ (−H₂O) (m/z 414 and 396) and ^{2,4}A₄ (m/z 516). However, these four fragments can also represent the (reduced) C4-oxidized “GX” or “XG” in addition to the “L” unit (all three are H₂P₁). The L unit was determined to be at the non-reducing end side based on: i) Y₄ (m/z 771), Z₄ (m/z 753) and ^{0,4}A₃ (m/z 353) ions represented the loss of non-deuterium-added hexaosyl units (H₁P₀, m/z loss of 162 or 180). This m/z loss confirms the presence of a galactosyl unit but not a glucosyl unit, as terminal glucosyl units either have one deuterium in the non-reducing end (m/z loss of 163 or 181) or in the reducing end of an alditol form (m/z loss of 165 or 183). ii) Y₁ (m/z 314) and Z₁ (m/z 296) fragments suggested that an X unit was present in the alditol form of a reducing end side, and thus the galactosyl unit can only be present in the middle or non-reducing end. iii) Taken into account the parental m/z value (representing the H₄P₂ structure) and the TXG structure (“XXXG-type” building blocks), the L unit can only be present in the non-reducing end.

Other non-“XXXG-type” TXG oligosaccharides (DP>9; m/z >1389.7) eluted after 45 min (**Fig. 3**) in the *MtLPMO9H*-TXG digest again indicate a “substitution-tolerant” cleavage, which were absent in the *MtLPMO9F*-TXG digest. Due to the complexity, their exact structures will not be elucidated further.

In the reduced *MtLPMO9F*-TXG digest, several originally C4-oxidized “GXXX-type” building blocks were identified, for example RD-C4ox-GXXX and RD-C4ox-GXXL (**Fig. 2**). In the MS² spectrum of RD-C4ox-GXXX (m/z 1065.5, $[M-H]^-$, **Fig. 4a**), Y₃ (m/z 902) and Z₃ (m/z 884) fragments indicated the loss of the reduced C4-oxidized G unit from the non-reducing end. The absence of an ion with m/z of 903 (difference of 162 compared to the parent m/z) suggested the absence of a galactosyl unit, as described above. In addition, other fragments from either β-(1→4)-glycosidic bond cleavage (Y₁, m/z 314; B₂, m/z 456; C₂, m/z 474; Z₂, m/z 590 and Y₂, m/z 608) or cross-ring cleavage (^{2,4}A₃, m/z 516; ^{0,3}A₃, m/z 678 and ^{0,2}A₃−H₂O, m/z 690) further confirmed the structure of RD-C4ox-GXXX. RD-C4ox-GXXL (m/z 1227.6, $[M-H]^-$, **Fig. 4b**) was identified in a similar way as described for RD-C4ox-GXXX. The diagnostic fragments Y₁₀ (m/z 1065) and Y₃ (m/z 1064) represented the loss of a reduced C4-oxidized G unit from the non-reducing end and a galactosyl unit, respectively. Other fragments from either β-(1→4)-glycosidic bond cleavage (B₂, m/z 456; Z₁, m/z 458; Y₁, m/z 476; Z₂, m/z 752 and Y₂, m/z 770) or cross-ring cleavage (^{2,4}A₃, m/z 516 and ^{0,2}A₃−H₂O, m/z 690) further reflected the structure of RD-C4ox-GXXL. These identified “GXXX-type” C4-oxidized TXG oligosaccharides, together with the absence of non-“XXXG-type” ones, suggested a “substitution-intolerant” cleavage of TXG by *MtLPMO9F*. Notably, the GXXX-type” C4-oxidized TXG oligosaccharides were also present in the *MtLPMO9H*-TXG digest (**Fig. 2**), which may indicate that *MtLPMO9H* preferably cleaved at the “non-reducing end” site of G unit, though cleavage next to a substituted glucosyl unit was also identified to occur. This preference of *MtLPMO9H* differs from the previously characterized *NcLPMO9M*, which has been shown to preferentially cleave next to substituted glucosyl units (25). It is speculated that this preference difference is

caused by the presence of a carbohydrate binding module 1 (CBM1) in *MtLPMO9H*, but not in *NcLPMO9M*. CBM1 could influence the binding of the LPMO catalytic domain to the xyloglucan polymer, and thus (partially) alter the cleavage preference of *MtLPMO9H*.

a) RD-C4ox-GXXX (m/z 1065.5)



b) RD-C4ox-GXXL (m/z 1227.6)

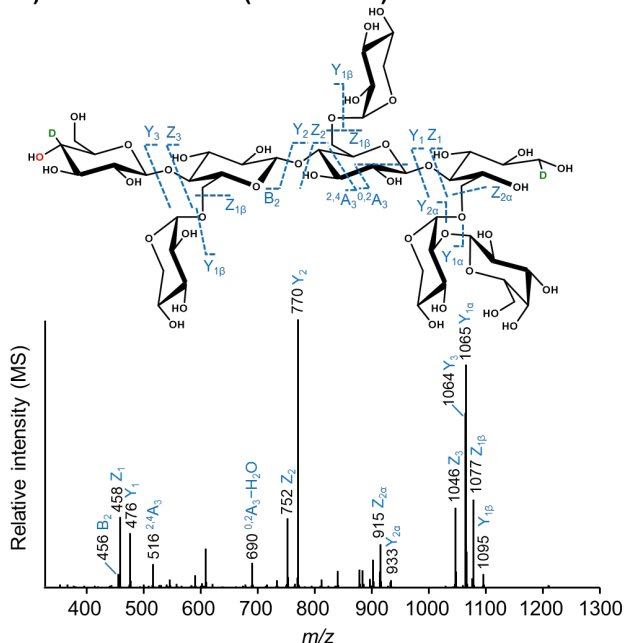


Fig. 4. Negative ion mode CID-MS² fragmentation patterns of reduced C4-oxidized TXG oligosaccharide. (a) RD-C4ox-GXXX (m/z 1065.5) and (b) RD-C4ox-GXXL (m/z 1227.6). Only the structures with glucosyl non-reducing end were used to demonstrate their structural elucidation.

3.3. MtLPMO9H and MtLPMO9F generates C4/C6-double oxidized xylogluco-oligosaccharides

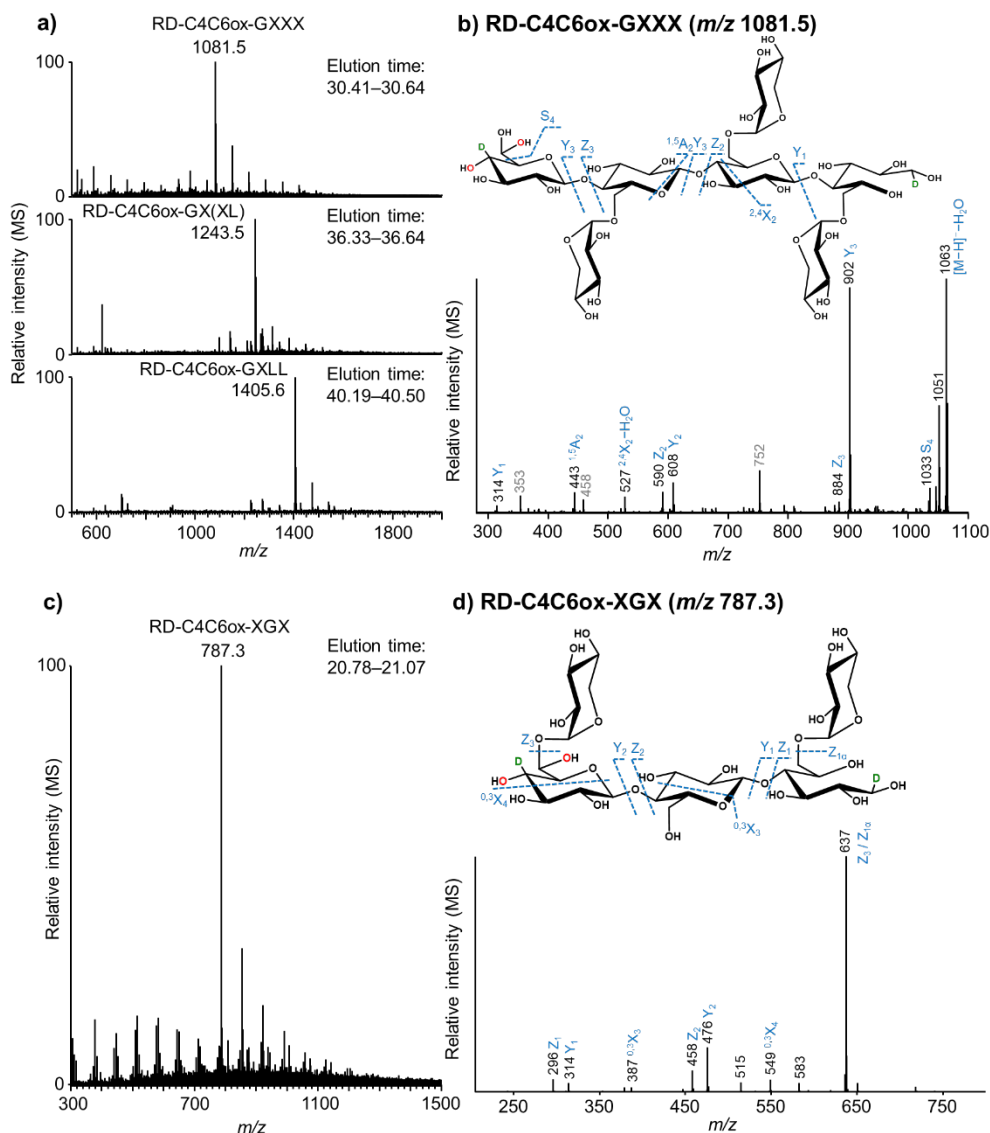


Fig. 5. (Negative ion mode) HILIC-ESI-MS and CID-MS² spectra of multiple (reduced) C4/C6-double oxidized TXG oligosaccharides. (a) MS spectra of m/z 1081.5 (elution time 30.41–30.64 min), m/z 1243.5 (elution time 36.33–36.64 min) and m/z 1405.6 (elution time 40.19–40.50 min), from top to bottom. (b) Negative ion mode CID-MS² fragmentation patterns of reduced C4/C6-double oxidized TXG oligosaccharide RD-C4C6ox-GXXX (m/z 1081.5). Fragments in grey color are tentatively proposed to be from other isomeric structures. (c) MS spectrum of m/z 787.3 (elution time 20.78–21.07 min). (d) Negative ion mode CID-MS² fragmentation patterns of reduced C4/C6-double oxidized TXG oligosaccharide RD-C4C6ox-XGX (m/z 787.3).

In addition to (reduced) C4-oxidized TXG oligosaccharides, (reduced) C4/C6-double oxidized TXG oligosaccharides were identified. **Fig. 5a** shows the full MS spectra of RD-C4C6ox-GXXX (m/z 1081.5, $[M-H]^-$), RD-C4C6ox-GX(XL) (m/z 1243.5, $[M-H]^-$) and RD-C4C6ox-GXLL (m/z 1405.6, $[M-H]^-$) detected in the *MtLPMO9F*-TXG digest (also present in the *MtLPMO9H*-TXG digest, but to a lesser extent (not shown)). The RD-C4C6ox-GXXX MS² spectrum (m/z 1081.5, $[M-H]^-$, **Fig. 5b**) is used as an example to demonstrate our structural elucidation. First, the reduced C4/C6-double oxidized unsubstituted glucosyl unit at the non-reducing end was identified by the m/z difference of 179 and 197, from Y_3 (m/z 902) and Z_3 (m/z 884) compared to the parent m/z (1081.5), respectively. The novel S-type ion (S_4 , m/z 1033), that indicates the loss of a C6-gem-diol structure in reduced C4/C6 double oxidized cello-oligosaccharides (19), was also found here. Other fragments, including Y_1 (m/z 314), $^{1,5}A_2$ (m/z 433), $^{2,4}X_2-H_2O$ (m/z 527), Z_2 (m/z 590) and Y_2 (m/z 608), indicated the loss of a reduced unsubstituted glucosyl unit originally having a C4-ketone and C6-gem-diol structures at the non-reducing end, further confirmed the structure of RD-C4C6ox-GXXX. Our previous study demonstrated that the C6-gem-diol structure, in the C4/C6-double oxidized cello-oligosaccharides, is formed via the oxygenation reaction (route I) of LPMOs, though the oxidation to C6-aldehyde followed by hydration to C6-gem-diol could not be excluded (route II) (19). Notably, in the RD-C4C6ox-GXXX MS² spectrum, several less abundant fragments (m/z 353, 458 and 752) were detected, which were, possibly, from the isomeric structure RD-C4C6ox-X(H_3P_2) (m/z 1081.5, $[M-H]^-$). As in these structures the C6-carbon atom is substituted with a xylosyl unit (at the non-reducing end G), C6-oxidation can here only occur via the oxygenation reaction (route I), and direct oxidation to a C6-aldehyde would not be possible (route II). In line with its substitution-tolerant cleavage behavior, *MtLPMO9H* generated a different C4/C6-double oxidized XGX unit (RD-C4C6ox-XGX, m/z 787.3, $[M-H]^-$, **Fig. 5c**), which provides further evidence for route I. Based on the MS and MS² spectra (**Fig. 5d**), insertion of a hydroxyl group on the substituted C6 atom was found in RD-C4C6ox-XGX. Therefore, we conclude that C6-oxidation of xyloglucan by AA9 LPMOs follows route I (oxygenation), and does not occur via route II (direct oxidation).

4. Conclusions

In this study, we characterized two AA9 *MtLPMOs*, having different active site segments, for their regioselectivity and xyloglucan cleavage profiles. We found that *MtLPMO9F* and *MtLPMO9H* both oxidatively cleaved cellulose and xyloglucan, while *MtLPMO9F* even displayed a broader substrate specificity. Using NaBD₄-reduction followed by HILIC-ESI-CID-MS/MS² analysis, we showed that *MtLPMO9F* released majorly C4-oxidized cello-oligosaccharides and C4/C6-double oxidized ones. In addition, C4/C6-double oxidized xylogluco-oligosaccharides were detected and formed via an oxygenation reaction. We further revealed that *MtLPMO9H* (+Seg1-Seg2) displayed xyloglucan "substitution-tolerant" cleavages, while *MtLPMO9F* (-Seg1+Seg2) displayed xyloglucan "substitution-intolerant" cleavages. These findings support the hypothesis that the configuration of active site segments in AA9 LPMOs correlate to their xyloglucan cleavage profiles.

5. References

1. Cherubini F. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Convers Manage*. 2010;51:1412-21.
2. Nanda S, Mohammad J, Reddy SN, Kozinski JA, Dalai AK. Pathways of lignocellulosic biomass conversion to renewable fuels. *Biomass Convers Biorefin*. 2014;4:157-91.
3. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, et al. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*. 2007;315:804-7.
4. Merino ST, Cherry J. Progress and challenges in enzyme development for biomass utilization. *Adv Biochem Eng Biot*. 2007;108:95-120.
5. Straathof AJ. Transformation of biomass into commodity chemicals using enzymes or cells. *Chem Rev*. 2014;114:1871-908.
6. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, et al. Fungal cellulases. *Chem Rev*. 2015;115:1308-448.
7. Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuel Bioprod Biorefin*. 2007;1:119-34.
8. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5:45.
9. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33:747-61.
10. Corrêa TL, dos Santos LV, Pereira GA. AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol*. 2016;100:9-16.
11. Berka RM, Grigoriev IV, Otilar R, Salamov A, Grimwood J, Reid I, et al. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol*. 2011;29:922-7.
12. Hinz SW, Pouvreau L, Joosten R, Bartels J, Jonathan MC, Wery J, et al. Hemicellulase production in *Chrysosporium lucknowense* C1. *J Cereal Sci*. 2009;50:318-23.
13. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490-D495.
14. Sabbadin F, Urresti S, Henrissat B, Avrova AO, Welsh LRJ, Lindley PJ, et al. Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes. *Science*. 2021;373:774.
15. Hangasky JA, Iavarone AT, Marletta MA. Reactivity of O₂ versus H₂O₂ with polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. 2018;115:4915-20.
16. Bissaro B, Varnai A, Rohr AK, Eijsink VGH. Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. *Microbiol Mol Biol Rev*. 2018;82.
17. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134:890-2.
18. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, et al. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289:2632-42.
19. Sun P, Laurent CVFP, Boerkamp JPV, van Erven G, Ludwig R, van Berkel WJH, Kabel MA. Regioselective C4 and C6 double oxidation of cellulose by lytic polysaccharide monooxygenases. *ChemSusChem*. Article in press.
20. Simmons TJ, Frandsen KEH, Ciano L, Tryfona T, Lenfant N, Poulsen JC, et al. Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat Commun*. 2017;8:1064.
21. Vaaje-Kolstad G, Forsberg Z, Loose JS, Bissaro B, Eijsink VG. Structural diversity of lytic polysaccharide monooxygenases. *Curr Opin Struct Biol*. 2017;44:67-76.
22. Frommhagen M, Westphal AH, van Berkel WJ, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbiol*. 2018;9:1080.
23. Frandsen KE, Lo Leggio L. Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes. *IUCrJ*. 2016;3:448-67.
24. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
25. Sun P, Laurent CVFP, Scheiblbrandner S, Frommhagen M, Kouzounis D, Sanders MG, et al. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnol Biofuels*. 2020;13:95.

26. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
27. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8:101.
28. Zhou H, Li T, Yu Z, Ju J, Zhang H, Tan H, et al. A lytic polysaccharide monooxygenase from *Myceliophthora thermophila* and its synergism with cellobiohydrolases in cellulose hydrolysis. *Int J Biol Macromol*. 2019;139:570-6.
29. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136:562-5.
30. Sun P, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng*. 2021;9:14124-33.
31. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
32. Frommhagen M, van Erven G, Sanders M, van Berkel WJH, Kabel MA, Gruppen H. RP-UHPLC-UV-ESI-MS/MS analysis of LPMO generated C4-oxidized gluco-oligosaccharides after non-reductive labeling with 2-aminobenzamide. *Carbohydr Res*. 2017;448:191-9.
33. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. *PLOS ONE*. 2018;13:e0202148.
34. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl Microbiol Biotechnol*. 2017;102:1281-95.
35. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
36. Karnaouri A, Muraleedharan MN, Dimarogona M, Topakas E, Rova U, Sandgren M, et al. Recombinant expression of thermostable processive MfEG5 endoglucanase and its synergism with MfLPMO from *Myceliophthora thermophila* during the hydrolysis of lignocellulosic substrates. *Biotechnol Biofuels*. 2017;10:126.
37. McNeil M, Darvill AG, Fry SC, Albersheim P. Structure and function of the primary cell walls of plants. *Annu Rev Biochem*. 1984;53:625-63.
38. Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr Res*. 2009;344:1879-900.
39. Hoffman M, Jia Z, Peña MJ, Cash M, Harper A, Blackburn II AR, et al. Structural analysis of xyloglucans in the primary cell walls of plants in the subclass *Asteridae*. *Carbohydr Res*. 2005;340:1826-40.
40. Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, et al. An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plant*. 1993;89:1-3.
41. Vincken JP, York WS, Beldman G, Voragen AG. Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiol*. 1997;114:9-13.
42. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Várnai A, et al. FgLPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Lett*. 2016;590:3346-56.
43. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al. *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
44. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnology*. 2011;7:214-23.
45. Sun P, Kadowaki MAS, Veersma R, Hilgers R, Magri S, Laurent CVFP, et al. Copper-dependent AA16 oxidoreductases boost cellulose active AA9 LPMOs from *Myceliophthora thermophila* C1. Chapter 8, this PhD thesis.
46. Agger JW, Isaksen T, Várnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A*. 2014;111:6287-92.

6. Supplementary information

6.1. Supplementary data

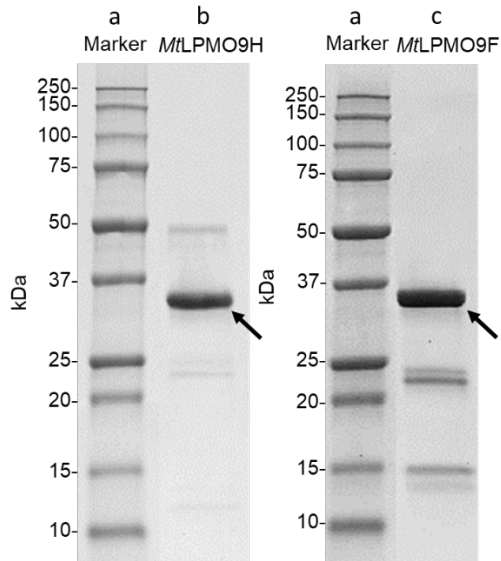


Fig. S1. SDS-PAGE analysis of (b) *MtLPMO9H*- and (c) *MtLPMO9F*-containing fractions after purification. The bands of *MtLPMO9H* and *MtLPMO9F* are indicated by arrows. In lanes a, the known molecular masses (kDa) of marker proteins are shown.

		Seg1									
<i>NcLPMO9M</i>	HGFVDNATI--GGQF-	YQFYQPYQDPFM----	GSPPDRISRKI-----	PGNGPVE-							
<i>FgLPMO9A</i>	HGHVESITV--GGTE-	YEGLNPGAAAFE----	NPRKELAAWFA----	TN-TDNGFVEP							
<i>MtLPMO9H</i>	HGHVSHIVV--NGVY-	YRNYDPTTDWYQ----	PNPPTVIGWTA----	ADQDNGFVEP							
<i>NcLPMO9C</i>	HTIFQKVS--NGAD-	QGQL-----	KGIRAP-----	ANNNPVT-							
<i>MtLPMO9E</i>	HAIFQKVS--NGAD-	QGSL-----	TGLRAP-----	NNNNPVQ-							
<i>MtLPMO9F</i>	HAIFQRVS--NGQD-	QGQL-----	KGVRAP-----	SSNSPIQ-							
		Seg2									
<i>NcLPMO9M</i>	-----DVTSLAIQCNA--D-S-A-----	P----	AKLHA-SAAAGSTVTLRWTI-								
<i>FgLPMO9A</i>	-----SAFGDADIICHR--G-A-E-----	N----	AVKSA-KVKAGEKITIKWDT-								
<i>MtLPMO9H</i>	-----NSFGTPDIIICHK--S A T-----	P----	GGGHA-TVAAGDKINIVWTP-								
<i>NcLPMO9C</i>	-----DVMSDDIICNA--V-T-M-----	K----	DSNVL-TVPAGAKVGHFWGHE								
<i>MtLPMO9E</i>	-----NVNSQDMICGQ--S-G-S-----	T----	SNTII-EVKAGDRIGAWYQHV								
<i>MtLPMO9F</i>	-----NVNDANMACNA--NIVYH-----	DNTII-	KVPAGARVGAWWQH V								
		Seg2									
<i>NcLPMO9M</i>	-----WP-DSHV-GPVITYMARC-PD----	TGCQDW-TPSASD-KVWF									
<i>FgLPMO9A</i>	-----WP-ESHK-GPVIDYLASCGS-----	AGCAKV-DKT--S-LKFF									
<i>MtLPMO9H</i>	-----EWP-ESHI-GPVIDYLAAC-N----	GDCETV-DKS--S-LRWF									
<i>NcLPMO9C</i>	IGGAAGPNDADN--PIAASHK-GPIMVYLAKV-D-----	NAATT-GTS--G-LKWF									
<i>MtLPMO9E</i>	IGGAQFPNDPDN--PIAKSHK-GPVMAYLAKV-D-----	NAATA-SKT--G-LKWF									
<i>MtLPMO9F</i>	IGGPQGANDPDN--PIAASHK-GPIQVYLAKV-D-----	NAATA-SPS--G-LKWF									
		Seg3									
<i>NcLPMO9M</i>	KIKEGGRE-GT----S-----	NVWA-ATPLMTAP--	ANYEYAI	P	S	C	L	K	P		
<i>FgLPMO9A</i>	KIAEAGMT-SG----S-----	GKFA-SDDLIAAG--	NTWEVTPTS	I	K	A					
<i>MtLPMO9H</i>	KIDGAGYD-KA----S-----	A----	GRWA-ADALRANG--	NSWLVI	P	S	D	L	K	A	
<i>NcLPMO9C</i>	KVAEAGLS-----N-----	GKWA-VDDLIAANN--	GWSYFDMPTC	I	A	P					
<i>MtLPMO9E</i>	KIWEDTFN-PS----T-----	KTWG-VDNLI	NNN--	GWYFNL	P	Q	C	I	A	D	
<i>MtLPMO9F</i>	KVAERGLN-----N-----	GVWA-VDELI	ANN--	GWHYFDL	P	S	C	V	A	P	
		Seg4									
<i>NcLPMO9M</i>	GYLLVRHEIIALHSAYSYP-----	GAQFYPGCHQLQVTGSG----	TKTPSS-G	L	-						
<i>FgLPMO9A</i>	GNVYLRHEIIALHAAGQEN-----	GAQNPQC	FNLEVESDG----	TAE-PA-	G	-					
<i>MtLPMO9H</i>	GNVYLRHEIIALHGAQSPN-----	GAQAYPQC	INLRVTGGG----	SNLPSG-	-						
<i>NcLPMO9C</i>	GQYLMRAELIALHNAGSQA-----	GAQFYIGCAQ	INVTGGG----	SASPSN-	T	-					
<i>MtLPMO9E</i>	GNVYLLRVEVLALHSAYSQG-----	QAQFYQSCAQ	INVSGGG----	SFTPAS-	T	-					
<i>MtLPMO9F</i>	GQYLMRVELLALHSASSPG-----	GAQFYMGCAQ	IEVTGSG----	TNSGSD-	F	-					
		Seg5									
<i>NcLPMO9M</i>	-----VSFPGAYKSTDP-GVTYDAYQ-----	AA--TYT-I	PGP-AV-FTC								
<i>FgLPMO9A</i>	-----VAGTSLYTASEK-GIVFDLYN-----	NP-T-SYP-I	PGP-KMNIAG								
<i>MtLPMO9H</i>	-----VAGTSLYKATDP-GILFNPYV-----	SS-P-DYT-V	PGP-AL-IAG								
<i>NcLPMO9C</i>	-----VSFPGAYSASDP-GILINITYGGSGKTDNGG-K-PYQ-I	PGP-AL-FTC									
<i>MtLPMO9E</i>	-----VSFPGAYSASDP-GILINITYGATGQPDNNG-Q-PYT-APG-AP-I	SC									
<i>MtLPMO9F</i>	-----VSFPGAYSANDP-GILLSIYDSSGKPNNGG-R-SYP-I	PGP-RP-I	SC								

Fig. S2. Structure-based amino acid sequence alignment of catalytic domains of *NcLPMO9M*, *FgLPMO9A*, *MtLPMO9H*, *NcLPMO9C*, *MtLPMO9E* and *MtLPMO9F*.

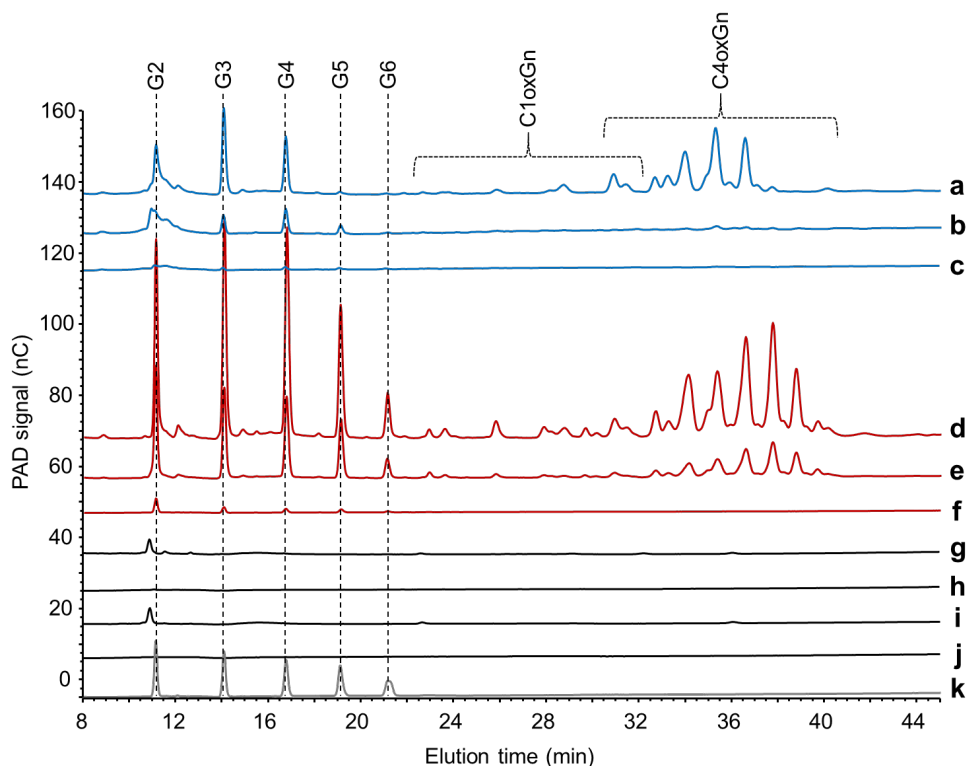


Fig. S3. HPAEC chromatograms of generated oligosaccharides from RAC incubations with *MtLPMO9H* and *MtLPMO9F* with and without *MtAA16A*. In the presence of Asc, incubation of RAC with *MtLPMO9H* with and without *MtAA16A* (d and e) resulted in the release of both C1- (C1oxGn) and C4-oxidized oligosaccharides (C4oxGn), while *MtLPMO9F* with and without *MtAA16A* (a and b) resulted in the generation of mainly C4oxGn. Cello-oligosaccharides standard mixture (k) is also shown. Control reactions were only RAC (j), RAC with Asc (i), RAC with *MtAA16A* in the presence (g) and absence (h) of Asc, *MtLPMO9H* incubated with RAC in the absence of Asc (f) and *MtLPMO9F* incubated with RAC in the absence of Asc (c). *MtAA16A* has been shown to boost AA9 *MtLPMOs* (Chapter 8), which is also observed for *MtLPMO9F* and *MtLPMO9H* here.

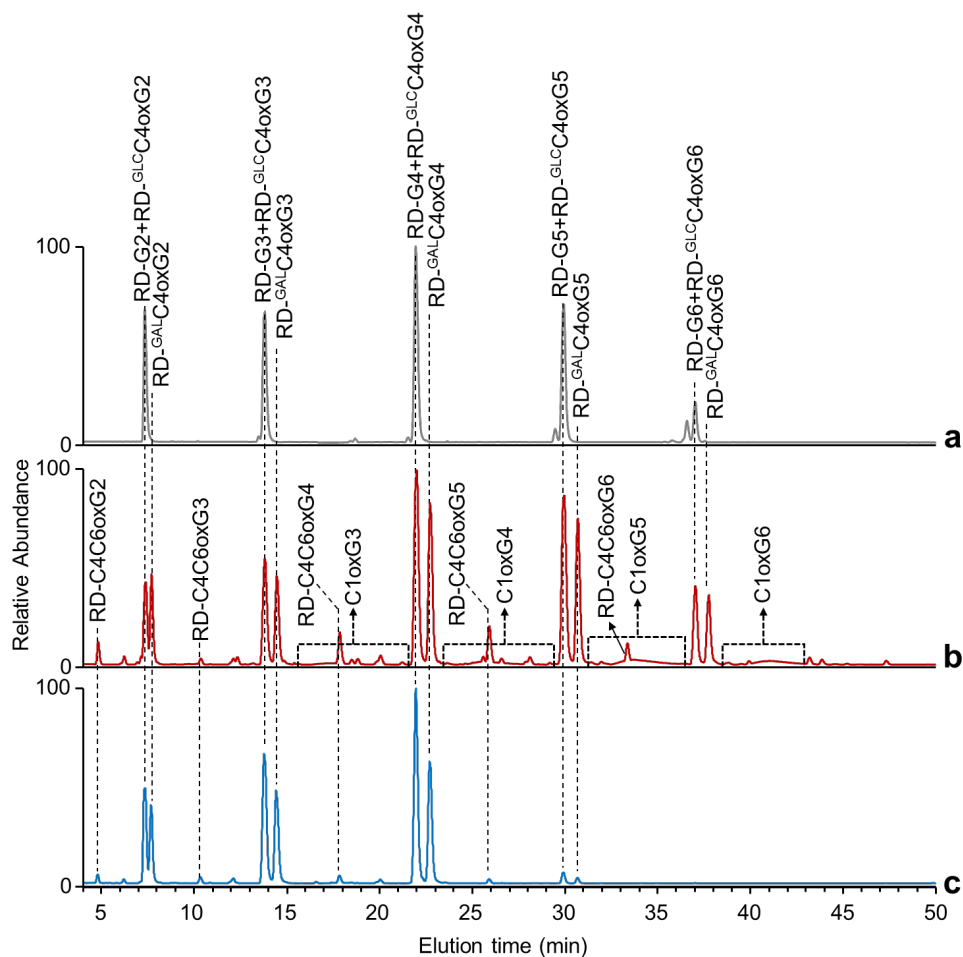


Fig. S4. HILIC-ESI-MS base-peak chromatograms of reduced (a) cello-oligosaccharide standard mixture (RD-G₂-6), (b) *MtAA16A-MtLPMO9H*-RAC digest in the presence of Asc and (c) *MtAA16A-MtLPMO9F*-RAC digest in the presence of Asc. RD-G_n: reduced non-oxidized cello-oligosaccharides; RD-GLC C4oxG_n: reduced C4-oxidized cello-oligosaccharides including the glucosyl end; RD-GLC C4oxG_n: reduced C4-oxidized cello-oligosaccharides including the galactosyl end; RD-C4C6oxG_n: reduced C4/C6-double oxidized cello-oligosaccharides.

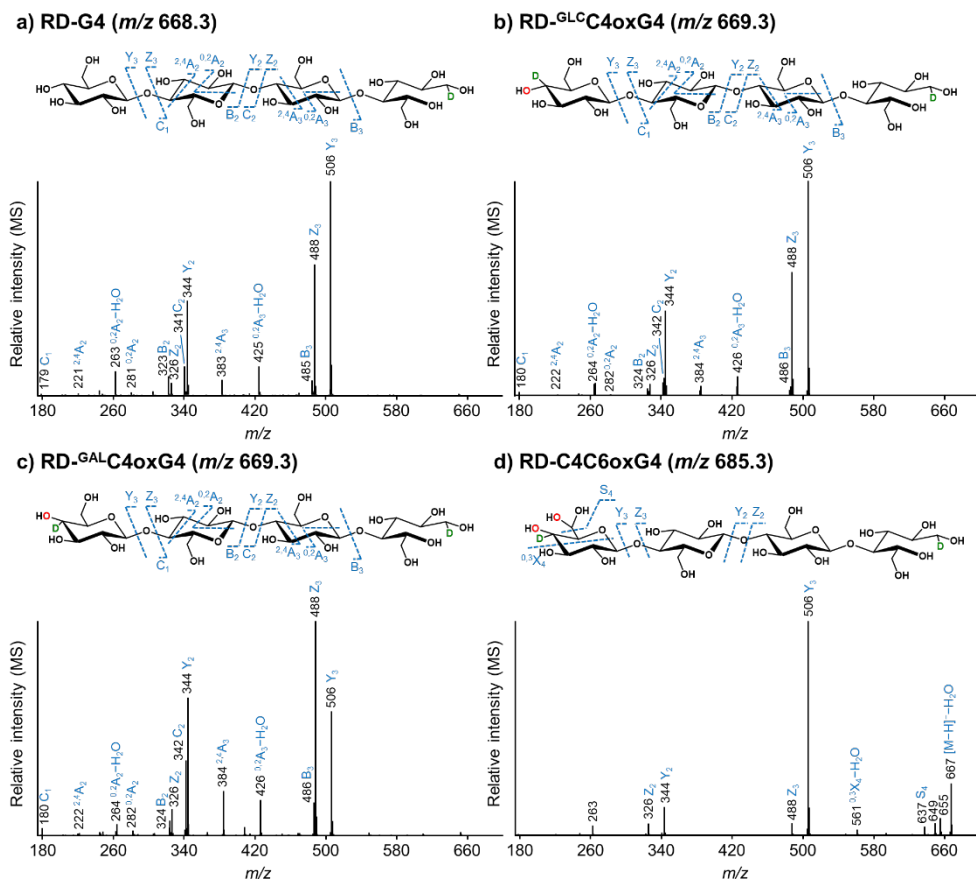


Fig. S5. Negative ion mode CID-MS² fragmentation patterns of DP4 of (a) reduced non oxidized cello-oligosaccharide (RD-G4, m/z 668.3), (b) reduced C4-oxidized cello-oligosaccharide with glucosyl residue (with one additional deuterium ion) at the non-reducing end side (RD-GLC4oxG4, m/z 669.3), (c) reduced C4-oxidized cello-oligosaccharide with galactosyl residue (with one additional deuterium ion) at the non-reducing end side (RD-GALC4oxG4, m/z 669.3) and (d) reduced C4/C6-double-oxidized cello-oligosaccharide (RD-C4C6oxG4, m/z 685.3) in the MtAA16A-MtLPMO9F-RAC digest.

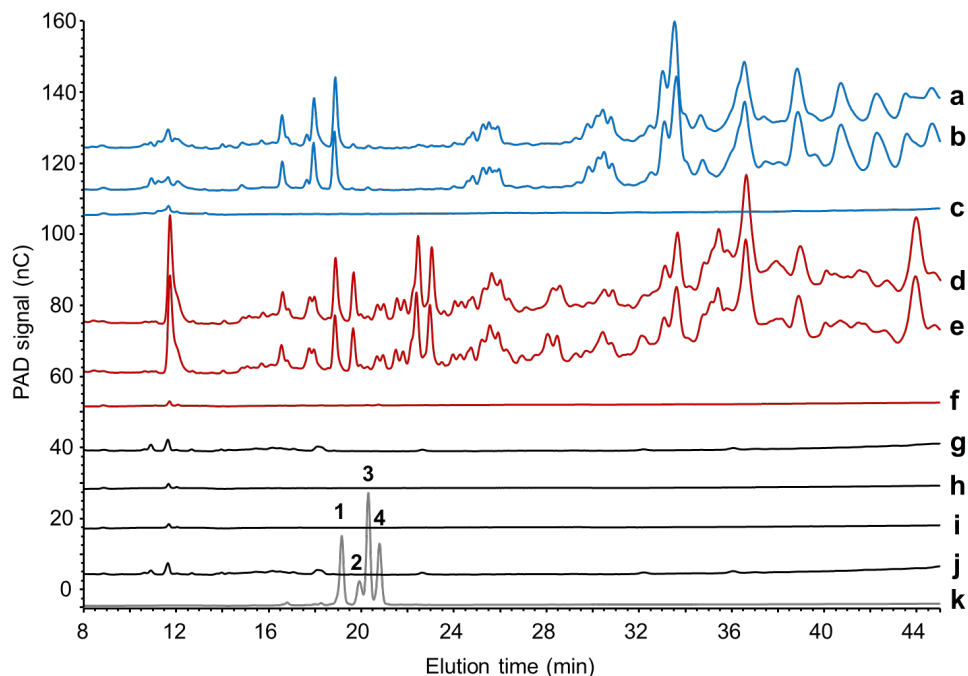


Fig. S6. HPAEC chromatograms of generated oligosaccharides from TXG incubations with *MtLPMO9H* and *MtLPMO9F* with and without *MtAA16A*. In the presence of Asc, incubation of TXG with *MtLPMO9H* (with and without *MtAA16A* (d and e)) and *MtLPMO9F* (with and without *MtAA16A* (a and b)) resulted in the release of series of (oxidized) TXG oligosaccharides. TXG oligosaccharide standard (k) is also shown (1=XXXG, 2=XLXG, 3=XXLG and 4=XLLG). Control reactions were only TXG (j), TXG with Asc (i), TXG with *MtAA16A* in the presence (g) and absence (h) of Asc, *MtLPMO9H* incubated with TXG in the absence of Asc (f) and *MtLPMO9F* incubated with TXG in the absence of Asc (c).

Chapter 8

Copper-dependent AA16 oxidoreductases boost cellulose active AA9 LPMOs from *Myceliophthora thermophila* C1

Peicheng Sun¹, Marco A.S. Kadowaki², Romy J. Veersma¹, Roelant Hilgers¹, Silvia Magri², Christophe V.F.P. Laurent^{3,4}, Roland Ludwig³, David Cannella², Leila Lo Leggio⁵, Willem J.H. van Berkel¹ and Mirjam A. Kabel¹

¹Laboratory of Food Chemistry, Wageningen University & Research, Wageningen, The Netherlands

²PhotoBioCatalysis Unit (BTL) École interfacultaire de Bioingénieurs (EIB), Université Libre de Brussels, Brussels, Belgium

³Biocatalysis and Biosensing Laboratory, Department of Food Science and Technology, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

⁴Institute of Molecular Modelling and Simulation, Department of Material Sciences and Process Engineering, BOKU–University of Natural Resources and Life Sciences, Vienna, Austria

⁵Biological Chemistry Section, Department of Chemistry, University of Copenhagen, Copenhagen, Denmark

Abstract

Copper-dependent oxidoreductases, classified into Auxiliary Activities (AA) families in the Carbohydrate-Active enZymes (CAZy) database, have been shown to play important roles in the enzymatic conversion of lignocellulosic plant biomass. The AA16 family was discovered in 2019 and the only member characterized so far is reported to be an C1-oxidizing lytic polysaccharide monooxygenase (LPMO) active on cellulose. In this work, we found that another AA16 member from *Myceliophthora thermophila* C1 (*MtAA16A*) did not display any oxidative cleavage of carbohydrate substrates. Interestingly, *MtAA16A* largely boosted the degradation of cellulose by four AA9 *MtLPMOs*. The same boosting effect was observed for AA16A from *Aspergillus nidulans* (*AnAA16A*). On the contrary, cellulose degradation by three AA9 LPMOs from *Neurospora crassa* (*NcLPMOs*) was not boosted by *MtAA16A* and *AnAA16A*. The boosting effect was found to relate to the production of H₂O₂ by AA16. It is proposed that H₂O₂ delivery to the AA9 *MtLPMOs* is facilitated by protein-protein interaction and that this interaction is absent between AA16s and AA9 *NcLPMOs*. Our findings provide new insights into the biological functions of copper-dependent enzymes and contribute to a further understanding of synergism within fungal enzymatic systems in degrading lignocellulose.

Keywords

Biomass; Lignocellulose; Biorefinery; Cellulose; Copper-dependent oxidoreductase; AA16; AA9 LPMO; Hydrogen peroxide

1. Introduction

Transition from a fossil-based society to a more sustainable one drives full valorization of lignocellulose-rich agricultural and forestry side-streams for the production of biofuels, biomaterials and biochemicals (1-3). Hereto, enzymatic driven degradation of lignocellulose (e.g., cellulose and hemicellulose) to their fermentable monosaccharides is an essential step (4-8). Enzymatic (hemi-)cellulose degradation includes the synergistic action of a wide range of (hemi-)cellulases and essential lytic polysaccharide monooxygenases (LPMOs) (9-11). LPMOs are currently classified as "Auxiliary Activities" (AA) families 9-11 and 13-17 in the Carbohydrate-Active enZymes (CAZy) database (<http://www.cazy.org>. (12)) (13, 14). The fungal AA9 family is currently the largest LPMO family. AA9 LPMOs have been characterized to oxidatively cleave cellulose, and to a certain extent hemicelluloses (15, 16). The AA16 family is a more recently proposed LPMO family with one characterized member from *Aspergillus aculeatus* (AaAA16) (17). AaAA16 has been indicated as C1-cellulose active LPMO (17), though its activity is much lower than that of C1-cellulose active LPMOs, such as MtLPMO9B (18) and NcLPMO9F (19). Hence, this study aims to get insight in the catalytic properties of other AA16 members.

AA16 members share common features with other LPMOs, for instance a copper-dependent active site coordinated by two histidines and a tyrosine (sometimes a phenylalanine in AA10) residue (16, 20, 21). This coordination is recognized as a "histidine brace", which is conserved in all LPMOs (21-24). Although the catalytic mechanism of LPMOs is not completely elucidated, it is well accepted that the catalytic cycle starts with a so-called "priming reduction" of Cu(II) to Cu(I) by external electron donors (15, 25). The external electron donor can be a chemical reductant, such as ascorbic acid (Asc) or pyrogallol (Pyg), or an enzyme, for example cellobiose dehydrogenase (CDH) (26-28). After this priming reduction, the catalytic reaction can follow two routes depending on the co-substrate, O₂ or H₂O₂ (25). The mechanistic details of these routes are still under debate as extensively reviewed elsewhere (15, 16, 20, 21, 23, 25, 29-31). In brief, in the O₂-dependent route, molecular oxygen binds to Cu(I) to form a series of copper-superoxide and hydroperoxide intermediates after receiving two electrons and two protons (30, 32). The highly reactive hydroperoxide intermediate hydroxylates the C1- and/or C4-carbon position of a carbohydrate substrate and generates Cu(II) for the next catalytic cycle (**Fig. 1, route 1**). Other studies suggested that two protons and only one electron are required and that Cu(II) is formed after one catalytic cycle (13, 21, 33-35). In the absence of a carbohydrate substrate, superoxide (or hydroperoxide) is released, resulting in the production of H₂O₂ and regeneration of the Cu(II) state (**Fig. 1, route 2**) (25, 32). The H₂O₂-dependent route (**Fig. 1, route 3**) proposed by recent studies suggests a peroxygenase reaction rather than a monooxygenase reaction (25). After the priming reduction, H₂O₂ binds to Cu(I) to form a Cu(I)···H₂O₂ intermediate, which serves as hydroxyl donor to hydroxylate the carbohydrate substrate, similarly as suggested for route 1 (**Fig. 1**) (25). The glycosidic bond next to the hydroxylated C-atom in the carbohydrate substrate is destabilized due to the additional oxygen atom, resulting in

a subsequent elimination reaction and, eventually, bond cleavage. (16, 26, 34). The products are oxidized and non-oxidized carbohydrates. In both O_2 and H_2O_2 dependent routes, the formation of the $Cu(I) \cdots H_2O_2$ intermediate is key, suggesting that H_2O_2 plays a critical role in the LPMO reaction. Even, H_2O_2 has been suggested to be the preferred co-substrate over O_2 , as turnover numbers obtained with H_2O_2 are in certain cases more than one order of magnitude higher compared to these obtained with O_2 (25, 35, 36). On the other hand, H_2O_2 induces oxidative damage of the copper-active site resulting in self-inactivation. Hence, the role of H_2O_2 is a balance between activation and deactivation, and at a delicate equilibrium, in such LPMO reactions (25, 31, 35-37).

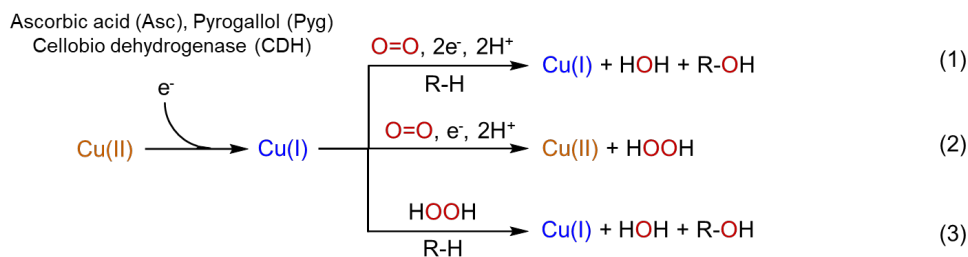


Fig. 1. Proposed catalytic mechanisms for LPMO reactions using O_2 ((1) and (2)) or H_2O_2 (3). In both O_2 and H_2O_2 dependent routes, C1- and/or C4-carbon position of carbohydrate ($R-H$) substrate is hydroxylated ($R-OH$) and a $Cu(I)$ is generated for the next catalytic circle. Adapted from Wang et al. (30).

In this study, we investigated the catalytic behavior of an AA16 protein, homologously expressed and produced from *Myceliophthora thermophila* C1 (*MtAA16A*). *MtAA16A* did not oxidatively cleave any carbohydrate substrates, but largely boosted AA9 *MtLPMOs* in degrading cellulose. The same boosting effect was observed for a similar AA16 enzyme from *Aspergillus nidulans* (*AnAA16A*). We suggest that AA16 proteins are not true LPMOs, but oxidoreductases assisting LPMOs in degrading lignocellulose. We have investigated the underlying mechanisms for our observations, which relate to H_2O_2 production and further H_2O_2 delivery via protein-protein interaction is hypothesized.

2. Materials and methods

2.1. Materials

Syringol, pyrogallol and ammonium acetate were purchased from Sigma-Aldrich (St. Louis, MO, USA). Cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexasose were purchased from Megazyme (Bray, Ireland). Regenerated amorphous cellulose (RAC) was prepared from Avicel® PH-101 (Sigma-Aldrich) as described previously (38). Ascorbic acid (Asc) and sorbitol were purchased from VWR International (Radnor, PA, USA). Other aromatic compounds used were purchased from Sigma-Aldrich or VWR International. Other carbohydrate substrates used were purchased from Sigma-Aldrich or Megazymes. All water used was produced by a Milli-Q system (Millipore, Molsheim, France), unless mentioned otherwise.

2.2. Expression, production and purification of *MtAA16A*

The gene encoding *MtAA16A* (MYCTH_2306267, UniProt ID: G2QH80) was homologously expressed in a low protease/low (hemi-) cellulose producing *Myceliophthora thermophila* C1 strain as described elsewhere (39, 40).

MtAA16A was purified by three subsequent chromatographic steps. Crude *MtAA16A*-rich fermentation broth was filtrated and dialyzed against 10 mM potassium phosphate buffer pH 7.6 before chromatographic purification. The dialyzed *MtAA16A* was purified by anion exchange chromatography (AEC) followed by size exclusion chromatography (SEC). Purification settings and elution programs of AEC and SEC have been described previously (41). The SEC-purified *MtAA16A*-containing fractions were further purified by cation exchange chromatography (42) on an ÄKTA-Micro preparative chromatography system (GE Healthcare). *MtAA16A*-containing fractions were loaded on a Resource Q column (30 × 16 mm internal diameter, GE Healthcare) pre-equilibrated with 10 mM sodium acetate buffer pH 3.8 (eluent A). The unbound fraction was firstly removed (one column volume). Eluent B was 10 mM sodium acetate buffer (pH 3.8) containing 500 mM NaCl (eluent B). Elution (flow rate of 1 mL/min) was performed as follows: from 0% to 30% B in two column volumes; 30% B for one column volume; next 30% to 100% B over two column volumes and finally 100% B for four column volumes. All fractions were collected and immediately stored on ice. Peak fractions (based on UV 280 nm) were adjusted to an approximate concentration of 2 mg/mL determined by BCA assay and analyzed by SDS-PAGE, as described previously (41), to determine the *MtAA16A* fractions. CEC-purified *MtAA16A*-containing fractions were combined and used as final enzyme stock solution. All CEC-purified *MtAA16A* fractions were aliquoted into 500 µL size and stored at -80 °C.

2.3. Cu(II) saturation of *MtAA16A*

Cu(II) saturation of *MtAA16A* was performed according to Loose et al. (43) with modifications. A pure *MtAA16A* stock solution (1 mg/mL, 500 µL) was incubated with a three-fold molar excess of Cu(II)SO₄ in 50 mM ammonium acetate pH 5.0 for 30 min at 25 °C under shaking at 600 rpm (Eppendorf ThermoMixer® C, Eppendorf, Hamburg, Germany). Excess Cu(II) was removed by a five-cycle washing-out procedure. For each washing step, 500 µL Cu(II)-saturated *MtAA16A* was concentrated below 50 µL using Amicon® Ultra-0.5 centrifugal filters (Sigma-Aldrich, St. Louis, Missouri, USA), and subsequently brought back to 500 µL by adding 50 mM ammonium acetate pH 5.0. The final concentration of excess Cu(II) was calculated lower than 0.7 pM. In this study, *MtAA16A* represents the Cu(II) saturated form, unless mentioned otherwise.

2.4. Incubation of AA9 LPMOs with AA16 enzymes on RAC

General incubation settings were 50 mM ammonium acetate (pH 5.0) buffer, 2 mg/mL RAC and 1 µM AA9 LPMO and/or 1 µM AA16 enzyme (*MtAA16A* and *AnAA16A*; expression, production and purification of *AnAA16A* is described in Supplementary information). The incubation has been performed in four ways:

i) Incubations (reaction volume of 300 μ L each) containing 1 μ M *MtLPMO9B* with and without 1 μ M *MtAA16A* or *AnAA16A* were done in the presence of 1 mM Pyg at 30 °C for 16 h;

ii) To monitor the generation of H_2O_2 and the oxidized cello-oligosaccharides over time, incubations (reaction volume of 1200 μ L each) containing 1 μ M *MtLPMO9B* with and without 1 μ M *MtAA16A* or *AnAA16A* were performed in the presence of 1 mM Pyg and 300 μ L of sample was taken out at 2, 4 and 6 h;

iii) Incubations (reaction volume of 600 μ L each) were initiated by adding 12 μ L aliquots of different H_2O_2 stock solutions (0, 500, 1250, 2500, 5000 and 10000 μ M) to reach H_2O_2 concentrations of 0, 10, 25, 50, 100 and 200 μ M, in the presence of 1 mM Asc and 1 μ M *MtLPMO9B* or *NcLPMO9M*. Every 1 h, 12 μ L of the different H_2O_2 stock solutions were added to the incubations (6 additions in total in the first 5 h). Two more samples containing 1 μ M *MtLPMO9B* and 1 μ M *MtAA16A* or *AnAA16A* were included in the presence of 1 mM Asc. The final reaction volume in these two incubations was adjusted (by adding water) to give the same enzyme concentrations as in incubations with H_2O_2 addition. At 6 h, 300 μ L of sample was taken out from each incubation and the rest of incubation was till 16 h.

iv) Incubations (reaction volume of 300 μ L each) were performed with 1 μ M *MtLPMO9E*, *MtLPMO9I*, *MtLPMO9H*, *NcLPMO9C*, *NcLPMO9M* or *NcLPMO9F* with and without 1 μ M *MtAA16A* or *AnAA16A* in the presence of 1 mM or 50 μ M Asc at 30 °C for 16 h.

All supernatants from the above incubations were collected and stored at -20 °C for further analysis. All incubations were performed in duplicate.

2.5. Determination of H_2O_2 by the ferric-xylenol orange assay

H_2O_2 level in the supernatants after 2, 4 and 6 h incubation of *MtLPMO9B* with/without *MtAA16A* and *AnAA16A* on RAC in the presences of Pyg was freshly determined by using the purchased Peroxide Assay Kit (Catalog Number: MAK311, Sigma-Aldrich). The assay was performed by following the protocol provided by the manufacture. Firstly, H_2O_2 standards (0, 3, 6, 9, 12, 18, 24 and 30 μ M) and detection reagent (mixing 1 volume of Reagent A with 100 volumes of Reagent B) were freshly prepared. Afterwards, 40 μ L of undiluted supernatants from the incubations and H_2O_2 standards was added into separate wells of a 96-well plate. Subsequently, 200 μ L of detection reagent was added into wells and the reaction was incubated for 30 min at room temperature. The absorbance of each sample at 585 nm was determined in a spectrophotometer. The H_2O_2 levels in samples were calculated based on a calibration curve generated by H_2O_2 standards. All measurements were performed in duplicates.

2.6. HPAEC-PAD analysis for oligosaccharide profiling

All supernatants from the incubations of AA9 LPMO (in the presence and absence of AA16s) with RAC were analyzed by high performance anion exchange chromatography (HPAEC). The analysis was performed on an ICS-5000 system (Dionex, Sunnyvale, CA, USA) equipped with a CarboPac PA-1 column (2 mm ID \times 250 mm; Dionex) in

combination with a CarboPac PA guard column (2 mm ID × 50 mm; Dionex). The system was further equipped with pulsed amperometric detection (PAD). Two mobile phases were (A) 0.1 M NaOH and (B) 1 M NaOAc in 0.1 M NaOH. The column temperature was set at 20 °C. The elution profile applied has previously been described (41, 44). Samples were diluted five times before analysis.

3. Results and discussion

3.1. *MtAA16A*: molecular mass, N-glycosylation and methylation of N-terminal histidine

MtAA16A was homologously produced by *M. thermophila* C1 and purified from the culture broth supernatant by three chromatographic techniques. The final *MtAA16A* preparation showed a major band (>90%, based on the band intensity) at 27 kDa in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, Fig. S1). Since the predicted molecular mass of *MtAA16A* based on the amino acid sequence without signal peptide (Fig. S2) is 18.4 kDa, glycosylation of *MtAA16A* was expected, as observed with other homologously produced MtLPMOs (44). Indeed, after treatment of *MtAA16A* with (*N*-acetyl- β -glucosaminyl)asparagine amidase (PNGase F), a major band at 19 kDa (Fig. S1) remained, indicating that *MtAA16A* was N-glycosylated.

Typical for homologously expressed fungal LPMOs is the methylated N-terminal histidine, of which the methylation is suggested to play a role in protection against auto-oxidation of the copper histidine braced active site (22, 45, 46). Reversed phase chromatography coupled to mass spectrometry (MS/MSⁿ) of a tryptic digest revealed that the N-terminal histidine of *MtAA16A* indeed was methylated (Met-H; Fig. S3). Further identification of peptides in the *MtAA16A* tryptic digest confirmed that the amino acid sequence of the *MtAA16A* protein was in accordance with the prediction based on gene annotation.

3.2. *MtAA16A* does not oxidatively cleave carbohydrates, but oxidizes syringol-like compounds

Since the AA16 family has been suggested to comprise catalytic LPMO-like enzymes, though this suggestion was based on a rather low C1-oxidative cleavage of cellulose observed for only one AA16 candidate (*AaAA16*; (17)), the catalytic LPMO-like behavior of *MtAA16A* was tested (Table S1). *MtAA16A* did not oxidatively cleave cellulose and none of the other carbohydrates tested, including cellopentaose, celohexaose, chitin, pectin, hemicelluloses and their combinations, could be (oxidatively) degraded (Table S1). Oxidative cleavage was neither observed by renewing Cu(II) saturation of *MtAA16A*, nor by varying the type of reducing agent or adding H₂O₂, and also not by increasing substrate or enzyme concentrations (Table S1). Therefore, we concluded that *MtAA16A* has no typical LPMO-like catalytic action towards cellulose and other (oligo-)polysaccharides tested.

At the same time, we asked ourselves whether *MtAA16A* actually is an enzyme or just a non-catalytic copper containing protein (47, 48). In 2018, Breslmayr and coworkers

developed the so-called 2,6-DMP assay based on the LPMO-driven conversion of 2,6-dimethoxyphenol (syringol) in the presence of H_2O_2 (49). To get further insight in its possible catalytic behavior, *MtAA16A* was screened by using the 2,6-DMP assay. AA9 LPMOs of the same fungus such as *MtLPMO9B* (44), *MtLPMO9E* (41, 50), *MtLPMO9I* (41) and *MtLPMO9H* (18, 51) served as reference and results are shown in Fig. S4. The 2,6-DMP-assay activity of all non-Cu(II)-saturated *MtLPMO9s* and non-Cu(II)-saturated *MtAA16A* ranged between 0.27–0.56 U/g. However, Cu(II)-saturated *MtAA16A* showed a much higher activity of 4.14 U/g, while Cu(II)-saturation of the *MtLPMO9s* did not increase their activity in the 2,6-DMP assay (data not shown). These results indicated that *MtAA16A* is a copper-dependent enzyme able to oxidize syringol (Syr) in the presence of H_2O_2 .

In addition to Syr, a number of aromatic compounds were tested as substrate for (Cu(II)-saturated) *MtAA16A* (Table S2). As analyzed by UHPLC-MS, 10 out of 28 compounds (marked in green in Table S2) were found to act as substrate for (Cu(II)-saturated) *MtAA16A*.

3.3. AA16 enzymes boost the *MtLPMO9B* oxidative cleavage of cellulose

We noticed that amongst the *MtAA16A* active aromatic compounds, several of them, such as pyrogallol (Pyg), catechol and sinapic acid, were also reported to act as electron donors for *MtLPMO9B* in oxidatively cleaving cellulose (44, 52). Out of curiosity, we mixed *MtAA16A* (Cu(II)-saturated) with *MtLPMO9B*, and with regenerated amorphous cellulose (RAC) and Pyg. *MtLPMO9B*-RAC(-Pyg) digests, with and without *MtAA16A*, were analyzed by HPAEC-PAD (**Fig. 2**). Firstly, in the controls no (oxidized) cellulose degradation products were detected: i) the incubation of *MtLPMO9B* with RAC in the absence of Pyg; ii) the *MtAA16A*-RAC digests, in absence or presence of Pyg (**Fig. 2a** and Fig. S5). Thus, it was concluded that both *Mt*-enzymes were free of hydrolytic side-activity and that, as mentioned in the text above, *MtAA16A* was not active towards RAC. Surprisingly, in the *MtAA16A* and Pyg containing *MtLPMO9B*-RAC digest (**Fig. 2a**), much more non-oxidized (Glc) and oxidized cello-oligosaccharides (GlcOx) were released, than in the same digest without *MtAA16A*. A comparable boosting effect was found by using another AA16 family member from *Aspergillus nidulans* (*AnAA16A*; expression, production and purification described in Supplementary information) (**Fig. 2a** and Fig. S5).

In an attempt to find out the underlying mechanism of the observed boosting effect, we hypothesized that, in the presence of reducing agent, AA16 enzymes produce H_2O_2 , which efficiently acts as co-substrate for the *MtLPMO9B* to oxidatively degrade RAC (**Fig. 1; route 3**). This hypothesis was based on two earlier findings. The first one is that *AaAA16* is reported to produce H_2O_2 in the presence of the reducing agent, which is a comparable reaction to other copper-dependent LPMOs (**Fig. 1, route 2**) (17, 25, 32). Production of H_2O_2 after Cu(II) reduction can even occur in presence of cellulose, which is not a substrate for the AA16 enzymes as was shown in this work. Secondly,

exogenous addition of H_2O_2 (1–200 μM ; **Fig. 1; route 3**) has been shown to increase the oxidative degradation of cellulose by AA9 LPMOs (53–55).

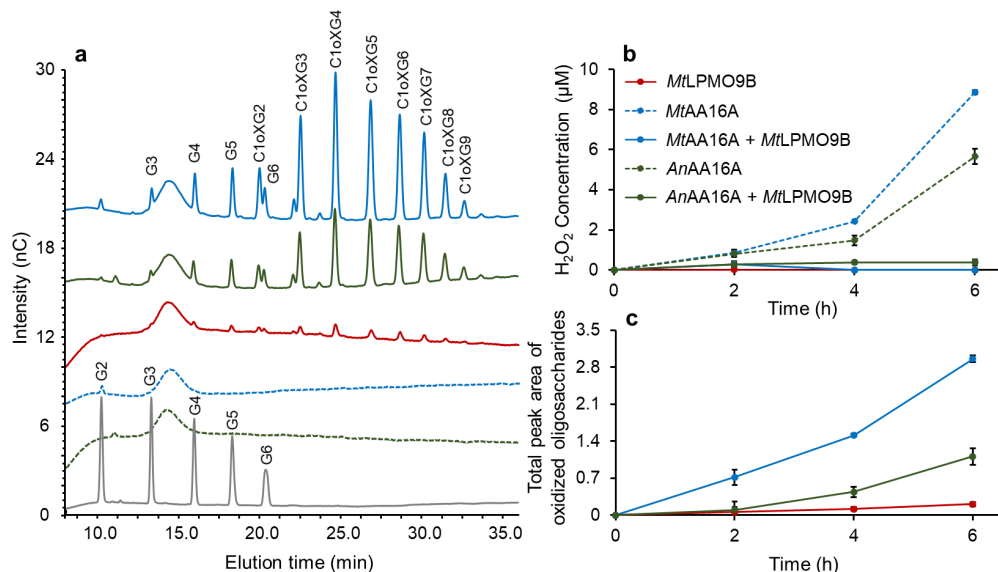


Fig. 2. (a) Oligosaccharide elution patterns determined by HPAEC, (b) H_2O_2 concentration over time determined by ferric-xylenol orange assay and (c) total peak area of oxidized oligosaccharides (GlcOx) over time determined by HPAEC are shown. RAC samples incubated with only *MtLPMO9B* (red line), only *MtAA16A* (blue dotted line), only *AnAA16A* (green dotted line), *MtAA16A* + *MtLPMO9B* (blue line) and *AnAA16A* + *MtLPMO9B* (green line) in the presence of Pyg. Standards of cello-oligosaccharides (DP2–6 (G2–G6), 1 $\mu\text{g}/\text{mL}$ each) are shown in grey. Annotation of non-oxidized (G2–6) and C1-oxidized cello-oligosaccharides (C1oxG2–9) is based on the previous study (41, 44). HPAEC chromatograms of control samples are shown in Fig. S5.

We measured the concentration of H_2O_2 in *MtLPMO9B*-RAC-Pyg digests, in absence and presence of *MtAA16A* or *AnAA16A* and/or *MtLPMO9B*. As shown in **Fig. 2b**, the amount of H_2O_2 increased in the *MtAA16A*- and *AnAA16A*-containing samples, in absence of *MtLPMO9B*. But in presence of *MtLPMO9B*, the H_2O_2 concentration remained low for both AA16-containing digests (**Fig. 2b**), while the amount of oxidized oligosaccharides substantially increased (**Fig. 2c**). These observations are in line with our hypothesis that AA16 enzymes produce H_2O_2 to boost *MtLPMO9B* activity.

3.4. H_2O_2 supply is less effective than AA16 enzyme supply in boosting catalytic action of *MtLPMO9B*

Next, it was studied whether exogenous H_2O_2 is as effective as AA16 enzyme supply in boosting the oxidative cellulose degradation by *MtLPMO9B*. Hereto, H_2O_2 was added in 6 equal aliquots at 6 successive time points during the incubation of *MtLPMO9B* with RAC. Continuous supply of H_2O_2 to increase AA9 LPMO catalytic action to oxidatively cleave cellulose has been shown effective in other studies (53–55). Furthermore, ascorbic acid (Asc) was used as reducing agent, known to be the most optimal for AA9 LPMOs including *MtLPMO9B* (16, 26, 30, 34, 44). In the presence of Asc, *MtLPMO9B* released oxidized products from RAC (6 h; **Fig. 3a**), and these products were increased

at 16 h (**Fig. 3b**). H_2O_2 supply (0, 10, 25, 50, 100 and 200 μM) boosted the *MtLPMO9B* action, seen from the increased amounts of oxidized products formed (6 h; **Fig. 3a** and Fig. S6). For H_2O_2 supply of 50 μM and higher, after 6 h of incubation, no additional oxidized products were formed (Fig. S6), which can be the result of a damaged active site of the *MtLPMO9B* (25, 36, 37). Notably, addition of AA16 enzymes resulted in the largest amounts of oxidized products formed from RAC by the *MtLPMO9B* (control reactions are shown in Fig. S7), which was approximately 2 times higher compared to the most optimal H_2O_2 concentration (50 μM , **Fig. 3**) supplied to the *MtLPMO9B*-RAC digest.

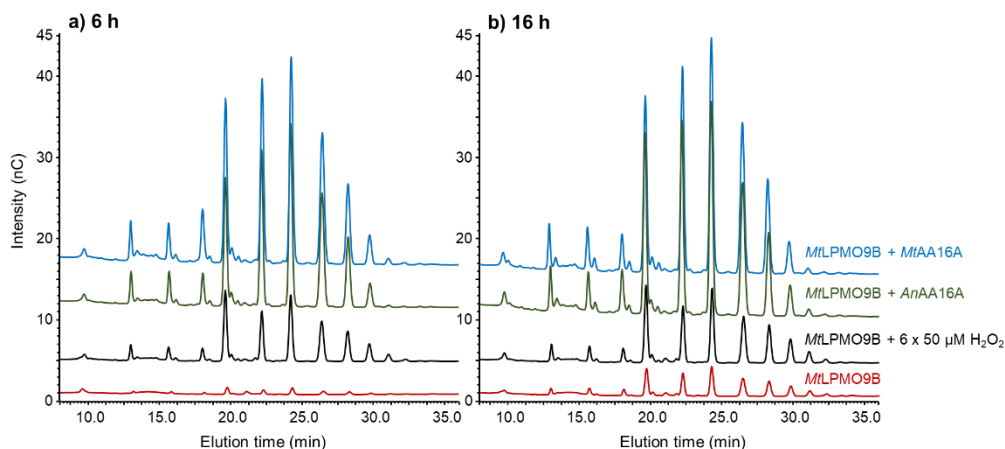


Fig. 3. HPAEC elution patterns of RAC samples incubated with only *MtLPMO9B* (red line), *MtLPMO9B* with the addition of 50 μM H_2O_2 (in total 6 times) (black line), *MtLPMO9B* + *AnAA16A* (green line) and *MtLPMO9B* + *MtAA16A* (blue line) in the presence of Asc at 6 h (A) and 16 h (B). HPAEC chromatograms of control samples are shown in Fig. S5 and S6. Compared to all concentration of H_2O_2 , highest activity was found when adding 50 μM H_2O_2 to *MtLPMO9B*-RAC digest at both 6 and 16 h. HPAEC chromatograms of *MtLPMO9B* with the addition of 0, 10, 25, 100 and 200 μM H_2O_2 (in total 6 times) in the presence of Asc are shown in Fig. S6.

3.5. AA16 enzymes boost other *MtLPMO9*s, but not AA9 LPMOs from *Neurospora crassa* (*NcLPMO9*s)

Next, we asked the question if simply the production of H_2O_2 is responsible for the increased *MtLPMO9B* activity towards cellulose in the presence of *MtAA16A*. In that case, *MtAA16A* should also boost other AA9 LPMOs. Hence, *MtAA16A* or *AnAA16A* was added to three other *MtLPMO*s and to three *NcLPMO*s, and the oxidative products formed from RAC, in the presence of Asc, were analyzed by HPAEC-PAD. As shown in **Fig. 4**, *MtAA16A* and *AnAA16A* boosted RAC degradation by *MtLPMO9E* (a), *MtLPMO9I* (b) and *MtLPMO9H* (c) (their expression, production and purification is described in Supplementary Information). However, addition of *MtAA16A* and *AnAA16A* did hardly boost the cellulose degradation by *NcLPMO9C* (d), *NcLPMO9B* (e) and *NcLPMO9F* (f) (their expression, production and purification is described in Supplementary Information). The latter observations are not in line with the above hypothesis that AA16 enzymes produce H_2O_2 to boost LPMO activity. To test if exogenously added H_2O_2

can boost *NcLPMO* activity, H_2O_2 (0, 10, 25, 50, 100 and 200 μM) was added in *NcLPMO*9M digestions under the same experimental setting and conditions as described in the previous section. The oxidative cleavage of RAC by *NcLPMO*9M was indeed boosted by H_2O_2 , with a maximum of up to 100 μM per addition (600 μM in total) (Fig. S8 and S9). Furthermore, *NcLPMO*9C has previously been shown to benefit from exogenous supply of H_2O_2 to oxidatively cleave RAC (55). It should be noted that, in that study, exogenous supply of H_2O_2 to *NcLPMO*9C was conducted under much lower Asc concentration (12 μM) (55) compared to the Asc used in this study (1 mM).

It might be that *NcLPMO*9C activity (in the presence of 1 mM Asc) cannot be boosted by AA16 anymore as it reached its maximum. Therefore, we tested all six AA9 LPMOs towards RAC with and without *MtAA16A* or *AnAA16A* in the presence of lower amount of Asc (50 μM). Similar to the results using 1 mM Asc, the boosting effect was seen in all three *MtLPMO*-AA16-RAC samples, but not convincingly in *NcLPMO*-AA16-RAC samples (Fig. S10). Apparently, merely H_2O_2 production by the AA16s cannot explain the boosting effect of *MtLPMOs*. Thus, we propose the challenging idea that *MtAA16A* and *AnAA16A* interact with *MtLPMO*9s, but not with *NcLPMO*9s, and that this interaction facilitates the transmission of H_2O_2 to the catalytic sites of the *MtLPMO*9s, instead of diffusing into the bulk solution.

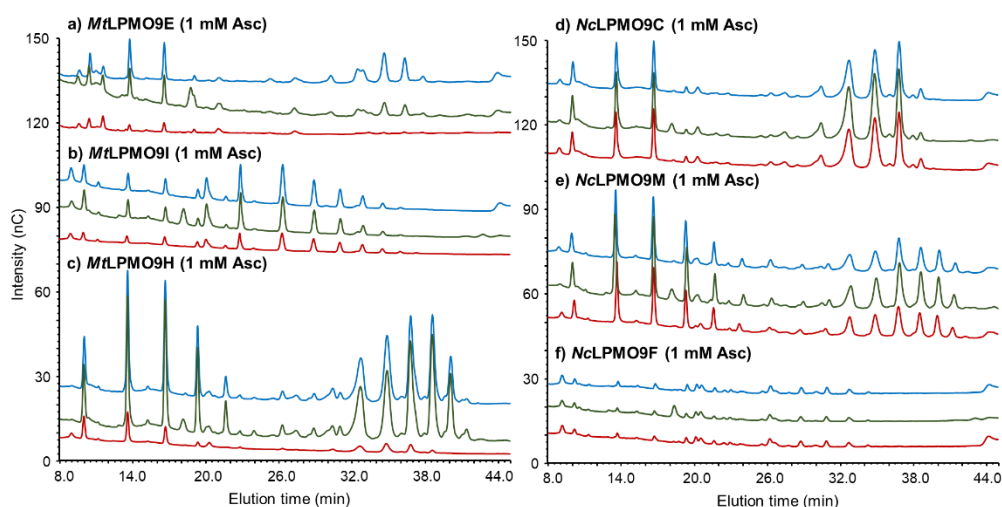


Fig. 4. HPAEC chromatograms of RAC samples incubated with various AA9 LPMOs ((a) *MtLPMO*9E; (b) *MtLPMO*9I; (c) *MtLPMO*9H; (d) *NcLPMO*9C; (e) *NcLPMO*9M and (f) *NcLPMO*9F) in the presence of 1 mM Asc after 16 h. RAC samples incubated with only LPMO, LPMO + *AnAA16A* and LPMO + *MtAA16A* are shown in red, green and blue lines, respectively.

3.6. Hypothetical AA9 LPMO-AA16 protein interaction facilitates the H_2O_2 -driven boosting effect

As mentioned above, we propose that the AA16 enzymes boosting effect on AA9 LPMO cellulose degradation is not only due to the fueling of H_2O_2 , but also facilitated by a hypothetical LPMO-AA16 protein interaction (**Fig 5a**). In the presence of O_2 and

reducing agent, AA16-Cu(II) can be reduced and, subsequently, H_2O_2 is produced. Similar reduction occurs on AA9 LPMOs, however, these LPMOs will not produce significant amounts of H_2O_2 when they bind to the carbohydrate substrate. Due to the hypothesized LPMO-AA16 protein interaction, LPMO and AA16 enzyme may stay close enough to ensure that AA16 enzyme-produced H_2O_2 is directly consumed by the LPMO. In this way, the LPMO reaction follows the H_2O_2 -dependent route as described in **Fig. 5d** and **Fig. 1, route 3**. Although O_2 is present, H_2O_2 has been shown to be the preferred co-substrate for LPMO (25, 35, 36). If LPMO-AA16 protein interaction is weak or even absent, AA16 enzyme produced H_2O_2 has a much lower chance to reach the LPMOs' active site, and will diffuse more easily into the bulk solution (**Fig. 5b**). In that case, the LPMO follows the O_2 -dependent route (**Fig. 5c** and **Fig. 1, route 1**), and no AA16 enzyme boosting effect will be observed. The AA16s studied here are supposed to interact with *Mt*LPMO9s, but not with *Nc*LPMOs, and thus they can boost the cellulose degradation by *Mt*LPMO9s, but not by *Nc*LPMOs.

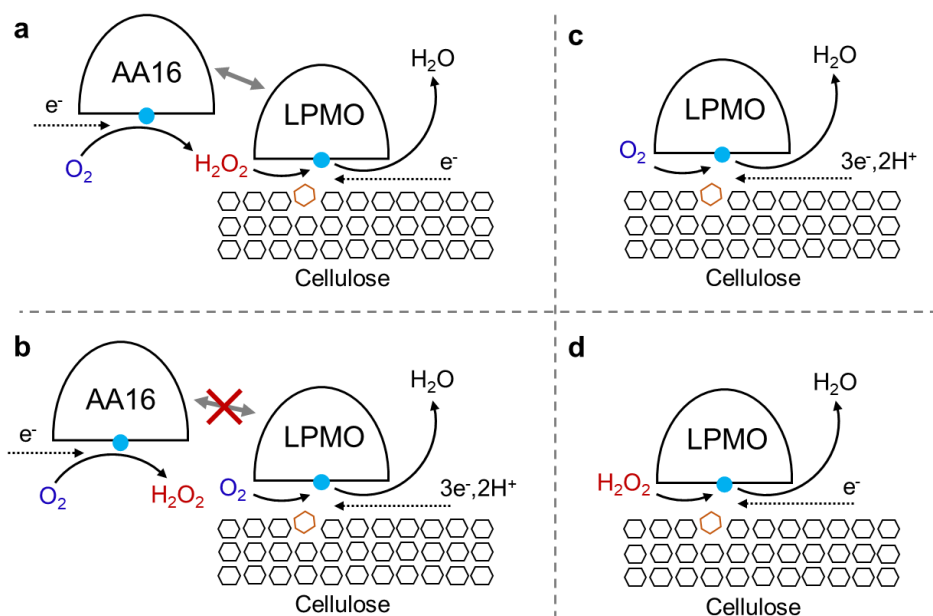


Fig. 5. Proposed mechanism for the AA16 boosting effect on the oxidative cleavage of cellulose by AA9 LPMOs. (a) When AA9 and AA16 protein interaction is present, AA16 produced H_2O_2 can easily reach AA9 LPMO to boost the oxidative cleavage. (b) When there is no (or weak) protein interaction between AA9 and AA16, AA16 produced H_2O_2 easily diffuses into bulk solution. Therefore, no (or weak) boosting effect will be observed. Simplified (c) O_2 - and (d) H_2O_2 -dependent reaction mechanisms (also described in **Fig. 1**) of AA9 LPMOs.

In the CAZy database, three putative AA16 proteins have been identified in the genome of *M. thermophila*, and zero candidates have been found in the genome of *N. crassa* (12). This may hint at a natural, evolution-driven difference, and might relate to our observation that the *M. thermophila* seems to need AA16, while *N. crassa* does not. This idea is strengthened by results from a recent study, in which Grieco et al. reported

that in the *M. thermophila* secretome one AA9 LPMO was detected together with one AA16 member and one AA3 CDH, when grown on partially delignified sugarcane bagasse (51). CDH is a well-known electron donating enzyme for AA9 LPMOs (19, 34, 56, 57), and AA16 may serve as a H₂O₂-producer in synergy with AA3 and AA9. In a very recent study, AA7 cellobiosaccharide dehydrogenase (CelDH7) has been described to transfer electrons for the AA9 LPMO catalytic reaction (58). In addition to CelDHs, the AA7 family contains oligosaccharide oxidases that oxidize cello-, chito- or xylo-oligosaccharides and produce H₂O₂ (59-61). Similar to AA3, AA7 family members have also been found to be co-expressed with LPMOs during growth on different plant biomass (62, 63). Co-expression with these classes of oxidoreductases again reflects that AA16 enzymes are expected to play an important role in fungal lignocellulosic biomass degradation.

4. Conclusions

In this work, we investigated a homologously produced MtAA16A from *M. thermophila* C1. MtAA16A did not oxidatively cleave carbohydrate substrates directly, but stimulated AA9 MtLPMOs in degrading cellulose through H₂O₂ production. A similar boosting effect was observed for AnAA16A from *A. nidulans*. The addition of MtAA16A or AnAA16A achieved a two times higher cellulose degradation yield compared to optimally supplied (exogenous) H₂O₂. Our findings provide new insights in the biological functions of copper-dependent enzymes and contribute to a further understanding of synergism within fungal enzymatic systems in degrading lignocellulose.

5. References

1. Cherubini F. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Convers Manage.* 2010;51:1412-21.
2. Sheldon RA. Green and sustainable manufacture of chemicals from biomass: state of the art. *Green Chem.* 2014;16:950-63.
3. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol.* 2015;33:747-61.
4. Straathof AJ. Transformation of biomass into commodity chemicals using enzymes or cells. *Chem Rev.* 2014;114:1871-908.
5. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, et al. Fungal cellulases. *Chem Rev.* 2015;115:1308-448.
6. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, et al. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science.* 2007;315:804-7.
7. Nanda S, Mohammad J, Reddy SN, Kozinski JA, Dalai AK. Pathways of lignocellulosic biomass conversion to renewable fuels. *Biomass Convers Biorefin.* 2014;4:157-91.
8. Kabel MA, Frommhagen M, Sun P, Schols HA. Modification of plant carbohydrates using fungal enzymes. In: Zaragoza Ó, Casadevall A, editors. *Encyclopedia of Mycology.* Oxford: Elsevier; 2021. p. 370-84.
9. Merino ST, Cherry J. Progress and challenges in enzyme development for biomass utilization. *Adv Biochem Eng Biot.* 2007;108:95-120.
10. Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod Biorefin.* 2007;1:119-34.
11. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels.* 2012;5:45.
12. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014;42:D490-5.

13. Ipsen JO, Hallas-Moller M, Brander S, Lo Leggio L, Johansen KS. Lytic polysaccharide monooxygenases and other histidine-brace copper proteins: structure, oxygen activation and biotechnological applications. *Biochem Soc Trans.* 2021;531-40.
14. Sabbadin F, Urresti S, Henrissat B, Avrova AO, Welsh LRJ, Lindley PJ, et al. Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes. *Science.* 2021;373:774.
15. Chylenski P, Bissaro B, Sorlie M, Rohr AK, Varnai A, Horn SJ, et al. Lytic polysaccharide monooxygenases in enzymatic processing of lignocellulosic biomass. *ACS Catal.* 2019;9:4970-91.
16. Frommhagen M, Westphal AH, van Berkel WJH, Kabel MA. Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases. *Front Microbio* 2018;9:1080.
17. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevret D, et al. AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels.* 2019;12:55.
18. Sun P, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng.* 2021;9:14124-33.
19. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels.* 2012;5:79.
20. Hemsworth GR, Davies GJ, Walton PH. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol.* 2013;23:660-8.
21. Tandrup T, Frandsen KEH, Johansen KS, Berrin J-G, Lo Leggio L. Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochem Soc Trans.* 2018;46:1431.
22. Quinlan RJ, Sweeney MD, Leggio LL, Otten H, Poulsen J-CN, Johansen KS, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad U S A.* 2011;108:15079-84.
23. Meier KK, Jones SM, Kaper T, Hansson H, Koetsier MJ, Karkehabadi S, et al. Oxygen activation by Cu LPMOs in recalcitrant carbohydrate polysaccharide conversion to monomer sugars. *Chem Rev.* 2017.
24. Frandsen KE, Lo Leggio L. Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes. *IUCrJ.* 2016;3:448-67.
25. Bissaro B, Rohr AK, Muller G, Chylenski P, Skaugen M, Forsberg Z, et al. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat Chem Biol.* 2017;13:1123-8.
26. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc.* 2012;134:890-2.
27. Kracher D, Forsberg Z, Bissaro B, Gangl S, Preims M, Sygmund C, et al. Polysaccharide oxidation by lytic polysaccharide monooxygenase is enhanced by engineered cellobiose dehydrogenase. *FEBS J.* 2020;287:897-908.
28. Kracher D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K, et al. Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science.* 2016;352:1098-101.
29. Bissaro B, Varnai A, Rohr AK, Eijssink VGH. Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. *Microbiol Mol Biol Rev.* 2018;82.
30. Wang BJ, Walton PH, Rovira C. Molecular mechanisms of oxygen activation and hydrogen peroxide formation in lytic polysaccharide monooxygenases. *ACS Catal.* 2019;9:4958-69.
31. Eijssink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnol Biofuels.* 2019;12:58.
32. Wang B, Wang Z, Davies GJ, Walton PH, Rovira C. Activation of O₂ and H₂O₂ by lytic polysaccharide monooxygenases. *ACS Catal.* 2020;10:12760-9.
33. Hedison TM, Breslmayr E, Shanmugam M, Karnpakdee K, Heyes DJ, Green AP, et al. Insights into the H₂O₂-driven catalytic mechanism of fungal lytic polysaccharide monooxygenases. *FEBS J.* 2020.
34. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol.* 2011;6:1399-406.
35. Hangasky JA, Iavarone AT, Marletta MA. Reactivity of O₂ versus H₂O₂ with polysaccharide monooxygenases. *Proc Natl Acad Sci U S A.* 2018;115:4915-20.

36. Kuusk S, Bissaro B, Kuusk P, Forsberg Z, Eijsink VGH, Sorlie M, et al. Kinetics of H₂O₂-driven degradation of chitin by a bacterial lytic polysaccharide monooxygenase. *J Biol Chem.* 2018;293:523-31.
37. Forsberg Z, Bissaro B, Gullesen J, Dalhus B, Vaaje-Kolstad G, Eijsink VGH. Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J Biol Chem.* 2018;293:1397-412.
38. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol. Biofuels.* 2015;8:101.
39. Punt PJ, Burlingame RP, Pynnonen CM, Olson PT, Wery J, Visser J, et al., *Chrysosporium lucknowense* protein production system patent WO/2010/107303. 2010.
40. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnol.* 2011;7:214-23.
41. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym.* 2020;234:115917.
42. Ceccherini S, Rahikainen J, Marjamaa K, Sawada D, Gronqvist S, Maloney T. Activation of softwood kraft pulp at high solids content by endoglucanase and lytic polysaccharide monooxygenase. *Ind Crops Prod.* 2021;166:113463.
43. Loose JS, Forsberg Z, Fraaije MW, Eijsink VG, Vaaje-Kolstad G. A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett.* 2014;588:3435-40.
44. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels.* 2016;9:186.
45. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sorlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci.* 2018;27:1636-50.
46. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure.* 2012;20:1051-61.
47. Garcia-Santamarina S, Probst C, Festa RA, Ding C, Smith AD, Conklin SE, et al. A lytic polysaccharide monooxygenase-like protein functions in fungal copper import and meningitis. *Nat Chem Biol.* 2020;16:337-44.
48. Labourel A, Frandsen KEH, Zhang F, Brouilly N, Grisel S, Haon M, et al. A fungal family of lytic polysaccharide monooxygenase-like copper proteins. *Nat Chem Biol.* 2020;16:345-50.
49. Breslmayr E, Hanzek M, Hanrahan A, Leitner C, Kittl R, Santek B, et al. A fast and sensitive activity assay for lytic polysaccharide monooxygenase. *Biotechnol Biofuels.* 2018;11:79.
50. Kadowaki MA, Várnai A, Jameson J-K, Leite AE, Costa-Filho AJ, Kumagai PS, et al. Functional characterization of a lytic polysaccharide monooxygenase from the thermophilic fungus *Myceliophthora thermophila*. *PLOS ONE.* 2018;13:e0202148.
51. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol.* 2020;8:1028.
52. Frommhagen M, Mutte SK, Westphal AH, Koetsier MJ, Hinz SW, Visser J, et al. Boosting LPMO-driven lignocellulose degradation by polyphenol oxidase-activated lignin building blocks. *Biotechnol Biofuels.* 2017;10:121.
53. Muller G, Chylenski P, Bissaro B, Eijsink VGH, Horn SJ. The impact of hydrogen peroxide supply on LPMO activity and overall saccharification efficiency of a commercial cellulase cocktail. *Biotechnol Biofuels.* 2018;11:209.
54. Hegnar OA, Petrovic DM, Bissaro B, Alfredsen G, Varnal A, Eijsink VGH. pH-Dependent relationship between catalytic activity and hydrogen peroxide production shown via characterization of a lytic polysaccharide monooxygenase from *Gloeophyllum trabeum*. *App Environ Microbiol.* 2019;85..
55. Petrovic DM, Varnal A, Dimarogona M, Mathiesen G, Sandgren M, Westereng B, et al. Comparison of three seemingly similar lytic polysaccharide monooxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J Biol Chem.* 2019;294:15068-81.

56. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol.* 2011;77:7007-15.
57. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci.* 2019;20:6219.
58. Momeni MH, Fredslund F, Bissaro B, Raji O, Vuong TV, Meier S, et al. Discovery of fungal oligosaccharide-oxidising flavo-enzymes with previously unknown substrates, redox-activity profiles and interplay with LPMOs. *Nat Commun.* 2021;12:1-13.
59. Ferrari AR, Rozeboom HJ, Dobruchowska JM, van Leeuwen SS, Vugts AS, Koetsier MJ, et al. Discovery of a xylooligosaccharide oxidase from *Myceliophthora thermophila* C1. *J Biol Chem.* 2016;291:23709-18.
60. Heuts DP, Janssen DB, Fraaije MW. Changing the substrate specificity of a chitoooligosaccharide oxidase from *Fusarium graminearum* by model-inspired site-directed mutagenesis. *FEBS Lett.* 2007;581:4905-9.
61. Xu F, Golightly EJ, Fuglsang CC, Schneider P, Duke KR, Lam L, et al. A novel carbohydrate: acceptor oxidoreductase from *Microdochium nivale*. *Eur J Biochem.* 2001;268:1136-42.
62. Berrin JG, Rosso MN, Abou Hachem M. Fungal secretomics to probe the biological functions of lytic polysaccharide monooxygenases. *Carbohydr Res.* 2017;448:155-60.
63. Nekiunaite L, Arntzen MO, Svensson B, Vaaje-Kolstad G, Abou Hachem M. Lytic polysaccharide monooxygenases and other oxidative enzymes are abundantly secreted by *Aspergillus nidulans* grown on different starches. *Biotechnol Biofuels.* 2016;9:187.
64. Van Soest Pv, Robertson J, Lewis B. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J Dairy Sci.* 1991;74:3583-97.

6. Supplementary information

6.1. Supplementary materials and methods

6.1.1. Expression, production and purification of *AnAA16A* and *AA9 LPMOs*

The *AnAA16A* gene (AN0778.2) was amplified directly from the *A. nidulans* genome. The oligonucleotides *AnAA16fw* (5'-ACAAC TAATTATTCGAAACGATGAAGCAGCTACCACCG-3') and *AnAA16rv* (5'-CCCTGAAAATAAAGATTCTCGCCGTTACCACTTCCACCAA-3') were used to remove the C-terminal extension region (residues 199-306) and maintain the native signal peptide. The removal of C-terminal extension region prior to protein production has been reported by Filiatrault-Chastel et al. (1). The *AnAA16* gene was cloned into a modified pPICZα vector as previously described (2). This construction allows the expression of a recombinant *AnAA16A* containing a cleavable C-terminal polyhistidine-tag. The *Pichia pastoris* X-33 (Invitrogen) was transformed by electroporation using a PmeI-linearized plasmid (pPICZT::*AnAA16A*) and selected on Yeast Extract–Peptone–Dextrose–Sorbitol (YPDS)-zeocin plates. The recombinant colonies were randomly picked, grown in BMMY medium and the gene expression confirmed by SDS-PAGE analysis of the supernatant content. The transformant showing the highest expression profile was grown in 40 mL of YPD medium overnight and inoculated in four Erlenmeyer flasks containing 0.5 L buffered glycerol-complex medium (BMGY) medium at 30 °C and 250 rpm until an OD600 of 2. The yeast cells from each flask were harvested and transferred to 0.1 L buffered methanol-complex medium (BMMY) and incubated at 30 °C and 250 rpm for 72 hours. Two percent absolute methanol was added every 24 hours to maintain recombinant protein production. The culture supernatant was filtered and the pH adjusted to 8.0 using the Tris-HCl buffer. The entire volume was loaded onto a 5-mL HiTrap Chelating HP column (GE Healthcare, Uppsala, Sweden) connected to an AKTA Start system (GE healthcare) equilibrated with 50 mM Tris-HCl pH 8.0 and 0.3 M NaCl (Buffer-A). *AnAA16A* was eluted using a linear gradient from 0 to 100 % of 1 M imidazole within 10 CVs. The fractions containing the purified enzyme were pooled and concentrated using Amicon Ultra 15 mL centrifugal filters (Merck Millipore, USA) with a cutoff of 10 kDa cutoff. The C-terminal His-tag cleavage and removal with Tobacco etch virus (TEV) protease was performed according to Kadowaki et al. (2). The non-tagged *AnAA16A* was then saturated with copper by incubating the protein solution with a 3-fold molar excess of Cu(II)SO₄ for 10 min at room temperature followed by a size exclusion chromatography on a pre-equilibrated HiLoad 16/60 Sephadex75 size exclusion column (GE Healthcare) in 50 mM Tris/HCl buffer pH 8.0 containing 150 mM NaCl. Protein purity was analyzed by SDS-PAGE using Coomassie Brilliant Blue G-250 staining (Sigma, Deisenhofen, Germany) and the concentration determined using the Bradford method using bovine serum albumin as standard.

The purification of *MtLPMO9E* (MYCTH_79765, UniProt ID: G2Q7A5) and *MtLPMO9I* (MTCTH_2299721, UniProt ID: G2Q774) has been described previously (3). The

production and purification of *MtLPMO9H* (MYCTH_46583, UniProt ID: G2Q9T3) have been described elsewhere (64).

The expression, production and purification of *NcLPMO9C* (NCU02916, UniProt ID: Q7SHI8), *NcLPMO9M* (NCU07898, UniProt ID: Q7SA19) and *NcLPMO9F* (NCU03328, UniProt ID: Q873G1) have been described previously (5, 6).

6.1.2. Deglycosylation of *MtAA16A* by PNGase F

Part of *MtAA16A* was incubated with PNGase F (N-Glycosidase F, Madison, Wisconsin, United States) to remove N-glycosylation. The deglycosylation procedure was performed using non-denaturing conditions based on the protocol provided by the manufacture. 4 μ L PNGase F was added to 36 μ L 50 mM ammonium bicarbonate (pH 7.8) containing 40 μ g of *MtAA16A*. The sample was incubated at 37 °C for 16 h under shaking of 600 rpm (Eppendorf ThermoMixer®C, Eppendorf, Hamburg, Germany).

6.1.3. Incubation of *MtAA16A* with various oligo and polysaccharide substrates

Carbohydrates (2 mg/mL each, types and sources were listed in Table S1) were dissolved in 50 mM ammonium acetate buffer (pH 5.0) with the addition of reducing agent (Table S1, 1 mM final concentration). Subsequently, *MtAA16A* was added to a concentration of 20 or 40 mg/g substrate (Table S1). Control reactions were performed without the addition of reducing agent. The incubation of 500 μ L was performed at 30 or 50 °C (Table S1) for 24 h in an Eppendorf ThermoMixer® C at 900 rpm (in a vertical orientation). In some reactions, 100 μ M H₂O₂ was added in the beginning of incubation. The incubations were stopped by immediately separating supernatants and pellets through centrifugation 22,000 $\times g$, 15 min at 4 °C in a table centrifuge. The supernatants were used for further analysis or stored at -20 °C.

6.1.4. Determination of *MtAA16A* and *MtLPMO9s* activity by 2,6-DMP assay

A spectrophotometric assay using 2,6-dimethoxyphenol as substrate (2,6-DMP assay) was performed according to Breslmayr et al. (7). The assay was carried out in a 96-well plate with total reaction volumes of 250 μ L. Each reaction contained 100 mM potassium phosphate buffer pH 6.0, 10 mM 2,6-DMP and 100 μ M H₂O₂ (7). After pre-incubation at 30 °C, *MtLPMO9B*, *MtLPMO9E*, *MtLPMO9H*, *MtLPMO9I* or *MtAA16A* were added to give a final concentration of 0.405 mg/mL, 0.49 mg/mL, 0.401 mg/mL, 0.448 mg/mL and 0.73 mg/mL, respectively. The control reaction was performed without enzymes. After the addition of enzyme, the 96-well plate was placed in a spectrophotometer (Infinite® F500, Tecan, Männedorf, Switzerland) and the absorbance was recorded at 480 nm for 600 s at 30 °C. The molar absorption coefficient of reaction product coerulignone (53,200 M⁻¹ cm⁻¹) (and the path length of 0.7 cm) was used to calculate the DMP assay activity.

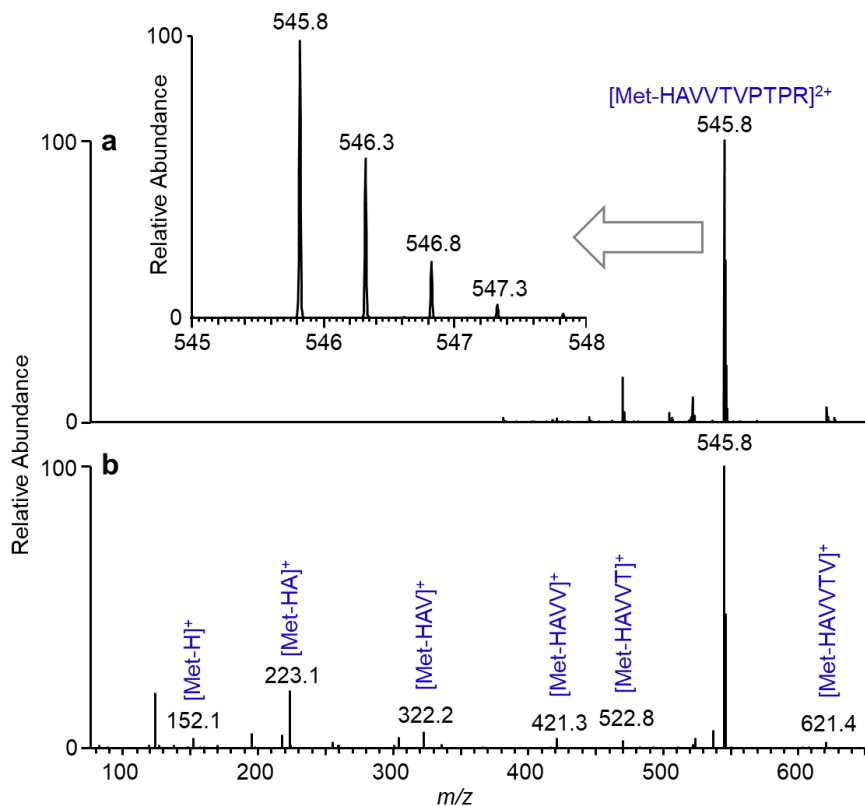


Fig. S3. LC-MS analysis of an N-terminal peptide from *MtAA16A* using positive mode. (a) Full MS showed that the N-terminal peptide was double charged. (b) MS/MS² fragmentation patterns confirmed that N-terminal peptide from *MtAA16A* was methylated.

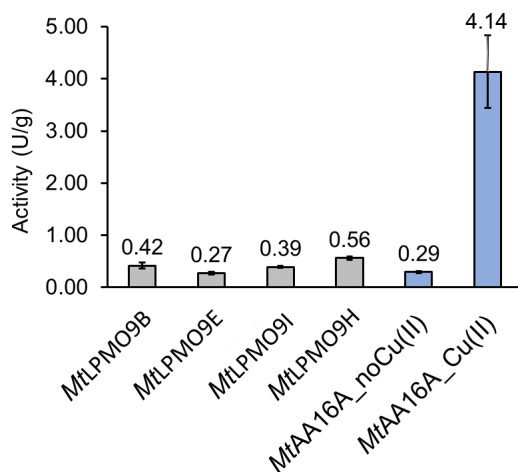


Fig. S4. Activity of *MtLPMO9*s and *MtAA16A* determined by using 2,6-DMP assay. For *MtAA16A*, an *MtAA16A* fraction directly after purification (*MtAA16A_noCu* (II)) and a copper-saturated fraction (*MtAA16A_Cu* (II)) were used.

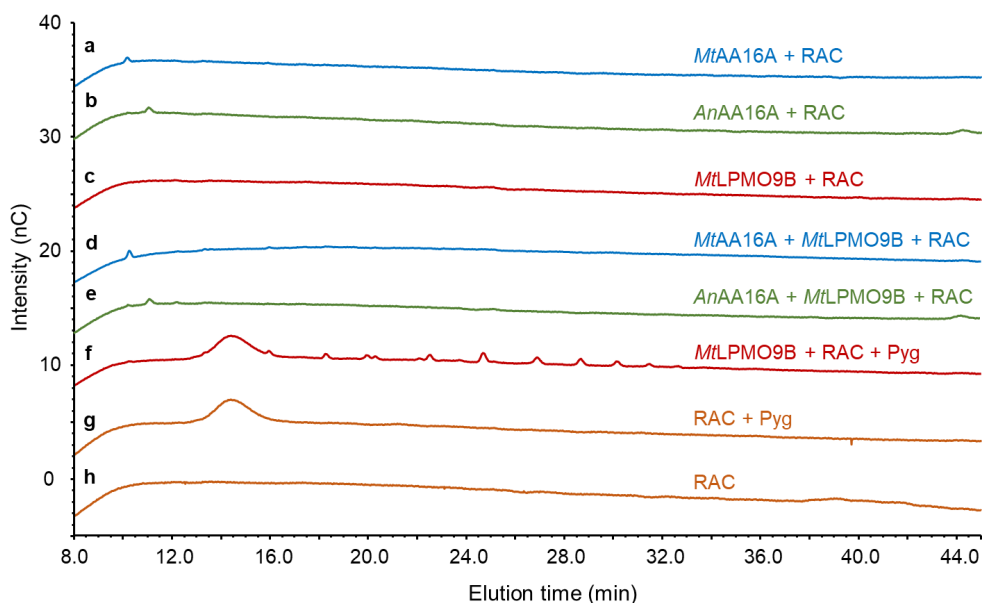


Fig. S5. HPAEC chromatograms of control reactions (in the presence of Pyg). The incubation of *MtAA16A* (a), *AnAA16A* (b), *MtLPMO9B* (c), *MtAA16A* + *MtLPMO9B* (d) and *AnAA16A* + *MtLPMO9B* (e) with RAC in the absence of Pyg did not show the generation of (detectable) cello-oligosaccharides, indicating the absence of hydrolytic side-activity. (f) The incubation of *MtLPMO9B* with RAC in the presence of Pyg showed the formation of both non-oxidized and oxidized cello-oligosaccharides. The incubation of only RAC in the presence of Pyg (g) did not cause the auto-degradation of RAC (h).

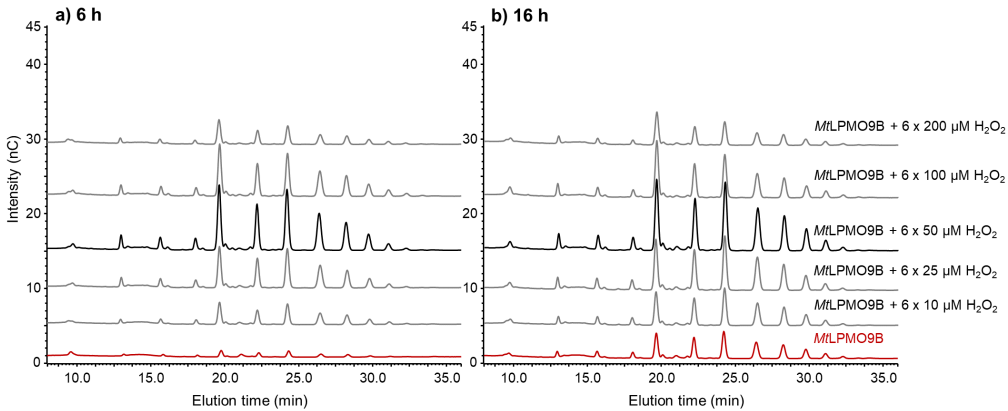


Fig. S6. HPAEC chromatograms of *MtLPMO9B*-RAC digests in the presence of Asc and addition of different concentrations of exogenous H_2O_2 (0, 10, 25, 50, 100 and 200 μM per time, 6 times in total) at 6 h (a) and 16 h (b).

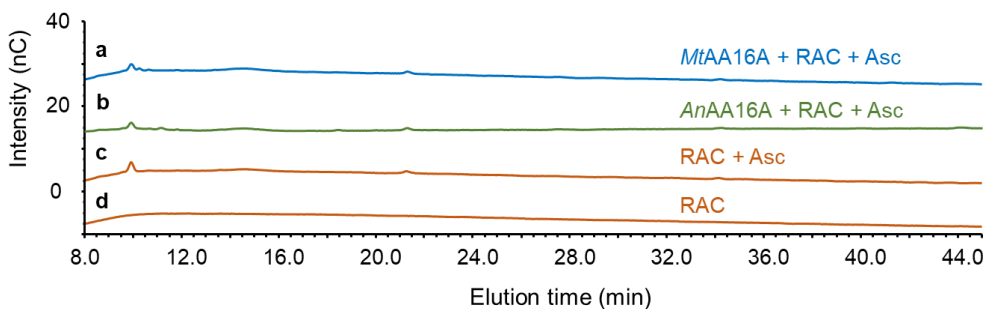


Fig. S7. HPAEC chromatograms of control reactions in the presence of Asc. The incubation of *MtAA16A* (a) and *AnAA16* (b) with RAC in the presence of Asc did not show the generation of (detectable) cello-oligosaccharides, indicating that these AA16s are not active towards RAC. The incubation of only RAC in the presence of Asc (c) did not cause the auto-degradation of RAC (d).

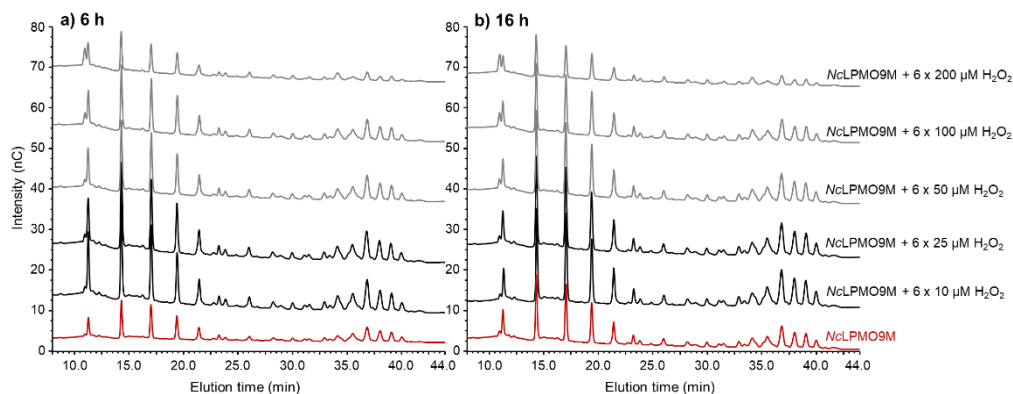


Fig. S8. HPAEC chromatograms of NcLPMO9M-RAC digests in the presence of Asc and addition of different concentration of exogenous H_2O_2 (0, 10, 25, 50, 100 and 200 μM per time, 6 times in total) at 6 h (a) and 16 h (b).

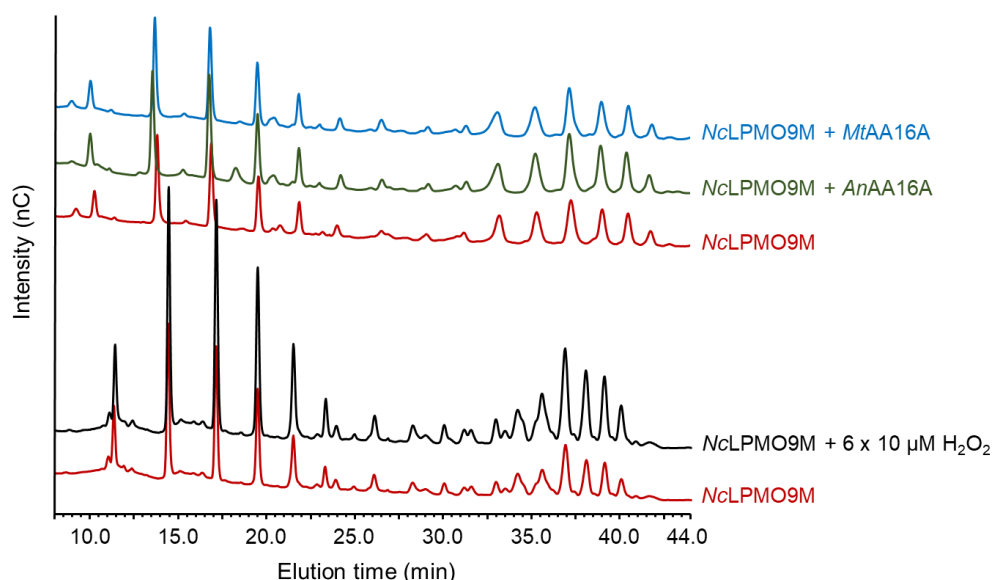


Fig. S9. HPAEC chromatograms of NcLPMO9M-RAC digests in the presence of Asc and addition of 10 μM H_2O_2 per time (6 times in total) or MtAA16A or AnAA16A at 16 h. Top three chromatograms are also shown in Fig. 4 and the bottom two chromatograms are also shown in Fig. S8. Due to the analysis was performed at different dates and systems, elution time of individual peaks is slightly different.

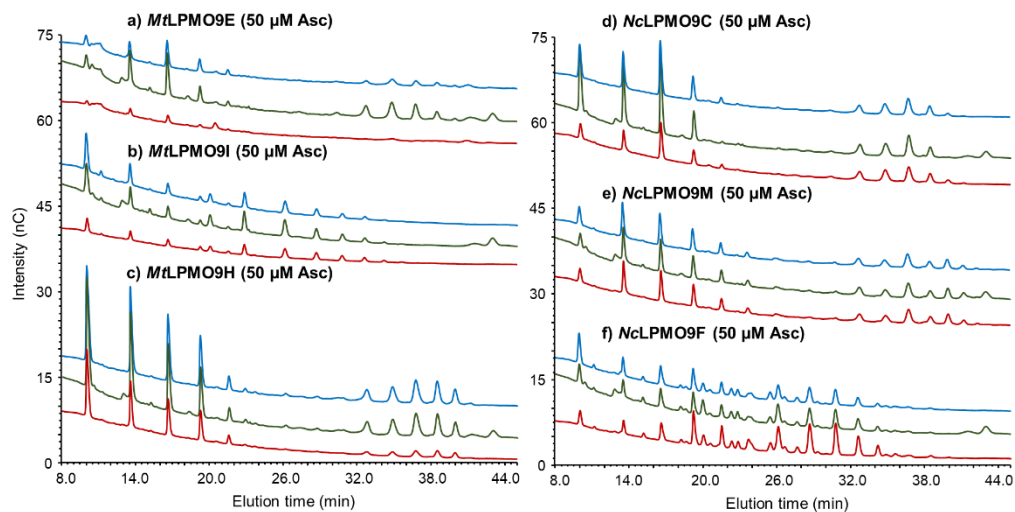


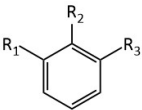
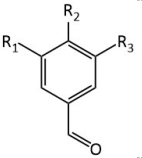
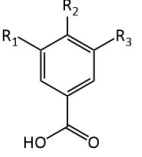
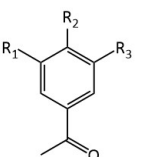
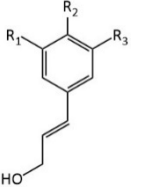
Fig. S10. HPAEC chromatograms of RAC samples incubated with various AA9 LPMOs ((a) *MtLPMO9E*; (b) *MtLPMO9I*; (c) *MtLPMO9H*; (d) *NcLPMO9C*; (e) *NcLPMO9M* and (f) *NcLPMO9F*) in the presence of 50 μM Asc after 16 h. RAC samples incubated with only LPMO, LPMO + AnAA16A and LPMO + *MtAA16A* are shown in red, green and blue lines, respectively.

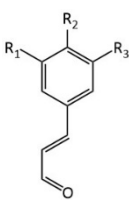
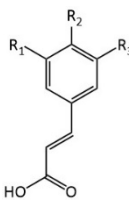
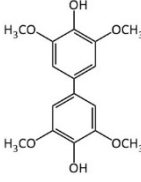
Table S1. Carbohydrates tested for screening MtAA16A activity under different conditions.

Substrate	Substrate concentration (mg/mL)	MtAA16A concentration ($\mu\text{g}/\text{mg}$ substrate)	Reducing agent	Temperature ($^{\circ}\text{C}$)	H ₂ O ₂ addition
RAC	2	10	Asc	50	-
	2	20	Asc	30	-
	2	20	Asc	30	50 μM
	2	20	Syr	30	-
	2	20	Syr	30	50 μM
	2	20	Gua	30	-
	2	20	Gua	30	50 μM
Avicel PH-101	2	20	Asc	30	-
	2	20	Asc	30	50 μM
Planetary ball milled	2	20	Asc	30	-
Avicel PH-101	2	20	Asc	30	50 μM
β -chitin from shrimp shell	2	10	Asc	50	-
	2	20	Asc	30	50 μM
β -chitin nano fiber	2	10	Asc	50	-
	2	20	Asc	30	50 μM
Regenerated amorphous β -chitin	5	30	Asc	30	-
	5	30	Asc	30	50 μM
Planetary ball milled β -chitin	5	30	Asc	30	-
	5	30	Asc	30	50 μM
β -chitosan	5	30	Asc	30	-
	5	30	Asc	30	50 μM
Regenerated amorphous β -chitosan	5	30	Asc	30	-
	5	30	Asc	30	50 μM
Planetary ball milled β -chitosan	5	30	Asc	30	-
	5	30	Asc	30	50 μM
Xylan from birch wood	2	10	Asc	50	-
	2	20	Asc	30	-
	2	20	Asc	30	50 μM
	2	20	Syr	30	-
	2	20	Syr	30	50 μM
	2	20	Gua	30	-
	2	20	Gua	30	50 μM
Xyloglucan from tamarind seed	2	20	Asc	30	-
	2	20	Asc	30	50 μM
	2	20	Syr	30	-
	2	20	Syr	30	50 μM
	2	20	Gua	30	-
	2	20	Gua	30	50 μM
Mannan from cassia	2	30	Asc	30	-
	2	30	Asc	30	50 μM
Galactan from potato	2	30	Asc	30	-
	2	30	Asc	30	50 μM
	2	30	Asc	30	-

Substrate	Substrate concentration (mg/mL)	MtAA16A concentration (µg/mg substrate)	Reducing agent	Temperature (°C)	H ₂ O ₂ addition
Laminarin from <i>Laminaria digitata</i>	2	30	Asc	30	50 µM
Glucomannan from konjac	2	30	Asc	30	-
	2	30	Asc	30	50 µM
Arabinan from sugar beet	2	30	Asc	30	-
	2	30	Asc	30	50 µM
Pectin (Degree of methylation (DM) 30) from lemon	2	10	Asc	50	-
	2	20	Asc	30	-
	2	20	Asc	30	50 µM
Pectin (DM96) from lemon	2	10	Asc	50	-
	2	20	Asc	30	-
	2	20	Asc	30	50 µM
Peptidoglycan from <i>Bacillus subtilis</i>	2	40	Asc	30	-
	2	40	Asc	30	50 µM
	2	40	Pyg	30	-
	2	40	Pyg	30	50 µM
Peptidoglycan from <i>Saccharomyces cerevisiae</i>	2	40	Asc	30	-
	2	40	Asc	30	50 µM
	2	40	Pyg	30	-
	2	40	Pyg	30	50 µM
Soy soluble polysaccharides	5	20	Asc	30	-
	5	20	Asc	30	50 µM
Fungal cell wall polysaccharides from <i>Pleurotus ostreatus</i>	5	20	Asc	30	-
	5	20	Asc	30	50 µM
RAC + Xylan	2 each	10	Asc	50	-
	2 each	20	Asc	30	-
	2 each	20	Asc	30	50 µM
	2 each	20	Syr	30	-
	2 each	20	Syr	30	50 µM
	2 each	20	Gua	30	-
	2 each	20	Gua	30	50 µM
RAC + Xyloglucan	2 each	10	Asc	50	-
	2 each	20	Asc	30	-
	2 each	20	Asc	30	50 µM
	2 each	20	Syr	30	-
	2 each	20	Syr	30	50 µM
	2 each	20	Gua	30	-
	2 each	20	Gua	30	50 µM
RAC + Pectin (DM30)	2 each	10	Asc	50	-
	2 each	20	Asc	30	-
	2 each	20	Asc	30	50 µM
RAC + Pectin (DM96)	2 each	10	Asc	50	-
	2 each	20	Asc	30	-

Table S2. Molecular structure, common name and International Union of Pure and Applied Chemistry (IUPAC) name of the 27 monomeric aromatic compounds and the dimeric aromatic compound hydrocoerulignone screened in this study for activity with *MtAA16A*. Compounds that can be oxidized by *MtAA16A* are indicated in green.

	Common name	IUPAC name	R ₁	R ₂	R ₃	Oxidized by <i>MtAA16A</i>
	I. Benzene	I. Benzene				
	II. Phenol	II. Hydroxybenzene				
	III. Anisole	III. Methoxybenzene	H	H	H	No
	IV. Catechol	IV. Benzene-1,2-diol	H	OH	H	No
	V. Veratrole	V. 1,2-Dimethoxybenzene	H	OCH ₃	H	No
	VI. Guaiacol	VI. 2-Methoxyphenol	H	OH	OH	Yes
	VII. Pyrogallol	VII. Benzene-1,2,3-triol	H	OH	OCH ₃	Yes
	VIII. Pyrogallol trimethyl ether	VIII. 1,2,3-Trimethoxybenzene	OH	OH	OH	Yes
	IX. Syringol	IX. 2,6-Dimethoxyphenol	OCH ₃	OCH ₃	OCH ₃	No
			OCH ₃	OH	OCH ₃	Yes
	I. 4-Formylphenol	I. 4-Hydroxybenzaldehyde	H	OH	H	No
	II. Vanillin	II. 4-Hydroxy-3-methoxybenzaldehyde	H	OH	OCH ₃	No
	III. Syringaldehyde	III. 4-Hydroxy-3,5-dimethoxybenzaldehyde	OCH ₃	OH	OCH ₃	No
	I. 4-Hydroxybenzoate	I. 4-Hydroxybenzoic acid	H	OH	H	No
	II. Vanillic acid	II. 4-Hydroxy-3-methoxybenzoic acid	H	OH	OCH ₃	No
	III. Syringic acid	III. 4-Hydroxy-3,5-dimethoxybenzoic acid	OCH ₃	OH	OCH ₃	Yes
	I. Piceol	I. 1-(4-Hydroxyphenyl)ethanone				
	II. Acetovanillone	II. 1-(4-Hydroxy-3-methoxyphenyl)ethanone	H	OH	H	No
	III. Acetosyringone	III. 1-(4-Hydroxy-3,5-dimethoxyphenyl)ethanone	H	OH	OCH ₃	No
			OCH ₃	OH	OCH ₃	No
	I. <i>p</i> -Coumaryl alcohol	I. 4-[(<i>E</i>)-3-Hydroxyprop-1-enyl]phenol				
	II. Coniferyl alcohol	II. 4-[(<i>E</i>)-3-Hydroxyprop-1-enyl]-2-methoxyphenol	H	OH	H	No
	III. Sinapyl alcohol	III. 4-[(<i>E</i>)-3-Hydroxyprop-1-enyl]-2,6-dimethoxyphenol	H	OH	OCH ₃	No
			OCH ₃	OH	OCH ₃	Yes

	I. <i>p</i> -Coumaryl aldehyde	I. (<i>E</i>)-3-(4-Hydroxyphenyl)prop-2-enal				
	II. Coniferyl aldehyde	II. (<i>E</i>)-3-(4-Hydroxy-3-methoxyphenyl)prop-2-enal	H H OCH ₃	OH OH OH	H OCH ₃ OCH ₃	No No Yes
	III. Sinapylaldehyde	III. (<i>E</i>)-3-(4-Hydroxy-3,5-dimethoxyphenyl)prop-2-enal				
	I. <i>p</i> -Coumaric acid	I. (<i>E</i>)-3-(4-Hydroxyphenyl)prop-2-enoic acid				
	II. Ferulic acid	II. (<i>E</i>)-3-(4-Hydroxy-3-methoxyphenyl)prop-2-enoic acid	H H OCH ₃	OH OH OH	H OCH ₃ OCH ₃	No Yes Yes
	III. Sinapic acid	III. (<i>E</i>)-3-(4-Hydroxy-3,5-dimethoxyphenyl)prop-2-enoic acid				
	Hydro-coerulignone	4-(4-Hydroxy-3,5-dimethoxyphenyl)-2,6-dimethoxyphenol				Yes

6.3. Supplementary references

1. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevret D, et al. AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels*. 2019;12:55.
2. Kadowaki MAS, Magri S, de Godoy MO, Monclaro AV, Zarattini M, Cannella D. A fast and easy strategy for lytic polysaccharide monooxygenase-cleavable His6-Tag cloning, expression, and purification. *Enzyme Microb Technol*. 2021;143:109704.
3. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
4. Sun P, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng*. 2021;9:14124-33.
5. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
6. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
7. Breslmayr E, Hanzek M, Hanrahan A, Leitner C, Kittl R, Santek B, et al. A fast and sensitive activity assay for lytic polysaccharide monooxygenase. *Biotechnol Biofuels*. 2018;11:79.

Chapter 9

General discussion

1. General discussion outline

This thesis aimed to provide further insights into AA9 LPMOs from the fungus *Myceliophthora thermophila* C1 (*MtLPMOs*). In particular, regioselectivity, substrate cleavage profiles and specificity of *MtLPMOs* have been studied and further defined. The results obtained by using these *MtLPMOs* were combined with those obtained by using AA9 LPMOs from the fungus *Neurospora crassa* (*NcLPMOs*). This comparison allowed us to expand our knowledge of LPMOs within the AA9 family. In addition, we investigated AA16 oxidoreductases and highlighted their importance for lignocellulose degradation when added to AA9 LPMOs. In this chapter, the outcomes from the previous chapters are discussed in context of literature findings, and are divided into five topics: i) diversity of AA9 LPMOs, ii) improved chromatographic approach for analysis of oxidized oligosaccharides, iii) production and consumption of H₂O₂ by AA16 oxidoreductases and AA9 LPMOs, iv) lignin-derived electron donors for H₂O₂ mediated AA16/AA9 activity, and v) role of LPMOs in lignocellulosic biomass degradation and valorization.

2. Diversity of AA9 LPMOs

2.1. AA9 LPMOs differ in active site segment configuration and substrate specificity

In this thesis, we studied AA9 LPMOs from two fungi: *M. thermophila* C1 (*MtLPMOs*) and *N. crassa* (*NcLPMOs*). In the genome of *M. thermophila* C1, twenty-two putative LPMO genes are annotated in the AA9 family (**Chapter 1**). Until now, eleven AA9 *MtLPMOs* have been (partially) characterized, of which five (*MtLPMO9B* (*Mt9B⁺* and *Mt9B⁻*), *MtLPMO9E*, *MtLPMO9F*, *MtLPMO9H* and *MtLPMO9I*) were investigated in this thesis (**Table 1**). In addition, ten AA9 LPMOs from the fungus *N. crassa* (*NcLPMOs*) have been characterized, of which three (*NcLPMO9C* (and *NcLPMO9C^{ΔCBM}*), *NcLPMO9M* and *NcLPMO9F*) have been studied in this thesis (**Table 2**).

Table 1. Characterized LPMOs from *M. thermophila* (September 2021)¹. Active site segment numbers are explained in **Chapter 5, Fig. 1**.

Enzyme name	Active site segment	Regio-selectivity	Substrate specificity	Xyloglucan cleavage profile
<i>MtLPMO9D</i> ²	+Seg1-Seg2	C1	Cellulose (RAC and PASC) ¹⁷	n.d.
<i>MtLPMO9</i> (<i>MtLPMO9H</i>) ³	+Seg1-Seg2	C1, C4, C4/C6 ¹⁶	Cellulose (RAC, AVI, BC and CMC), xyloglucan	Substitution-tolerant
<i>MtLPMO9C</i> ⁴	-Seg1+Seg2	C4	Cellulose (RAC), xyloglucan ¹⁸ , mixed β-(1→3, 1→4)-linked glucan ¹⁸	n.d.
<i>MtLPMO9E</i> ⁵	-Seg1+Seg2	C4, C4/C6 ¹⁶	Cellulose (RAC), xyloglucan, cello-oligosaccharides (DP≥5), glucomannan ¹⁸ ,	Substitution-intolerant

			mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan ¹⁸	
<i>MtLPMO9F</i> ⁶	-Seg1+Seg2	C4, C4/C6 ¹⁶	Cellulose (RAC, AVI, BC and CMC), xyloglucan, cello-oligosaccharides (DP \geq 5), glucomannan ¹⁸ , mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan ¹⁷	Substitution-intolerant
<i>MtLPMO9G</i> ⁷	-Seg1+Seg2 ¹⁵	C4	Cellulose (RAC) ¹⁷	n.d.
<i>MtLPMO9B</i> (<i>Mt9B</i>) ⁸	-Seg1+Seg2+Seg3	C1	Cellulose (RAC, AVI, BC and CMC)	Inactive
<i>Mt9B</i> ⁹	-Seg1+Seg2+Seg3	C1	Cellulose (RAC, AVI, BC and CMC)	Inactive
<i>MtLPMO9I</i> ¹⁰	-Seg1+Seg2+Seg3	C1	Cellulose (RAC, AVI and BC)	Inactive
<i>MtLPMO9A</i> ¹¹	-Seg1-Seg2 ¹⁴	C1, C4	Cellulose (RAC), xylan associated to cellulose, xyloglucan ¹⁸ , mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan ¹⁸	Active ¹⁸
<i>MtLPMO9L</i> ¹²	-Seg1-Seg2 ¹⁴	C1	Cellulose (PASC and AVI)	Inactive
MYCTH_112089 ¹³	-Seg1-Seg2 ¹⁴	C1	Cellulose (PASC) ¹⁷	n.d.

¹Abbreviations: RAC, regenerated amorphous cellulose; AVI, Avicel; PASC, phosphoric acid swollen cellulose; BC, bacterial cellulose; CMC, carboxymethyl cellulose; n.d., not determined

²MYCTH_92668, UniProt ID: G2QAB5, production host *M. thermophila* C1, no CBM (1);

³MYCTH_46583, UniProt ID: G2Q9T3, with CBM1, production host *P. pastoris* (2, 3); production host *M. thermophila* C1 (4-7); ⁴MYCTH_100518, UniProt ID: G2QA92, production host *M. thermophila* C1, no CBM (8); production host *P. pastoris* (9); ⁵MYCTH_79765, UniProt ID: G2Q7A5, production host *M. thermophila* C1, no CBM (5, 10, 11); ⁶MYCTH_111088, UniProt ID: G2Q9F7, production host *M. thermophila* C1, with CBM1 (6, 10); ⁷MYCTH_55803, UniProt ID: n.d., production host *M. thermophila* C1, no CBM (10); ⁸MYCTH_80312, UniProt ID: G2QCJ3, production host *M. thermophila* C1, with CBM1 (4, 12); ⁹MYCTH_80312, UniProt ID: G2QCJ3, production host *M. thermophila* C1, no CBM (4); ¹⁰MYCTH_2299721, UniProt ID: G2Q774, production host *M. thermophila* C1, no CBM (11); ¹⁰MYCTH_85556, UniProt ID: G2QNT0, production host *M. thermophila* C1, no CBM (12, 13); ¹²MYCTH_103537, UniProt ID: n.d., production host *P. pastoris*, no CBM (14); ¹³MYCTH_112089, UniProt ID: G2QI82 (9)

¹⁴Based on the reported short L3 loop and L2 loop

¹⁵Unpublished

¹⁶Defined in this thesis (see **Chapter 4**)

¹⁷Only tested on cellulose

¹⁸Minor activity

Table 2. Characterized LPMOs from *N. crassa* (September 2021)¹. Active site segment numbers are explained in **Chapter 5, Fig. 1**.

Enzyme name	Active site segment	Regio-selectivity	Substrate specificity	Xyloglucan cleavage profile
NcLPMO9B ²	+Seg1-Seg2 ¹²	C1, C4	Cellulose (PASC) ¹⁴	n.d.
NcLPMO9M ³	+Seg1-Seg2	C1, C4, C4/C6 ¹³	Cellulose (RAC, AVI, BC and CMC), xyloglucan	Substitution-tolerant
NcLPMO9A ⁴	-Seg1+Seg2	C4	Cellulose (RAC and PASC), xyloglucan, glucomannan	Substitution-intolerant
NcLPMO9C ⁵	-Seg1+Seg2	C4, C4/C6 ¹³	Cellulose (RAC, AVI and CMC), xyloglucan, cello-oligosaccharides (DP≥4), glucomannan ¹⁵ , mixed β-(1→3, 1→4)-linked glucan ¹⁵	Substitution-intolerant
NcLPMO9C ΔCBM ⁶	-Seg1+Seg2	C4, C4/C6 ¹³	Cellulose (RAC, AVI and CMC), xyloglucan ¹⁶	Substitution-intolerant
NcLPMO9D ⁷	-Seg1+Seg2	C4	Cellulose (RAC and PASC), xyloglucan, glucomannan	Substitution-intolerant
NcLPMO9E ⁸	-Seg1-Seg2 ¹²	C1	Cellulose (RAC)	n.d.
NcLPMO9F ⁹	-Seg1-Seg2	C1	Cellulose (PASC, AVI and CMC)	Inactive
NcLPMO9G ¹⁰	-Seg1-Seg2 ¹²	C1	Cellulose (PASC) ¹⁴	n.d.
NcLPMO9J ¹¹	-Seg1-Seg2 ¹²	C1	Cellulose (PASC) ¹⁴	n.d.

¹Abbreviations: RAC, regenerated amorphous cellulose; AVI, Avicel; PASC, phosphoric acid swollen cellulose; BC, bacterial cellulose; CMC, carboxymethyl cellulose; n.d., not determined. All NcLPMOs were produced in *P. pastoris*. NcLPMO9D, NcLPMO9E and NcLPMO9M were also produced in *N. crassa* in addition to *P. pastoris*

²NCU07760, UniProt ID: Q7S111, with CBM1, (9); ³NCU07898, UniProt ID: Q7SA19, no CBM (4, 5, 7, 9, 15, 16); ⁴NCU02240, UniProt ID: Q7S439, with CBM1 (9, 17, 18); ⁵NCU02916, UniProt ID: Q7SHI8, with CBM1 (5, 7, 9, 18-26); ⁶NCU02916, UniProt ID: Q7SHI8, no CBM (26); ⁷NCU01050, UniProt ID: Q1K8B6, no CBM (9, 15, 16); ⁸NCU08760, UniProt ID: Q7RWN7, with CBM1 (9, 15, 16, 18, 21); ⁹NCU03328, UniProt ID: Q1K4Q1, no CBM (7, 9, 21, 24-26); ¹⁰NCU00836, UniProt ID: Q7SCJ5, with CBM1 (9); ¹¹NCU01867, UniProt ID: Q7SHD9, with CBM1 (21, 24);

¹²Based on the sequence alignment and determined L2, L3 and LS loops

¹³Defined in this thesis (see **Chapter 4**)

¹⁴Only tested on cellulose

¹⁵Minor activity

¹⁶Only tested on cellulose and xyloglucan

All reported AA9 MtLPMOs and NcLPMOs show oxidative cleavage of cellulose (**Table 1 and 2**). So far most studies only provided data on whether cellulose was (or was not) a substrate for oxidative cleavage by LPMOs, regardless the type of cellulose used. In this thesis, we demonstrated that AA9 LPMOs generated different oxidized product

profiles when subjected to different types of cellulose (**Chapter 2**). Thus, in addition to the more general cellulose-specificity of AA9 LPMOs, the exact type of cellulose used will define how these LPMOs bind to their substrate, and hence, define the product profile.

In addition to the distinct cellulose cleavage profiles, one of the most interesting findings in this thesis was that the use of different AA9 LPMOs resulted in different xyloglucan cleavage profiles: “inactive”, “substitution-intolerant” and “substitution-tolerant” (**Chapter 6 and 7**). These xyloglucan cleavage profiles were further correlated to the active site segment configurations of the enzymes, which were phylogenetically grouped into four clusters. Interestingly, *Mt*LPMOs and *Nc*LPMOs are spread over all four clusters (**Table 1 and 2**), indicating that *M. thermophila* and *N. crassa* can make use of different LPMOs with diverse functions and roles for cellulose and xyloglucan degradation.

It was also found that some xyloglucan-active LPMOs, such as *Nc*LPMO9C, *Nc*LPMO9M and *Mt*LPMO9E, generated a much higher yield of oxidized oligosaccharides from xyloglucan than from cellulose. Although this observation is based on a rough estimation and comparison of signal intensities in HPAEC, it still points out that xyloglucan is possibly the preferred substrate for some LPMOs. Of course the solubility of cellulose is lower than of xyloglucan, which may give less accessibility of cellulose that might explain the lower yield of degradation products. However, the xyloglucan preference over cellulose is also strengthened by HPSEC results, which showed that a more complete degradation of xyloglucan was observed for *Nc*LPMO9C, *Nc*LPMO9M (**Chapter 6, Fig. 2**) and *Mt*LPMO9E (**Fig. 1**). Nevertheless, the substrate preference indicates that different active site segment configurations may result in different binding affinity to cellulose, xyloglucan, and even cello-(xylogluco-)oligosaccharides and xylan. The obtained knowledge about the substrate preference can be used to target different LPMOs towards specific carbohydrate substrates to improve the degradation efficiency.

Next to xyloglucan activity, AA9 LPMO activity towards xylan has been demonstrated. *Ls*LPMO9A (from *Lentinus similis*), *Mc*LPMO9H (from *Malbranchea cinnamomea*) and *Mt*LPMO9A were shown to cleave xylan, but the latter two enzymes only catalyze the cleavage when xylan is associated with cellulose (13, 27, 28). In lignocellulose, xylan is the most abundant type of hemicellulose and is known to intertwist with cellulose to form a complex. Therefore, it is not surprising that fungi produce xylan-active LPMOs to open up the xylan structure in order to reach the cellulose. Until now, catalytic performance of LPMOs towards xylan, xylan degradation profiles and identification of oxidized xylo-oligosaccharides are not well understood. It has been speculated that xylan-active LPMOs also show xylan-specificity regarding the (in)tolerance of xylan substitutions, similar to the aforementioned xyloglucan-specificity (29).

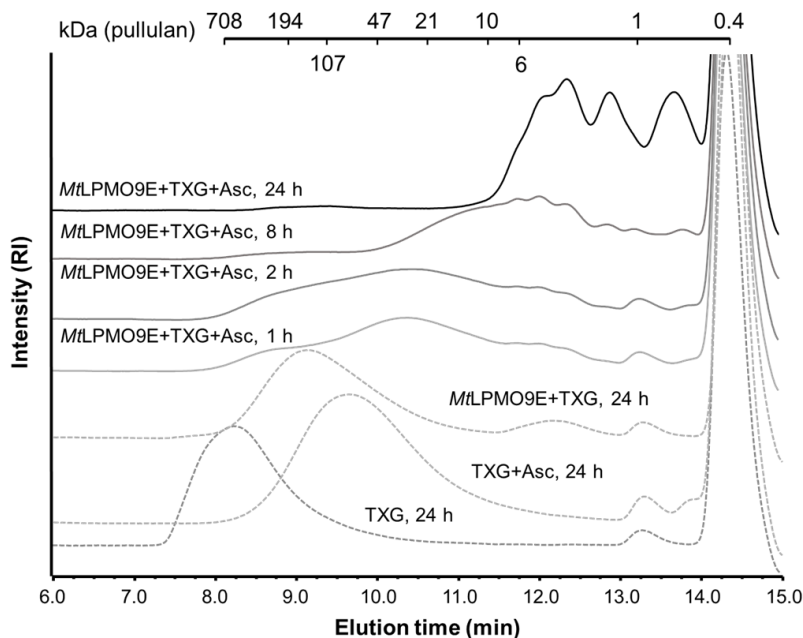


Fig. 1. Molecular weight (MW) distribution of tamarind seed xyloglucan (TXG) digests (24 h) in presence of ascorbic acid (Asc) from *MtlPMO9E*, analyzed by HPSEC-RI. MW distributions of control digests (only TXG, TXG+Asc, and *MtlPMO9E*+TXG; dotted lines) are also shown. MWs are indicated based on calibration with a series of pullulan standards (0.4–708 kDa).

Other hemicelluloses such as glucomannan and mixed β -(1 \rightarrow 3, 1 \rightarrow 4)-linked glucan can also be cleaved by some of the *MtlPMOs* and *NclPMOs*. Other fungi, such as *P. anserina* and *M. cinnamomea*, were also found to contain multiple LPMOs having different substrate specificity, and overall covering degradation of various hemicelluloses in addition to cellulose (27, 30). These fungi are known to degrade a variety of complex plant cell wall polysaccharides, and thus they are expected to obtain a powerful enzyme “toolbox”. In addition, it has been shown that varying the composition of substrates and growth conditions results in expression of different LPMOs (and other carbohydrate-active enzymes), again suggesting the diverse function and roles in degrading target polysaccharides (3, 27, 30–33). Overall, AA9 LPMO hemicellulose activities reveal their importance in “opening” or “loosening” complex lignocellulose structures. To fully map and understand their natural functions, it would be helpful to produce all AA9 LPMOs annotated in one fungal genome and characterize their catalytic repertoire.

2.2. N-terminal His1 methylation and glycosylation may influence LPMO activity and stability

As summarized in **Table 1 and 2**, *MtlPMOs* and *NclPMOs* showed comparable diversity in structural configuration, regioselectivity and substrate specificity. However, we also observed very different characteristics between *MtlPMOs* and *NclPMOs* used in this thesis. First, *MtlPMOs* were homologically produced in *M. thermophila* C1, while

NcLPMOs were heterologously produced in the yeast *P. pastoris*. As a result, *MtLPMOs* have a methylated N-terminal histidine (His1), while *NcLPMOs* do not have this modification (13, 21, 26) (**Chapter 5**). Petrović and coworkers showed that methylation of His1 in *TaLPMO9A* does not influence the substrate specificity, product profiles, and LPMO thermal and operational stability (18). The only significant difference suggested by Petrović and coworkers is that the methylation improves the resistance of the catalytic site against oxidative damage (18). When adding up to 200 mM H₂O₂, the non-methylated form of *TaLPMO9A* showed a faster reaction rate in the early time points (<10 min) compared to the methylated *TaLPMO9A*. However, after 30 min both types of *TaLPMO9A* became inactivated, but the inactivation was more severe in the non-methylated *TaLPMO9A*.

Table 3. Summary of N-terminal His1 methylation, glycosylation and DMP assay activity of *MtLPMOs* and *NcLPMOs* used in this thesis. In green: N-glycosylation occurs near the active site.

	DMP assay activity (U/g)	MW difference (kDa) ¹	Predicted N-glycosylation site ²	Predicted O-glycosylation site ³	Glycosylation occurrence based on MW difference	Residues coordinating the copper
<i>MtLPMO9B</i> ⁴	0.4	3.4	No	Ser222, Ser225, Thr226, Ser230, etc.	Yes, only O-type	MethHis1, His79, Tyr170
<i>MtLPMO9E</i> ⁵	0.3	0.7	Asn175	No	No	MethHis1, His83, Tyr168
<i>MtLPMO9F</i> ⁶	n.d.	4.1	No	Ser217, Ser227, Ser239, Ser241, etc.	Yes, only O-type	MethHis1, His84, Tyr167
<i>MtLPMO9H</i> ⁷	0.6	1.8	No	Ser212, Thr217, Ser229, Ser230, etc.	Yes, only O-type	MethHis1, His87, Tyr173
<i>MtLPMO9I</i> ⁸	0.4	1.1	Asn28	Ser30, Thr33, Ser34, Thr41	No	MethHis1, His84, Tyr158
<i>NcLPMO9C</i> ⁹	28.8	22.8	Asn173	Thr99, Ser183, Thr226, Ser232, etc.	Yes, both N- and O-type	His1, His83, Tyr166
<i>NcLPMO9F</i> ¹⁰	2.2 ¹²	0.8	No	Ser45, Thr46, Ser52	No	His1, His72, Tyr156
<i>NcLPMO9M</i> ¹¹	39.1	7.8	Asn6	Ser104, Thr128, Ser188	Yes, (likely) only N-type	His1, His83, Tyr171

¹Difference between theoretical molecular weight (MW) and apparent MW displayed by SDS-PAGE

²Determined on NetNGlyc 1.0 Server (34); ³Determined on NetOGlyc 4.0 Server (35)

⁴Theoretical MW 30.6 kDa, apparent MW 34 kDa; ⁵Theoretical MW 24.3 kDa, apparent MW 25 kDa;

⁶Theoretical MW 29.9 kDa, apparent MW 34 kDa; ⁷Theoretical MW 33.2 kDa, apparent MW 35 kDa;

⁸Theoretical MW 23.9 kDa, apparent MW 25 kDa; ⁹Theoretical MW 34.2 kDa, apparent MW 57 kDa;

¹⁰Theoretical MW 23.2 kDa, apparent MW 24 kDa; ¹¹Theoretical MW 24.2 kDa, apparent MW 32 kDa;

¹²Value is from Breslmayr et al., 2018 (36)

Apart from the methylation, *MtLPMOs* and *NcLPMOs* differ in type and degree of glycosylation (**Table 3**). In general, enzymes produced in filamentous fungi such as *M. thermophila*, *T. reesei* and *A. nidulans* usually have N-/O-linked glycosylation (37-39). N-glycosylation occurs on a tripeptide sequon, Asn-Xaa-Ser/Thr (Xaa is not a Pro), while O-glycosylation occurs on a Ser and/or a Thr residue (40, 41). *MtLPMO9B*, *MtLPMO9H* and *MtLPMO9F* that all have a linker and a CBM1 only had O-glycosylation (**Table 3**), which is often present in the linker (42-44). The difference between their theoretical molecular weight (MW) and apparent MW displayed by SDS-PAGE confirmed that the produced proteins were glycosylated. *MtLPMO9E* had only one predicted N-glycosylation site (Asn175). However, the MW difference was only 0.7 kDa, indicating that N-glycosylation was not present. Regarding *MtLPMO9I*, although the MW difference was 1.1 kDa and three O-glycosylation sites (no N-glycosylation site) were predicted, it is not expected that this protein is glycosylated, as O-glycosylation has been reported to hardly occur in catalytic domains of CAZymes (39, 43).

In *NcLPMO9C* and *NcLPMO9M*, the differences between theoretical and SDS-PAGE determined MWs were much larger (22.8 kDa and 7.8 kDa) compared to *MtLPMOs* (**Table 3**). It has been shown that enzymes produced in yeast (e.g., *P. pastoris*) have a higher degree of N-glycosylation or O-glycosylation if a linker is present (21, 30, 39, 42, 43, 45-48). Therefore, the MW differences observed with *NcLPMO9C* and *NcLPMO9M* are explained by the high degree of N-/O-glycosylation and N-glycosylation, respectively. The MW difference in *NcLPMO9F* was only 0.8 kDa, indicating that no O-glycosylation occurred, which was also observed from a previous batch of *NcLPMO9F* (PMO-03328) (21).

Taken the above described diverse glycosylation into account, it is intriguing to note that high DMP assay activity was observed for the two highly N-glycosylated *NcLPMOs* (**Table 3**). A detailed investigation showed that the N-glycosylation sites in *NcLPMO9C* (Asn173) and *NcLPMO9M* (Asn6) are close to key amino acid residues (Tyr166 in *NcLPMO9C* and His1 in *NcLPMO9M*) coordinating the copper ion in the active site. Although the predicted N-glycosylation site in *MtLPMO9E* (Asn175) is also close to the key residue Tyr168, the produced *MtLPMO9E* is not N-glycosylated (**Table 3**). Hence, possibly, N-glycosylation near the active site protects the enzyme against oxidative damage by reactive oxygen species, and thus might increase the LPMO activity and stability. The influence of (N-)glycosylation has been, so far, mainly studied for glycosidic hydrolases, which is comprehensively reviewed elsewhere (43). Overall, several studies showed that N-glycosylation increases the catalytic activity, thermal stability and enzyme solubility of glycosidic hydrolases (37, 39, 42-45, 47-51). Although the underlying mechanisms for explaining the effects of N-glycosylation on glycosidic hydrolases (and other types of enzymes) are not really clear, some studies suggest that N-glycosylation promotes the folding of protein and limits the conformational mobility either around the active site or the whole protein structure, and thus increases the stability of the enzyme (39, 48, 52, 53). These features could also hold for LPMOs, though LPMOs have very distinct active site structures, catalytic mechanisms and co-substrates than glycosidic hydrolases. Nevertheless, more

research for understanding the effect of glycosylation on LPMOs is highly recommended. Hereto, it would be of interest to compare the inactivation in RAC-digests of the same LPMO having or lacking N-glycosylation near the active site. Production of the latter LPMO could be achieved via replacing the target asparagine by site-directed mutagenesis.

2.3. Functional subclassification of the AA9 family

The classification of CAZymes is based on the amino acid sequence similarity including at least one biochemically characterized family member (54, 55). Although the overall function of a family can be predicted based on the characterized members, the apparent substrate specificity of each individual candidate can be very different. The difficulties in predicting substrate specificity, and even functional roles, are especially problematic for AA9 LPMOs. Based on amino acid sequence similarity, AA9 (formerly GH61) family members all have a conserved N-terminal His residue, a second His residue in a "His-Xaa-Gly-Pro" motif, and a catalytic site Tyr residue in the "Gln-Xaa-Tyr-Xaa-Xaa-Cys" motif (**Chapter 1**) (9). However, AA9 LPMOs share low overall sequence identity and high diversity of regioselectivity and substrate specificity (**Chapter 5–7**, (9, 12, 20, 27, 30, 56–60). Therefore, establishment of subfamilies of the AA9 family is recommended to better group and predict the function of individual LPMO, which can be based on similarities in protein structures, and/or based on function. Busk and coworkers developed an algorithm using a peptide pattern recognition (PPR) approach based on a set of short conserved sequences (58). By using this algorithm, AA9 is divided into sixteen subfamilies, which differ in the amino acid regions 100–114, 115–119 and 160–172 (58). However, this classification did not receive much attention in later research, probably due to the complicated classification principle and lack of correlation to the observed substrate specificity and regioselectivity. Vu and coworkers classified AA9 LPMOs into four subgroups (PMO1, C1 oxidation; PMO2, C4, oxidation; PMO3, C1 and C4 oxidation; and PMO3*, C1 oxidation) based on sequence similarity and regioselectivity (9). However, incongruous regioselectivity has been reported in each subgroup. For example, *Pa*LPMO9H classified in subgroup PMO2 but it shows C1- and C4-oxidation (30). *Mt*LPMO9D belongs to PMO3 subgroup shows only C1-oxidation (1). Furthermore, this subclassification does not provide insight into substrate specificity, since hemicellulosic activities were not well recognized at that time. In addition to the classification based on sequence similarity, the Enzyme Commission number (EC number) classification (61) categorizes AA9 LPMOs based on regioselectivity and substrate specificity, including lytic cellulose monooxygenase (C1-hydroxylating, EC 1.14.99.54), lytic cellulose monooxygenase (C4-dehydrogenating, EC 1.14.99.56), and a preliminary lytic xyloglucan monooxygenase (EC 1.14.99.B10) (61). This EC number classification system has so far only provided very brief characteristics of AA9 LPMOs, from which it is difficult to predict the function and catalytic performance of putative candidates.

In this thesis, based on the work of our collaborators from the Ludwig Lab in Vienna, AA9 LPMOs were divided based on active site segment (Seg1 to Seg5) configuration and phylogeny, which seemed useful to predict xyloglucan specificity (**Chapter 5–7**)

(62-64). Based on such structural features, we propose to classify AA9 LPMOs into four subfamilies (Level 1): AA9_1: \sim Seg1 \sim Seg2, only active on cellulose; AA9_2: \sim Seg1 \sim Seg2, active on cellulose and xyloglucan (substitution-tolerant); AA9_3: \sim Seg1 \sim Seg2, active on cellulose and xyloglucan (substitution-intolerant); AA9_4: \sim Seg1 \sim Seg2 \sim Seg3, only active on cellulose. Moreover, it is of interest to look further into xylan specificity, and whether this correlates to the suggested active site segment classification. As described above, the xylan-active LPMOs could also gain xylan-specificity regarding the (in)tolerance of xylan substitutions, which may also be modulated by the active site segment configuration. If other pronounced hemicellulose specificities are recognized in the future, the corresponding subclasses can be added later. For example, the hemicellulose specificity on glucomannan, though it was less studied, could also be correlated to the active site segment configuration.

The proposed subclassification highlights the importance of structure-based similarity and substrate (binding) specificity, which are critical characteristics for studying and applying enzymes. Subclassification based on substrate specificity has already been applied in the AA3 glucose-methanol-choline (GMC) oxidoreductases family that is divided into 4 subfamilies: AA3_1 (mostly cellobiose dehydrogenases), AA3_2 (including both aryl-alcohol oxidase and glucose 1-oxidase), AA3_3 (alcohol oxidase) and AA3_4 (pyranose 2-oxidase) (55). As the subclassification of AA9 LPMOs requires a more comprehensive characterization of substrate specificity, the development of high-throughput enzyme expression, production, purification and substrate characterization approaches is highly recommended. In future research, it would be desirable to establish a "substrate-heatmap" which can be added in the CAZy database not only for AA9 LPMOs but also for other CAZymes.

3. Improved chromatographic approach for analysis of oxidized oligosaccharides

Until the beginning of this PhD project, analysis of LPMO-generated (oxidized) oligosaccharides mostly relied on HPAEC-PAD and MALDI-TOF-MS or direct infusion ESI-MS. HPAEC-PAD cannot be used for structural determination since oxidized oligosaccharide standards are not always available, and MALDI-TOF-MS (or direct infusion ESI-MS) cannot distinguish isomers (**Chapter 1**). Other techniques such as PGC-MS, RP-UHPLC-MS and HILIC-MS were also used, but their application for oligosaccharide separation and structural determination was far from optimal. In **Chapter 3, 4 and 6**, we developed and applied a HILIC-ESI-CID-MS/MS² method, with and without prior NaBD₄ reduction, to analyze LPMO-cellulose or xyloglucan digests. By using this methodology, we are now able to separate complicated mixtures of non-oxidized and oxidized oligosaccharides, from which the novel C4/C6-double oxidized cello-oligosaccharides (**Chapter 4**) and the C4-oxidized glucosyl units substituted with xylosyl (and galactosyl) residues (**Chapter 6 and 7**) have been identified for the first time. Compared to the methods and assays summarized in **Chapter 1 Table 3**, HILIC provides sufficient separation (without reduction) and much better separation (with reduction) of especially C4-oxidized oligosaccharides (in different isomeric forms). In addition, CID-MS/MS² generates signature fragmentation

patterns of different types of oligosaccharides for unambiguous structural elucidation. Although the above achievement largely promotes the characterization of LPMO degradation products, the following points are of interest to be further addressed and developed in future research:

I) Without reduction, C4-oxidized (xylogluco-)oligosaccharides co-elute together (and with α/β -anomers of non-oxidized ones), and elute as rather broad peaks (>3 min) in HILIC, which complicates their identification. However, non-reduced oligosaccharides display fragmentation mass spectra having less fragments and more predictable patterns, compared to the reduced oligomers, which facilitates the identification. Therefore, improving the separation of non-reduced C4-oxidized oligosaccharides is still desired in HILIC. Although after reduction of the oligosaccharides a much better separation is achieved, reduction of C4-oxidized oligosaccharides still results in two types containing either galactosyl or glucosyl residues at the non-reducing end, which again complicates their separation and analysis. One way to improve the separation of the oligosaccharides with galactosyl or glucosyl ends is modifying the elution conditions. In this thesis, we have tried different combinations of solvents (e.g., water/methanol), elution profiles, column temperature and flowrates. However, we only tested amide HILIC column. Other resin types, such as polyethyleneimine (PIE) and silica can be investigated in the future research. On the other hand, it is worth to investigate whether the reduction reaction can be steered to either glucosyl or galactosyl ends by using a different reducing agent (e.g., lithium aluminum hydride) (65).

II) Quantification of LPMO-oxidized products is highly important for studying LPMOs, for example, to compare the kinetics, catalytic activity and performance over time. Quantification of C1-oxidized products released from cellulosic substrates has been well developed based on the gluconic acid and cellobionic acid, as they are stable in HPAEC condition and standards are commercially available (1) (**Chapter 2**). In addition, C1-oxidized hemicellulosic oligosaccharides are not well defined and determined so far, hence, the requirement of C1-oxidized hemicellulosic oligosaccharide standards becomes a point of attention for the near future. However, quantification of C4-oxidized oligosaccharides has been a major challenge for studying C4-oxidizing LPMOs. C4-oxidized products are not stable in HPAEC conditions, and no C4-oxidized (mono-)oligosaccharide standards (both oxidized cello- and hemicello-oligosaccharides) are available. The HILIC-MS technique has the potential to semi-quantify reduced C4-oxidized oligosaccharides, which can be done by quantification of the MS peak areas of individual products. This quantification of MS peak areas in HILIC-MS is comparable to the of quantitation of peak areas of C1-oxidized cello-oligosaccharides in HPAEC as described in **Chapter 2**. The advantage for quantification of MS peak areas is that this can be applied for more complicated e.g., C4-oxidized xylogluco-oligosaccharides, as their structures are known from MS/MS² spectra. Once the separation of non-reduced C4-oxidized oligosaccharides is much improved, it can also be applied for the quantification of non-reduced products directly. Regarding the quantification of C4-oxidized ends released from cellulose, the most widely used HPAEC technique for (mono)oligosaccharide quantification cannot be used, since C4-oxidized

ends are not stable in alkaline eluents (17, 23). Again, reduction provides an opportunity as the quantification of C4-oxidized ends can be done by quantification of galactose assuming that the equilibrium shifts to galactose completely.

4. Production and consumption of H₂O₂: the dual ability of AA16 oxidoreductases and AA9 LPMOs

In **Chapter 8**, two copper-dependent oxidoreductases from the AA16 family (*MtAA16A* and *AnAA16A*) were studied. They were found to not oxidatively cleave carbohydrates, though they share similarities with (AA9) LPMOs (e.g., His-brace and copper-dependent active site). We also found that in the presence of a proper reducing agent, both AA9 LPMOs and AA16 oxidoreductases produce H₂O₂, and can oxidize a range of aromatic compounds (**Chapter 8**). This dual ability of production and consumption of H₂O₂ is also observed with LPMOs from other AA families.

Initially, LPMOs have been revealed to carry out a monooxygenase reaction with polysaccharide substrates meaning that one oxygen atom from dioxygen (O₂) is inserted into reaction products (66, 67). In 2013, Kittl and coworkers found that LPMOs generate H₂O₂ as a side-product once the Cu(II) is reduced and no proper carbohydrate substrate is present (21). The mechanism of this LPMO side reaction has been explained by the conversion of the Cu(I)···hydroperoxide intermediate to Cu(II) and H₂O₂ when no substrate is available (68). Based on this side reaction, Kittl and coworkers developed the so-called Amplex Red assay for LPMOs (21). This assay measures the level of LPMO-produced H₂O₂ by using the horseradish peroxidase (HRP) coupled conversion of Amplex Red to resurofin (21). Subsequently, this assay has been widely used for screening the substrate specificity and kinetics of LPMOs (21, 30, 69–80). In 2018, Breslmayr and coworkers found that AA9 LPMOs were able to oxidize various aromatic compounds such as gallic acid, sinapic acid and syringol in the presence of H₂O₂, similar to a peroxidase reaction (36). According to the authors, this research was inspired by the finding that H₂O₂ can act as a co-substrate for the LPMO-mediated oxidative cleavage of polysaccharides. However, the latter reaction was proven to be a peroxygenase reaction (81). Therefore, it remains questionable if the Amplex Red assay truly reflects the LPMO H₂O₂-producing ability. In principle, the peroxidase-like ability of LPMOs (or AA16s in **Chapter 8**) means that they also convert Amplex Red by using the self-produced H₂O₂, which will influence the outcome of the assay. Because of that, we used a non-enzymatic Ferric-xylenol assay (**Chapter 8**), which directly monitors H₂O₂ levels. However, H₂O₂ assay will never reflect the actual LPMO activity due to the dual H₂O₂ producing and consuming ability. Therefore, direct measurement of oxidized degradation products as emphasized in this thesis is highly recommended.

In addition to the unclear peroxidase reaction mechanism, another intriguing question is what the determinants are to steer the dual H₂O₂ producing and consuming ability of LPMOs and AA16s. It can be speculated that the dual ability depends on the types of reducing agents and the H₂O₂ level in the reaction bulk solution. Regarding the H₂O₂ level, it can be assumed that the peroxidase reaction is preferred at high H₂O₂

concentrations, while H_2O_2 production is preferred at low H_2O_2 concentrations. Furthermore, it remains to be answered how the LPMO and AA16 peroxidase reactions correlate with or contribute to lignocellulose degradation. The further understanding of this dual ability will provide more insights into the functions and roles of these copper-dependent oxidoreductases.

5. Lignin-derived aromatic compound/ H_2O_2 -mediated AA16 oxidoreductase and AA9 LPMO reaction

As discussed in the previous section, AA16 oxidoreductases and AA9 LPMOs can catalyze the peroxidase-mediated oxidation of syringol. As syringol is a lignin-derived aromatic compound, we were curious if the peroxidase reaction is still carried out by AA9 LPMOs in the presence of cellulose. Intriguingly, oxidative cleavage of RAC was observed in the incubation of *Mt*LPMO9B (1 μM) with RAC (2 mg/mL), syringol (Syr, 1 mM) and exogenous H_2O_2 (100 μM), while no oxidized products were detected in the same incubation without exogenous H_2O_2 (**Fig. 2**). Similar results were found for *Nc*LPMO9C and *Nc*LPMO9M. Based on these observations, it was firstly hypothesized that RAC-unbound LPMO (in **Chapter 2** we showed that around 40% of *Mt*LPMO9B did not bind to RAC) carries out a peroxidase reaction to generate syringol radicals that oxidatively cleave RAC. However, replacement of *Mt*LPMO9B with catalytic amounts of HRP did not result in the oxidative cleavage of RAC. In addition, laccase is known to oxidize syringol using O_2 to generate syringol radicals (82, 83). Therefore, we also replaced *Mt*LPMO9B with catalytic amounts of the purified laccase from *Trametes versicolor* (84, 85). Again, no oxidative cleavage of RAC was observed. These results indicate that syringol radicals do not directly cause the oxidative cleavage of RAC, and suggest that oxidative cleavage of RAC results from the LPMO reaction which is initiated with the reduction of Cu(II) by syringol/ H_2O_2 , but not syringol alone.

Because *Mt*AA16A and *An*AA16A can boost *Mt*LPMO9B activity towards RAC (**Chapter 8**), we performed the same incubations with syringol in the presence of 1 μM *Mt*AA16A or 1 μM *An*AA16A. Without addition of H_2O_2 , no oxidative cleavage of RAC was observed in two *Mt*LPMO9B-AA16-RAC incubations, supporting that syringol radicals do not induce the cleavage of RAC as described above (**Fig. 3**). With addition of H_2O_2 , the boosting effect of both AA16s on *Mt*LPMO9B-RAC cleavage was found again. However, the underlying mechanism of the boosting effect seems to differ from that described in **Chapter 8**, where we found that H_2O_2 -production from the reduced Cu(II) in AA16s activates the AA9 LPMOs, helped by a putative LPMO-AA16 protein interaction. Here, the boosting effect was found in the presence of both H_2O_2 and syringol. It is hard to believe that syringol with H_2O_2 reduces the Cu(II) in both AA16s and *Mt*LPMO9B, followed by subsequent production of H_2O_2 from the reduced AA16s to fuel *Mt*LPMO9B, which are two contradictory pathways. Therefore, the hypothesis that syringol/ H_2O_2 reduces the AA9 LPMOs is still not proven. In addition, if the hypothesis is correct, how syringol/ H_2O_2 reduce the AA9 LPMOs (and AA16s) and whether it correlates to their peroxidase activity remains to be elucidated. The latter correlation is partially disapproved by the observation that no oxidative products were found in the incubations where AA16s were replaced with HRP.

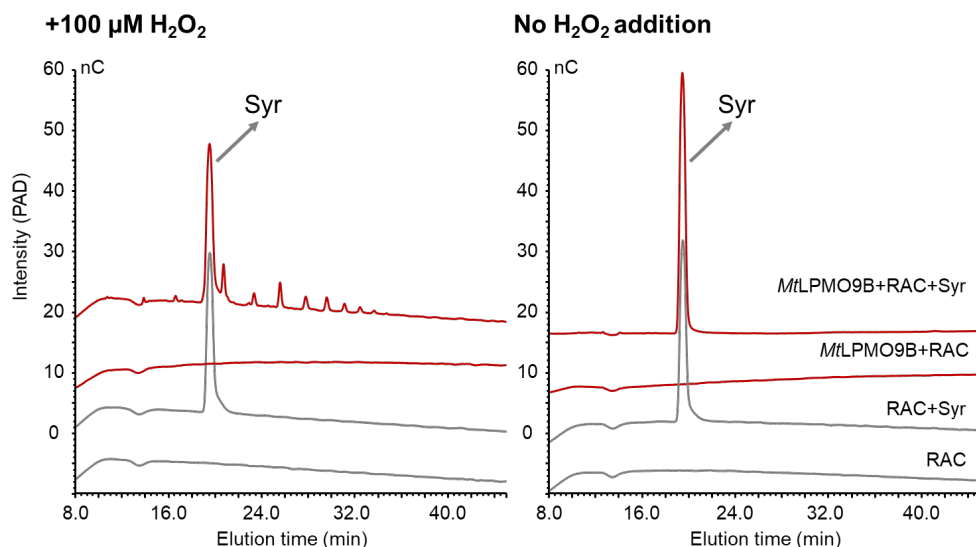


Fig. 2. HPAEC chromatograms of RAC samples incubation with *MtLPMO9B* (1 μ M) in the absence or presence of 1 mM syringol (Syr) with the addition of 100 μ M H₂O₂ (left) or no addition of H₂O₂ (right).

We further tested if other lignin-derived compounds/H₂O₂ in the incubation of *MtLPMO9B* with or without *MtAA16A* or *AnAA16A* can stimulate the oxidative cleavage of RAC. In **Chapter 8**, *MtAA16A* has been shown to oxidize other syringol-like compounds in the presence of H₂O₂ including syringic acid, sinapyl alcohol, sinapaldehyde and sinapic acid, which all have 4-hydroxy-3,5-dimethoxy moiety. However, replacement of syringol with these compounds (and guaiacol and ferulic acid) did not result in the oxidative cleavage of RAC with exogenous H₂O₂. These unexpected results reflect the complex synergy between LPMOs and AA16s, which deserves more study to understand their roles in lignocellulose degradation. A similar finding was reported by Hegnar and coworkers, who showed that 2,3-dihydroxybenzoic acid (2,3-DHBA) can drive *GtLPMO9B* reaction towards PASC only with the addition of H₂O₂ (86). However, they demonstrated that 2,3-DHBA can directly reduce the *GtLPMO9B* at pH 8.0 and 9.0 (pH range of 6.0–9.0 tested) evidenced by the detection of H₂O₂ formed. As all experiments described in **Chapter 8** were performed at pH 5.0, we do not know if the above mentioned lignin-derived aromatic compounds will serve as reductants at higher pH. Although mechanistic insights were not investigated and discussed by Hegnar and coworkers, the pH dependency of the LPMO-(AA16)-reductant system provides a direction for future study.

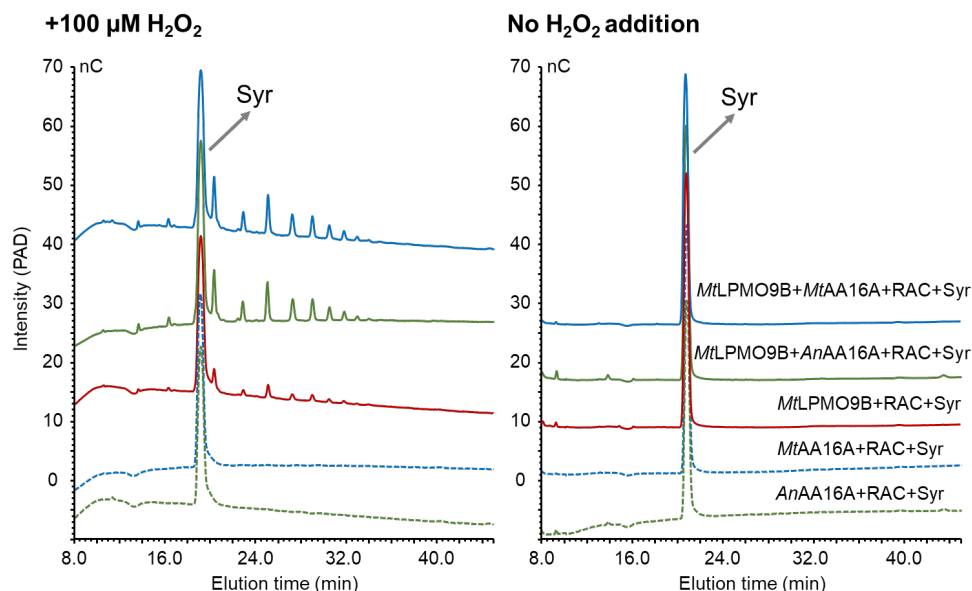


Fig. 3. HPAEC chromatograms of RAC samples incubation with *MtLPMO9B* (1 μ M), *MtLPMO9B* (1 μ M)+*MtAA16A* (1 μ M), and *MtLPMO9B* (1 μ M)+*AnAA16A* (1 μ M) in the absence or presence of 1 mM syringol (Syr) with the addition of 100 μ M H_2O_2 (left) or no addition of H_2O_2 (right).

The above discussed syringol/ H_2O_2 -mediated LPMO and AA16 reactions emphasize the importance of H_2O_2 in lignocellulose degradation. In **Chapter 8**, the correlation of AA9 LPMOs and AA16s with other AA enzymes was discussed. Here, H_2O_2 producing enzymes in these AA families will be highlighted, based on the excellent review from Bissaro and coworkers (87), together with some recently published work. Extracellular H_2O_2 producing enzymes in lignocellulolytic systems reported so far include cellobiose dehydrogenase (CDH; AA3_1) (15, 24, 88-90); glucose oxidase (GOX; AA3_2) (91, 92), aryl-alcohol oxidase (AAO; AA3_2) (93), pyranose 2-oxidase (P2O; AA3_4) (94) and methanol oxidase (MOX; AA3_3) (95). These enzymes belong to the so called glucose-methanol-choline (GMC) superfamily of flavoenzymes (Section 1.3) (55, 96). In addition to AA3 GMC enzymes, for instance, glyoxal oxidase (GLOX; AA5_1) (97-100) alcohol oxidase (AlcOx; AA5_2) (101), and oligosaccharide oxidases (AA7) (102-104) have been described to produce H_2O_2 using various substrates. In a recent study, AA7 cellooligosaccharide dehydrogenase was discovered and its interplay with LPMO as an electron donor in cellulose degradation has been demonstrated, though it may also produce H_2O_2 to fuel the LPMO reaction (103). In another study, Perna and coauthors reported that laccase-catalyzed oxidation of lignin leads to the production of up to 25 μ M H_2O_2 , which can sufficiently fuel LPMO in one-pot reactions (105).

H_2O_2 has also been described to be used by lignin-active peroxidases for lignin degradation. These lignin-degrading enzymes include lignin peroxidases (LiPs), manganese peroxidases (MnPs) and versatile peroxidases (VPs), and their structures, functions and catalytic reactions are reviewed elsewhere (109). It can be imagined that the H_2O_2 produced can be used for two routes: i) lignin-degradation and ii) LPMO

(and AA16) oxidative degradation on recalcitrant polysaccharides using the degraded lignin fragments (e.g., syringol-like structures).

6. LPMOs in lignocellulosic biomass degradation and valorization: current status and future perspectives

6.1. Saccharification of lignocellulosic biomass for production of biofuels and biochemicals

6.1.1. Various commercial cellulase cocktails already supplemented with LPMOs

In lignocellulosic biomass, three major components cellulose, hemicellulose and lignin are intertwined and cross-linked in a complex network. Therefore, lignocellulosic biomass is usually subjected to a pretreatment to increase the accessibility for enzymes in the subsequent saccharification process (110-113). Saccharification of lignocellulosic biomass is widely performed by commercial enzyme cocktails, which are produced by (genetically modified) filamentous fungi (28, 114-117). Already in 2007, the addition of an AA9 LPMO (previously GH61B from *T. terrestris*) to a classical cellulolytic enzyme cocktail Celluclast® 1.5 L was found to increase the yield of cellulose degradation up to 1.4-fold (118). Afterwards, AA9 LPMOs were supplemented to commercial cellulolytic enzyme cocktails such as Cellic® CTec2 and Cellic® CTec3 (119-122). Soon thereafter, synergy of LPMOs with cellulolytic hydrolases including endoglucanases (EG), cellobiohydrolases (CBH) and β -glucosidases (35) became a hot topic of investigation (2, 14, 62, 122-131). Most of these studies have shown that the addition of LPMOs increases the cellulose degradation with a reduced overall enzyme dose, which is a major cost benefit in industrial applications (132, 133). Detailed outcomes and comparison of these studies are comprehensively reviewed elsewhere (87, 134-138).

6.1.2. Application-tailored (hemi-)cellulase cocktails supplemented with specific LPMOs

It can be expected that future (LPMO-supplemented) enzyme cocktails for saccharification in biorefinery aim for a “cost-effective” process, meaning to further improve the polysaccharide degradation efficiency with decreased enzyme loadings. This goal might be partially achieved, as suggested by the outcomes and findings in this thesis, by selecting and incorporating LPMOs having the “appropriate” regioselectivity and substrate specificity.

The LPMO-directed regioselective oxidation of cellulose can influence its synergy with different glycosidic hydrolases. It has been reported that reducing end active GH7 CBHs display an impeded effect with a C1-(reducing end) oxidizing LPMO *Tt*AA9 (139), but a boosting effect with a C1/C4-oxidizing *Ta*AA9A and a C4-oxidizing *Ls*AA9A (130). It has also been shown that C1-oxidizing LPMOs obtain strong synergy with non-reducing end active CBHs from the GH6 family (140). In addition, double C4/C6-oxidation described in **Chapter 4** may also influence the cellulose degradation, as C6-

oxidation was reported to have a negative effect on cellulase activity (141-143). These studies strongly indicate that the selection of “appropriate” regioselective LPMOs in order to control the “correct” oxidation position should be taken into account, which is essential for supplementation of LPMOs in enzyme cocktails.

Beyond the regioselectivity, the different cellulose cleavage product profiles of LPMOs described in **Chapter 2** could also influence the synergy with cellulases. There were already studies indicating that the LPMO boosting or impeding effects on cellulases depend on the types of cellulose, and it is supposed that these effects could be moderated by the processivity of cellulases and by the mode-of-action of LPMOs (14, 126, 130, 144, 145). As shown in **Chapter 2**, oxidized degradation product profiles from different cellulose types reflect the catalytic behavior of LPMOs, providing the “reference” for designing, e.g., proper cellulase cocktails to degrade different cellulose sources.

Design of application-tailored enzyme cocktails can benefit from the characterization of hemicellulose specificity and further assignments of hemicellulosic mode-of-action. It is known that cellulose is still associated with hemicelluloses such as xylan, xyloglucan and glucomannan even after pretreatment (146-148), and that the additional hemicellulosic side-activities from cellulases might contribute to the overall cellulose degradation (149-151). Therefore, LPMOs with broad substrate specificity especially active on (cellulose-associated) xylan, such as *MtLPMO9A*, *LsLPMO9A* and *McLPMO9H*, or active on glucomannan such as *MtLPMO9E* and *NcLPMO9C*, are interesting to study in relation to cellulose degradation from real lignocellulosic biomass. In addition, supplementation of different xyloglucan-active LPMOs in hemicellulose cocktails also has great potential for primary cell wall- or xyloglucan-rich biomass degradation.

6.2. Production of (novel) functional materials and compounds

6.2.1. LPMO-assisted nanocellulose production

In addition to lignocellulose degradation, LPMOs have been used to assist the production of functional materials and compounds. Several studies have investigated the LPMO-assisted production of nanocellulose (152-161), which is a valuable material used as emulsion stabilizer and gas barrier, also even in food industry (162-165).

6.2.2. LPMO-oxidized oligosaccharides as potential health promoters

In a recent study, Li and coauthors reported that LPMO produces oxidized Konjac glucomannan oligosaccharides having prebiotic effects (166). They showed that addition of oxidized Konjac glucomannan oligosaccharides largely increases growth of faecal microbiota and production of short chain fatty acids in *in vitro* fermentation (166). Similar to glucomannan oligosaccharides, xylogluco-oligosaccharides are also potential prebiotics (167-169). It would be of interest to investigate the interaction of LPMO produced oxidized xylogluco-oligosaccharides with gut microbiota, and elucidate the structure-specific prebiotic effects based on the characterized oxidized xylogluco-oligosaccharide structures described in **Chapter 5 and 6**. In the last month of this

PhD project, AA17 LPMOs have been discovered to oxidatively cleave pectin (homogalacturonan) (170). As pectin oligosaccharides have been shown to provide health benefits by influencing gut microbiota composition and metabolism, it can be expected that LPMO-oxidized pectin oligosaccharides have potential as novel prebiotics as well. (171-175).

6.2.3. Supplementation of LPMOs in fiber-degrading enzyme cocktails as feed additives

In animal nutrition, fiber-degrading enzyme cocktails enriched in xylanase and glucanase have been used as feed additives. These enzymes are active towards non-starch polysaccharides, which are suggested to decrease the overall digesta viscosity to stimulate the diffusion of nutrients, and to improve the accessibility for starch digestion (176-180). In addition, fiber-degrading enzymes release, e.g., arabinoxylan- and β -glucan-oligosaccharides, which can be fermented by microbiota in gastrointestinal tract, and hence, as described above, promote gut health of animals (178-181). It can be expected that LPMOs have great potential to be supplemented with current fiber-degrading enzymes, since (hemicellulose-active) LPMOs not only promote opening-up of non-starch polysaccharide structures, but might also release potential functional and health-promoting oligosaccharides.

6.2.4. Exploration of LPMO-oxidized oligosaccharides in food, pharmaceutical and medical applications

So far, LPMO-oxidized oligosaccharides have not been well investigated for food, pharmaceutical and medical applications. One of the reasons is that the yield of oxidized oligosaccharides is rather low. Once the yield can be significantly increased by, for example, protein engineering of LPMOs, it can be expected that these LPMO-oxidized oligosaccharide may serve as novel and functional ingredients. For example, it is suggested that starch-derived oxidized oligosaccharides (aldonic acids) can be applied for cosmetic and dermatological uses (182). Due to the similar structure, oxidized cello- and chito-oligosaccharide aldonic acids might be used for comparable applications. Production of these valuable compounds from (plant cell wall) biomass instead of (edible) starch crops could be of great interest for future sustainable industrial applications.

All in all, LPMOs have huge potential to be applied in industrial lignocellulosic biomass degradation and valorization, and their further characterization presents a challenge for future research.

7. References

1. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, et al. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monoxygenases. *Appl Microbiol Biotechnol*. 2017;102:1281-95.
2. Karnaouri A, Muraleedharan MN, Dimarogona M, Topakas E, Rova U, Sandgren M, et al. Recombinant expression of thermostable processive MteG5 endoglucanase and its synergism with MtlPMO from *Myceliophthora thermophila* during the hydrolysis of lignocellulosic substrates. *Biotechnol Biofuels*. 2017;10:126.
3. Grieco MAB, Haon M, Grisel S, de Oliveira-Carvalho AL, Magalhaes AV, Zingali RB, et al. Evaluation of the enzymatic arsenal secreted by *Myceliophthora thermophila* during growth on sugarcane bagasse with a focus on LPMOs. *Front Bioeng Biotechnol*. 2020;8:1028.
4. Sun P, Valenzuela SV, Chunkruea P, Pastor FJ, Laurent CVFP, Ludwig R, et al. Oxidized product profiles of AA9 LPMOs depend on the type of cellulose. *ACS Sustain Chem Eng*. 2021;9:14124-33.
5. Sun P, Laurent CVFP, Ludwig R, van Berkel WJH, Kabel MA. Regioselective C4 and C6 double oxidation of cellulose by lytic polysaccharide monoxygenases. Chapter 4, this PhD thesis.
6. Sun P, de Munnik M, van Berkel WJH, Kabel MA. Extending the diversity of *Myceliophthora thermophila* LPMOs: two different xyloglucan cleavage profiles. Chapter 7, this PhD thesis.
7. Sun P, Kadowaki MAS, Veersma R, Hilgers R, Magri S, Laurent CVFP, et al. Copper-dependent AA16 oxidoreductases boost cellulose active AA9 LPMOs from *Myceliophthora thermophila* C1. Chapter 8, this PhD thesis.
8. Frommhagen M, van Erven G, Sanders M, van Berkel WJH, Kabel MA, Gruppen H. RP-UHPLC-UV-ESI-MS/MS analysis of LPMO generated C4-oxidized gluco-oligosaccharides after non-reductive labeling with 2-aminobenzamide. *Carbohydr Res*. 2017;448:191-9.
9. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monoxygenases. *J Am Chem Soc*. 2014;136:562-5.
10. Frommhagen M. Lytic polysaccharide monoxygenase from *Myceliophthora Thermophila* C1 [PhD thesis]: Wageningen University and Research; 2017.
11. Sun P, Frommhagen M, Kleine Haar M, van Erven G, Bakx E, van Berkel WJH, et al. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydr Polym*. 2020;234:115917.
12. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken J-P, et al. Lytic polysaccharide monoxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels*. 2016;9:186.
13. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, et al. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monoxygenase. *Biotechnol Biofuels*. 2015;8:101.
14. Zhou H, Li T, Yu Z, Ju J, Zhang H, Tan H, et al. A lytic polysaccharide monoxygenase from *Myceliophthora thermophila* and its synergism with cellobiohydrolases in cellulose hydrolysis. *Int J Biol Macromol*. 2019;139:570-6.
15. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monoxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6:1399-406.
16. Li X, Beeson IV WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monoxygenases. *Structure*. 2012;20:1051-61.
17. Westereng B, Arntzen MO, Achmann FL, Varnai A, Eijssink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monoxygenases acting on cellulose. *J Chromatogr A*. 2016;1445:46-54.
18. Petrovic DM, Bissaro B, Chylenski P, Skaugen M, Sorlie M, Jensen MS, et al. Methylation of the N-terminal histidine protects a lytic polysaccharide monoxygenase from auto-oxidative inactivation. *Protein Sci*. 2018;27:1636-50.
19. Courtade G, Wimmer R, Rohr AK, Preims M, Felice AKG, Dimarogona M, et al. Interactions of a fungal lytic polysaccharide monoxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc Natl Acad Sci U S A*. 2016;113:5922-7.
20. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Varnai A, et al. Structural and functional characterization of a lytic polysaccharide monoxygenase with broad substrate specificity. *J Biol Chem*. 2015;290:22955-69.

21. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5:79.
22. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, et al. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci U S A*. 2014;111:6287-92.
23. Isaksen T, Westereng B, Achmann FL, Agger JW, Kracher D, Kittl R, et al. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289:2632-42.
24. Kracher D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K, et al. Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science*. 2016;352:1098-101.
25. Sun P, Laurent CVFP, Scheiblbrandner S, Frommhagen M, Kouzounis D, Sanders MG, et al. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnol Biofuels*. 2020;13:95.
26. Laurent CVFP, Sun P, Scheiblbrandner S, Csarman F, Cannazza P, Frommhagen M, et al. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *Int J Mol Sci*. 2019;20:6219.
27. Huttner S, Varnai A, Petrovic DM, Bach CX, Anh DTK, Thanh VN, et al. Specific xylan activity revealed for AA9 lytic polysaccharide monooxygenases of the thermophilic fungus *Malbranchea cinnamomea* by functional characterization. *Appl Environ Microbiol*. 2019;85:e01408-19.
28. Simmons TJ, Frandsen KEH, Ciano L, Tryfona T, Lenfant N, Poulsen JC, et al. Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat Commun*. 2017;8:1064.
29. Kolenová K, Vršanská M, Biely P. Mode of action of endo- β -1, 4-xylanases of families 10 and 11 on acidic xyloligosaccharides. *J Biotechnol*. 2006;121:338-45.
30. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnol Biofuels*. 2015;8:90.
31. de Gouvêa PF, Gerolamo LE, Bernardi AV, Pereira LMS, Uyemura SA, Dinamarco TM. Lytic polysaccharide monooxygenase from *Aspergillus fumigatus* can improve enzymatic cocktail activity during sugarcane bagasse hydrolysis. *Protein Pept Lett*. 2019;26:377-85.
32. Jagadeeswaran G, Gainey L, Prade R, Mort AJ. A family of AA9 lytic polysaccharide monooxygenases in *Aspergillus nidulans* is differentially regulated by multiple substrates and at least one is active on cellulose and xyloglucan. *Appl Microbiol Biotechnol*. 2016;100:4535-47.
33. Ladeveze S, Haon M, Villares A, Cathala B, Grisel S, Herpoel-Gimbert I, et al. The yeast *Geotrichum candidum* encodes functional lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2017;10:215.
34. Gupta R, Brunak S, editors. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*; 2001.
35. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KTB, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J*. 2013;32:1478-88.
36. Breslmayr E, Hanzek M, Hanrahan A, Leitner C, Kittl R, Santek B, et al. A fast and sensitive activity assay for lytic polysaccharide monooxygenase. *Biotechnol Biofuels*. 2018;11:79.
37. Rubio MV, Zubieta MP, Cairo JPLF, Calzado F, Leme AFP, Squina FM, et al. Mapping N-linked glycosylation of carbohydrate-active enzymes in the secretome of *Aspergillus nidulans* grown on lignocellulose. *Biotechnol biofuels*. 2016;9:1-19.
38. Visser H, Joosten V, Punt PJ, Gusakov AV, Olson PT, Joosten R, et al. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. *Ind Biotechnol*. 2011;7:214-23.
39. Qin YQ, Wei XM, Liu XM, Wang TH, Qu YB. Purification and characterization of recombinant endoglucanase of *Trichoderma reesei* expressed in *Saccharomyces cerevisiae* with higher glycosylation and stability. *Protein Expression Purif*. 2008;58:162-7.
40. Kasturi L, Chen HG, Shakin-Eshleman SH. Regulation of N-linked core glycosylation: Use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors. *Biochem J*. 1997;323:415-9.

41. Mohl JE, Gerken T, Leung MY. Predicting mucin-type O-glycosylation using enhancement value products from derived protein features. *J Theor Comput Chem*. 2020;19:2040003.
42. Bailey MJ, Siikaaho M, Valkeajarvi A, Penttila ME. Hydrolytic properties of two cellulases of *Trichoderma Reesei* expressed in yeast. *Biotechnol Appl Biochem*. 1993;17:65-76.
43. Beckham GT, Dai Z, Matthews JF, Momany M, Payne CM, Adney WS, et al. Harnessing glycosylation to improve cellulase activity. *Curr Opin Biotechnol*. 2012;23:338-45.
44. Gilkes NR, Henrissat B, Kilburn DG, Miller RC, Warren RAJ. Domains in microbial β -1,4-glycanases—sequence conservation, function, and enzyme families. *Microbiol Rev*. 1991;55:303-15.
45. Agrawal D, Kaur B, Brar KK, Chadha BS. An innovative approach of priming lignocellulosics with lytic polysaccharide mono-oxygenases prior to saccharification with glycosyl hydrolases can economize second generation ethanol process. *Bioresour Technol*. 2020;308:123257.
46. Bernardi AV, Gerolamo LE, de Gouvea PF, Yonamine DK, Pereira LMS, de Oliveira AHC, et al. LPMO AfAA9_B and cellobiohydrolase AfCel6A from *A. fumigatus* boost enzymatic saccharification activity of cellulase cocktail. *Int J Mol Sci*. 2021;22:276.
47. Imperiali B, Rickert KW. Conformational implications of asparagine-linked glycosylation. *P Natl Acad Sci U S A*. 1995;92:97-101.
48. Yanez E, Carmona TA, Tiemblo M, Jimenez A, Fernandez-Lobato M. Expression of the *Schwanniomyces occidentalis* SWA2 amylase in *Saccharomyces cerevisiae*: role of N-glycosylation on activity, stability and secretion. *Biochem J*. 1998;329:65-71.
49. Adney WS, Jeoh T, Beckham GT, Chou Y-C, Baker JO, Michener W, et al. Probing the role of N-linked glycans in the stability and activity of fungal cellobiohydrolases by mutational analysis. *Cellulose*. 2009;16:699-709.
50. Couturier M, Feliu J, Haon M, Navarro D, Lesage-Meessen L, Coutinho PM, et al. A thermostable GH45 endoglucanase from yeast: impact of its atypical multimodularity on activity. *Microb Cell Fact*. 2011;10:103.
51. Kallas AM, Piens K, Denman SE, Henriksson H, Faldt J, Johansson P, et al. Enzymatic properties of native and deglycosylated hybrid aspen (*Populus tremula* x *tremuloides*) xyloglucan endotransglycosylase 16A expressed in *Pichia pastoris*. *Biochem J*. 2005;390:105-13.
52. Joao HC, Scragg IG, Dwek RA. Effects of glycosylation on protein conformation and amide proton exchange rates in RNase B. *FEBS Lett*. 1992;307:343-6.
53. Mer G, Hietter H, Lefèvre J-F. Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated proteinase inhibitor. *Nat Struct Biol*. 1996;3:45-53.
54. Henrissat B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 1991;280:309-16.
55. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490-5.
56. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134:890-2.
57. Beeson WT, Vu VV, Span EA, Phillips CM, Marletta MA. Cellulose degradation by polysaccharide monooxygenases. *Annu Rev Biochem*. 2015;84:923-46.
58. Busk PK, Lange L. Function-based classification of carbohydrate-active enzymes by recognition of short, conserved peptide motifs. *Appl Environ Microbiol*. 2013;79:3380-91.
59. Lenfant N, Hainaut M, Terrapon N, Drula E, Lombard V, Henrissat B. A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res*. 2017;448:166-74.
60. Moses V, Hatherley R, Bishop OT. Bioinformatic characterization of type-specific sequence and structural features in auxiliary activity family 9 proteins. *Biotechnol Biofuels*. 2016;9:1-17.
61. Kotyk A. IUPAC—IUBMB Joint Commission on Biochemical Nomenclature (JCBN) and Nomenclature Committee of IUBMB (NC-IUBMB). *Folia Microbiol*. 1999;44:243-6.
62. Chen X, Zhang X, Long L, Ding S. Comparison of C4-oxidizing and C1/C4-oxidizing AA9 LPMOs in substrate adsorption, H₂O₂-driven activity and synergy with cellulase on celluloses of different crystallinity. *Carbohydr Polym*. 2021;269:118305.
63. Monclaro AV, Petrović DM, Alves GS, Costa MM, Midorikawa GE, Miller RN, et al. Characterization of two family AA9 LPMOs from *Aspergillus tamarii* with distinct activities on xyloglucan reveals structural differences linked to cleavage specificity. *PLOS ONE*. 2020;15:e0235642.
64. Shi Y, Chen K, Long L, Ding S. A highly xyloglucan active lytic polysaccharide monooxygenase *EpLPMO9A* from *Eupenicillium parvum* 4-14 shows boosting effect on hydrolysis of complex lignocellulosic substrates. *Int J Biol Macromol*. 2021;167:202-13.

65. Itsuno S. Enantioselective reduction of ketones. *Organic Reactions*. 2004;52:395-576.
66. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, et al. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20:1479-83.
67. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, et al. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330:219-22.
68. Wang BJ, Wang ZF, Davies GJ, Walton PH, Rovira C. Activation of O₂ and H₂O₂ by lytic polysaccharide monooxygenases. *ACS Catal*. 2020;10:12760-9.
69. Basotra N, Dhiman SS, Agrawal D, Sani RK, Tsang A, Chadha BS. Characterization of a novel lytic polysaccharide monooxygenase from *Malbranchea cinnamomea* exhibiting dual catalytic behavior. *Carbohydr Res*. 2019;478:46-53.
70. Bissaro B, Kommedal E, Røhr ÅK, Eijsink VG. Controlled depolymerization of cellulose by light-driven lytic polysaccharide oxygenases. *Nat Commun*. 2020;11:1-12.
71. Cheng C, Haider J, Liu P, Yang J, Tan Z, Huang T, et al. Engineered LPMO significantly boosting cellulase-catalyzed depolymerization of cellulose. *J Agric Food Chem*. 2020;68:15257-66.
72. Eijsink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnol Biofuels*. 2019;12:58.
73. Forsberg Z, Bissaro B, Gullesen J, Dalhus B, Vaaje-Kolstad G, Eijsink VGH. Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J Biol Chem*. 2018;293:1397-412.
74. Hegnar OA, Petrovic DM, Bissaro B, Alfredsen G, Várnai A, Eijsink VG. pH-dependent relationship between catalytic activity and hydrogen peroxide production shown via characterization of a lytic polysaccharide monooxygenase from *Gloeophyllum trabeum*. *Appl Environ Microbiol*. 2019;85:e02612-e02618.
75. Kont R, Bissaro B, Eijsink VG, Våljamäe P. Kinetic insights into the peroxygenase activity of cellulose-active lytic polysaccharide monooxygenases (LPMOs). *Nat Commun*. 2020;11:1-10.
76. Ogunyewo OA, Randhawa A, Gupta M, Kaladhar VC, Verma PK, Yazdani SS. Synergistic action of a lytic polysaccharide monooxygenase and a cellobiohydrolase from *Penicillium funiculosum* in cellulose saccharification under high-level substrate loading. *Appl Environ Microbiol*. 2020;86:e01769-20.
77. Singh RK, Blossom BM, Russo D, Van Oort B, Croce R, Jensen P, et al. Thermal unfolding and refolding of a lytic polysaccharide monooxygenase from *Thermoascus aurantiacus*. *RSC Adv*. 2019;9:29734-42.
78. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevret D, et al. AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels*. 2019;12.
79. Kuusk S, Bissaro B, Kuusk P, Forsberg Z, Eijsink VGH, Sørle M, et al. Kinetics of H₂O₂-driven degradation of chitin by a bacterial lytic polysaccharide monooxygenase. *J Biol Chem*. 2018;293:523-31.
80. Petrovic DM, Varnai A, Dimarogona M, Mathiesen G, Sandgren M, Westereng B, et al. Comparison of three seemingly similar lytic polysaccharide monooxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J Biol Chem*. 2019;294:15068-81.
81. Bissaro B, Rohr AK, Muller G, Chylenski P, Skaugen M, Forsberg Z, et al. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat Chem Biol*. 2017;13:1123-8.
82. Chen S-C, Wu P-H, Su Y-C, Wen T-N, Wei Y-S, Wang N-C, et al. Biochemical characterization of a novel laccase from the basidiomycete fungus *Cerrena sp.* WR1. *Protein Eng Des Sel*. 2012;25:761-9.
83. Edens WA, Goins TQ, Dooley D, Henson JM. Purification and characterization of a secreted laccase of *Gaeumannomyces graminis var. tritici*. *Appl Environ Microbiol*. 1999;65:3071-4.
84. Hilgers R, Kabel MA, Vincken J-P. Reactivity of *p*-coumaroyl groups in lignin upon laccase and laccase/HBT treatments. *ACS Sustain Chem Eng*. 2020;8:8723-31.
85. Hilgers R, van Dam A, Zuilhof H, Vincken J-P, Kabel MA. Controlling the competition: boosting laccase/HBT-catalyzed cleavage of a β -O-4' linked lignin model. *ACS Catal*. 2020;10:8650-9.
86. Hegnar OA, Petrovic DM, Bissaro B, Alfredsen G, Varnai A, Eijsink VGH. pH-Dependent relationship between catalytic activity and hydrogen peroxide production shown via

- characterization of a lytic polysaccharide monooxygenase from *Gloeophyllum trabeum*. Appl Environ Microbiol. 2019;85.
87. Bissaro B, Varnai A, Rohr AK, Eijsink VGH. Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. Microbiol Mol Biol Rev. 2018;82.
 88. Breslmayr E, Laurent CVFP, Scheiblbrandner S, Jerkovic A, Heyes DJ, Oostenbrink C, et al. Protein conformational change is essential for reductive activation of lytic polysaccharide monooxygenase by cellobiose dehydrogenase. ACS Catal. 2020;10:4842-53.
 89. Kracher D, Forsberg Z, Bissaro B, Gangl S, Preims M, Sygmund C, et al. Polysaccharide oxidation by lytic polysaccharide monooxygenase is enhanced by engineered cellobiose dehydrogenase. FEBS J. 2020;287:897-908.
 90. Tan TC, Kracher D, Gandini R, Sygmund C, Kittl R, Haltrich D, et al. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. Nat Commun. 2015;6:7542.
 91. Kelley RL, Reddy CA. Purification and characterization of glucose-oxidase from ligninolytic cultures of *Phanerochaete chrysosporium*. J Bacteriol. 1986;166:269-74.
 92. Liu S, Oeljeklaus S, Gerhardt B, Tudzynski B. Purification and characterization of glucose oxidase of *Botrytis cinerea*. Physiol Mol Plant Pathol. 1998;53:123-32.
 93. Guillen F, Martinez AT, Martinez MJ. Substrate-specificity and properties of the aryl-alcohol oxidase from the ligninolytic fungus *Pleurotus eryngii*. Eur J Biochem. 1992;209:603-11.
 94. Daniel G, Volc J, Kubatova E. Pyranose oxidase, a major source of H₂O₂ during wood degradation by *Phanerochaete chrysosporium*, *Trametes versicolor*, and *Oudemansiella mucida*. Appl Environ Microbiol. 1994;60:2524-32.
 95. Nishida A, Eriksson KE. Formation, purification, and partial characterization of methanol oxidase, a H₂O₂-producing enzyme in *Phanerochaete chrysosporium*. Biotechnol Appl Biochem. 1987;9:325-38.
 96. Ferreira P, Carro J, Serrano A, Martinez AT. A survey of genes encoding H₂O₂-producing GMC oxidoreductases in 10 *Polyporales* genomes. Mycologia. 2015;107:1105-19.
 97. Kersten PJ. Glyoxal oxidase of *Phanerochaete chrysosporium*-its characterization and activation by lignin peroxidase. Proc Natl Acad Sci U S A. 1990;87:2936-40.
 98. Kersten PJ, Kirk TK. Involvement of a new enzyme, glyoxal oxidase, in extracellular H₂O₂ production by *Phanerochaete chrysosporium*. J Bacteriol. 1987;169:2195-201.
 99. Yin DL, Urresti S, Lafond M, Johnston EM, Derikvand F, Ciano L, et al. Structure-function characterization reveals new catalytic diversity in the galactose oxidase and glyoxal oxidase family. Nat Commun. 2015;6:10197.
 100. Whittaker MM, Kersten PJ, Nakamura N, SandersLoehr J, Schweizer ES, Whittaker JW. Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. J Biol Chem. 1996;271:681-7.
 101. Cregg JM, Madden K, Barringer K, Thill G, Stillman C. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. Mol Cell Biol. 1989;9:1316-23.
 102. Lee MH, Lai WL, Lin SF, Hsu CS, Liaw SH, Tsai YC. Structural characterization of glucooligosaccharide oxidase from *Acremonium strictum*. Appl Environ Microbiol. 2005;71:8881-7.
 103. Momeni MH, Fredslund F, Bissaro B, Raji O, Vuong TV, Meier S, et al. Discovery of fungal oligosaccharide-oxidising flavo-enzymes with previously unknown substrates, redox-activity profiles and interplay with LPMOs. Nat Commun. 2021;12:1-13.
 104. Vuong TV, Vesterinen AH, Foumani M, Juvonen M, Seppala J, Tenkanen M, et al. Xylo- and cello-oligosaccharide oxidation by gluco-oligosaccharide oxidase from *Sarocladium strictum* and variants with reduced substrate inhibition. Biotechnol Biofuels. 2013;6:148.
 105. Perna V, Meyer AS, Holck J, Eltis LD, Eijsink VGH, Agger JW. Laccase-catalyzed oxidation of lignin induces production of H₂O₂. ACS Sustain Chem Eng. 2020;8:831-41.
 106. Mate DM, Alcalde M. Laccase: a multi-purpose biocatalyst at the forefront of biotechnology. Microb Biotechnol. 2017;10:1457-67.
 107. Eggert C, Temp U, Eriksson K-EL. Laccase is essential for lignin degradation by the white-rot fungus *Pycnoporus cinnabarinus*. FEBS Lett. 1997;407:89-92.
 108. Jones SM, Solomon EI. Electron transfer and reaction mechanism of laccases. Cell Mol Life Sci. 2015;72:869-83.
 109. Pollegioni L, Tonin F, Rosini E. Lignin-degrading enzymes. FEBS J. 2015;282:1190-213.
 110. Hendriks ATWM, Zeeman G. Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour Technol. 2009;100:10-8.
 111. Kumar AK, Sharma S. Recent updates on different methods of pretreatment of lignocellulosic feedstocks: a review. Bioresour Bioprocess. 2017;4:7.

112. Kumar P, Barrett DM, Delwiche MJ, Stroeve P. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Ind Eng Chem Res.* 2009;48:3713-29.
113. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, et al. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour Technol.* 2005;96:673-86.
114. Berka RM, Grigoriev IV, Otilar R, Salamov A, Grimwood J, Reid I, et al. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol.* 2011;29:922-7.
115. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb Cell Fact.* 2016;15:1-13.
116. Hinz SW, Pouvreau L, Joosten R, Bartels J, Jonathan MC, Wery J, et al. Hemicellulase production in *Chrysosporium lucknowense* C1. *J Cereal Sci.* 2009;50:318-23.
117. Stricker AR, Mach RL, De Graaff LH. Regulation of transcription of cellulases-and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*). *Appl Microbiol Biotechnol.* 2008;78:211-20.
118. Merino ST, Cherry J. Progress and challenges in enzyme development for biomass utilization. *Adv Biochem Eng Biot.* 2007;108:95-120.
119. Cannella D, Chia-wen CH, Felby C, Jørgensen H. Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol Biofuels.* 2012;5:1-10.
120. Cannella D, Jorgensen H. Do new cellulolytic enzyme preparations affect the industrial strategies for high solids lignocellulosic ethanol production? *Biotechnol Bioeng.* 2014;111:59-68.
121. Du J, Song WX, Zhang X, Zhao J, Liu GD, Qu YB. Differential reinforcement of enzymatic hydrolysis by adding chemicals and accessory proteins to high solid loading substrates with different pretreatments. *Bioprocess Biosystems Eng.* 2018;41:1153-63.
122. Sun FF, Hong J, Hu J, Saddler JN, Fang X, Zhang Z, et al. Accessory enzymes influence cellulase hydrolysis of the model substrate and the realistic lignocellulosic biomass. *Enzyme Microb Technol.* 2015;79-80:42-8.
123. Berlin A, Maximenko V, Gilkes N, Saddler J. Optimization of enzyme complexes for lignocellulose hydrolysis. *Biotechnol Bioeng.* 2007;97:287-96.
124. Calderaro F, Keser M, Akeroyd M, Bevers LE, Eijsink VGH, Varnai A, et al. Characterization of an AA9 LPMO from *Thielavia australiensis*, TausLPMO9B, under industrially relevant lignocellulose saccharification conditions. *Biotechnol Biofuels.* 2020;13:195.
125. Costa TH, Eijsink VGH, Horn SJ. The use of lytic polysaccharide monooxygenases in anaerobic digestion of lignocellulosic materials. *Biotechnol Biofuels.* 2019;12:1-14.
126. Eibinger M, Ganner T, Bubner P, Rosker S, Kracher D, Haltrich D, et al. Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J Biol Chem.* 2014;289:35929-38.
127. Muller G, Varnai A, Johansen KS, Eijsink VG, Horn SJ. Harnessing the potential of LPMO-containing cellulase cocktails poses new demands on processing conditions. *Biotechnol Biofuels.* 2015;8:187.
128. Pierce BC, Agger JW, Zhang ZH, Wichmann J, Meyer AS. A comparative study on the activity of fungal lytic polysaccharide monooxygenases for the depolymerization of cellulose in soybean spent flakes. *Carbohydr Res.* 2017;449:85-94.
129. Sanhueza C, Carvajal G, Soto-Aguilar J, Lienqueo ME, Salazar O. The effect of a lytic polysaccharide monooxygenase and a xylanase from *Gloeophyllum trabeum* on the enzymatic hydrolysis of lignocellulosic residues using a commercial cellulase. *Enzyme Microb Technol.* 2018;113:75-82.
130. Tokin R, Ipsen JØ, Westh P, Johansen KS. The synergy between LPMOs and cellulases in enzymatic saccharification of cellulose is both enzyme- and substrate-dependent. *Biotechnol Lett.* 2020;42:1975-84.
131. Velasco J, Pellegrini VDA, Sepulchro AGV, Kadowaki MAS, Santo MCE, Polikarpov I, et al. Comparative analysis of two recombinant LPMOs from *Aspergillus fumigatus* and their effects on sugarcane bagasse saccharification. *Enzyme Microb Technol.* 2021;144:109746.
132. Aden A, Foust T. Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol. *Cellulose.* 2009;16:535-45.
133. Joelsson E, Erdei B, Galbe M, Wallberg O. Techno-economic evaluation of integrated first- and second-generation ethanol production from grain and straw. *Biotechnol Biofuels.* 2016;9:535-45.

134. Chylenski P, Bissaro B, Sorlie M, Rohr AK, Varnai A, Horn SJ, et al. Lytic polysaccharide monooxygenases in enzymatic processing of lignocellulosic biomass. *ACS Catal.* 2019;9:4970-91.
135. Corrêa TL, dos Santos LV, Pereira GA. AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol.* 2016;100:9-16.
136. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol.* 2015;33:747-61.
137. Johansen KS. Discovery and industrial applications of lytic polysaccharide mono-oxygenases. *Biochem Soc Trans.* 2016;44:143-9.
138. Singhanian RR, Dixit P, Patel AK, Giri BS, Kuo CH, Chen CW, et al. Role and significance of lytic polysaccharide monooxygenases (LPMOs) in lignocellulose deconstruction. *Bioresour Technol.* 2021;335:125261.
139. Keller MB, Badino SF, Rojel N, Sorensen TH, Kari J, McBrayer B, et al. A comparative biochemical investigation of the impeding effect of C1-oxidizing LPMOs on cellobiohydrolases. *J Biol Chem.* 2021;296:100504.
140. Tokin R, Ipsen JØ, Westh P, Johansen KS. The synergy between LPMOs and cellulases in enzymatic saccharification of cellulose is both enzyme-and substrate-dependent. *Biotechnol Lett.* 2020;42:1975-84.
141. Kato Y, Habu N, Yamaguchi J, Kobayashi Y, Shibata I, Isogai A, et al. Biodegradation of β -(1 \rightarrow 4)-linked polyglucuronic acid (cellouronic acid). *Cellulose.* 2002;9:75-81.
142. Xu F, Ding H, Tejirian A. Detrimental effect of cellulose oxidation on cellulose hydrolysis by cellulase. *Enzyme Microb Technol.* 2009;45:203-9.
143. Yang J, Tu MB, Xia CL, Keller B, Huang Y, Sun FF. Effect of fenton pretreatment on C1 and C6 oxidation of cellulose and its enzymatic hydrolyzability. *ACS Sustain Chem Eng.* 2019;7:7071-9.
144. Eibinger M, Sattolkow J, Ganner T, Plank H, Nidetzky B. Single-molecule study of oxidative enzymatic deconstruction of cellulose. *Nat Commun.* 2017;8:894.
145. Keller MB, Badino SF, Blossom BM, McBrayer B, Borch K, Westh P. Promoting and impeding effects of lytic polysaccharide monooxygenases on glycoside hydrolase activity. *ACS Sustain Chem Eng.* 2020;8:14117-26.
146. Busse-Wicher M, Grantham NJ, Lyczakowski JJ, Nikolovski N, Dupree P. Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochem Soc Trans.* 2016;44:74-8.
147. Penttilä PA, Varnai A, Pere J, Tammelin T, Salmen L, Siika-aho M, et al. Xylan as limiting factor in enzymatic hydrolysis of nanocellulose. *Bioresour Technol.* 2013;129:135-41.
148. Yu L, Lyczakowski JJ, Pereira CS, Kotake T, Yu X, Li A, et al. The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage. *Plant Physiol.* 2018;178:1011-26.
149. Chylenski P, Forsberg Z, Stahlberg J, Varnai A, Lersch M, Bengtsson O, et al. Development of minimal enzyme cocktails for hydrolysis of sulfite-pulped lignocellulosic biomass. *J Biotechnol.* 2017;246:16-23.
150. Hu JG, Arantes V, Saddler JN. The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnol Biofuels.* 2011;4:36.
151. Varnai A, Huikko L, Pere J, Siika-aho M, Viikari L. Synergistic action of xylanase and mannanase improves the total hydrolysis of softwood. *Bioresour Technol.* 2011;102:9096-104.
152. Gupta VK, Kubicek CP, Berrin JG, Wilson DW, Couturier M, Berlin A, et al. Fungal enzymes for bio-products from sustainable and waste biomass. *Trends Biochem Sci.* 2016;41:633-45.
153. Han XS, Bi R, Khatri V, Oguzlu H, Takada M, Jiang JG, et al. Use of endoglucanase and accessory enzymes to facilitate mechanical pulp nanofibrillation. *ACS Sustain Chem Eng.* 2021;9:1406-13.
154. Karnaouri A, Jalvo B, Moritz P, Matsakas L, Rova U, Hofft O, et al. Lytic polysaccharide monooxygenase-assisted preparation of oxidized-cellulose nanocrystals with a high carboxyl content from the tunic of marine invertebrate *Ciona intestinalis*. *ACS Sustain Chem Eng.* 2020;8:18400-12.
155. Li ZQ, Zhang Y, Anankanbil S, Guo Z. Applications of nanocellulosic products in food: Manufacturing processes, structural features and multifaceted functionalities. *Trends Food Sci Technol.* 2021;113:277-300.

156. Moreau C, Tapin-Lingua S, Grisel S, Gimbert I, Le Gall S, Meyer V, et al. Lytic polysaccharide monoxygenases (LPMOs) facilitate cellulose nanofibrils production. *Biotechnol Biofuels*. 2019;12.
157. Mudedla SK, Vuorte M, Veijola E, Marjamaa K, Koivula A, Linder MB, et al. Effect of oxidation on cellulose and water structure: a molecular dynamics simulation study. *Cellulose*. 2021;28:3917-33.
158. Muraleedharan MN, Karnaouri A, Piatkova M, Ruiz-Caldas MX, Matsakas L, Liu B, et al. Isolation and modification of nano-scale cellulose from organosolv-treated birch through the synergistic activity of LPMO and endoglucanases. *Int J Biol Macromol*. 2021;183:101-9.
159. Perzon A, Blossom BM, Felby C, Jeoh T, Hitomi A, Ulvskov P, et al. Cellulose nanofibrils as assay substrates for cellulases and lytic polysaccharide monoxygenases. *ACS Appl Nano Mater*. 2020;3:6729-36.
160. Rossi BR, Pellegrini VOA, Cortez AA, Chiromito EMS, Carvalho AJF, Pinto LO, et al. Cellulose nanofibers production using a set of recombinant enzymes. *Carbohydr Polym*. 2021;256:117510.
161. Villares A, Moreau C, Bennati-Granier C, Garajova S, Foucat L, Falourd X, et al. Lytic polysaccharide monoxygenases disrupt the cellulose fibers structure. *Sci Rep*. 2017;7:40262.
162. Moreau C, Villares A, Capron I, Cathala B. Tuning supramolecular interactions of cellulose nanocrystals to design innovative functional materials. *Ind Crops Prod*. 2016;93:96-107.
163. Lavoine N, Desloges I, Dufresne A, Bras J. Microfibrillated cellulose—its barrier properties and applications in cellulosic materials: A review. *Carbohydr Polym*. 2012;90:735-64.
164. Azizi Samir MAS, Alloin F, Dufresne A. Review of recent research into cellulosic whiskers, their properties and their application in nanocomposite field. *Biomacromolecules*. 2005;6:612-26.
165. Kalashnikova I, Bizot H, Cathala B, Capron I. New Pickering emulsions stabilized by bacterial cellulose nanocrystals. *Langmuir*. 2011;27:7471-9.
166. Li F, Sun XJ, Yu W, Shi CC, Zhang XY, Yu HB, et al. Enhanced konjac glucomannan hydrolysis by lytic polysaccharide monoxygenases and generating prebiotic oligosaccharides. *Carbohydr Polym*. 2021;253:117241.
167. Graubner S, Zverlov V, Schwarz W, Kornberger P, Andreeßen B, Herlet J, et al. Technische Universitaet Muenchen, assignee. Method for preparing xyloglucan-oligosaccharides. Germany 2019.
168. Chen H, Jiang X, Li S, Qin W, Huang Z, Luo Y, et al. Possible beneficial effects of xyloglucan from its degradation by gut microbiota. *Trends Food Sci Technol*. 2020;97:65-75.
169. Mishra A, Malhotra AV. Tamarind xyloglucan: a polysaccharide with versatile application potential. *J Mater Chem*. 2009;19:8528-36.
170. Sabbadin F, Urresti S, Henrissat B, Avrova AO, Welsh LR, Lindley PJ, et al. Secreted pectin monoxygenases drive plant infection by pathogenic oomycetes. *Science*. 2021;373:774-9.
171. Li W, Zhang K, Yang H. Pectin alleviates high fat (lard) diet-induced nonalcoholic fatty liver disease in mice: possible role of short-chain fatty acids and gut microbiota regulated by pectin. *J Agric Food Chem*. 2018;66:8015-25.
172. Beukema M, Faas MM, de Vos P. The effects of different dietary fiber pectin structures on the gastrointestinal immune barrier: impact via gut microbiota and direct effects on immune cells. *Exp Mol Med*. 2020;52:1364-76.
173. Elshahed MS, Miron A, Aprotosoiaie AC, Farag MA. Pectin in diet: Interactions with the human microbiome, role in gut homeostasis, and nutrient-drug interactions. *Carbohydr Polym*. 2020;117388.
174. Larsen N, Bussolo de Souza C, Krych L, Barbosa Cahú T, Wiese M, Kot W, et al. Potential of pectins to beneficially modulate the gut microbiota depends on their structural properties. *Front Microbiol*. 2019;10:223.
175. Maxwell EG, Belshaw NJ, Waldron KW, Morris VJ. Pectin—an emerging new bioactive food polysaccharide. *Trends Food Sci Technol*. 2012;24:64-73.
176. Fadel A, Mahmoud AM, Ashworth JJ, Li W, Ng YL, Plunkett A. Health-related effects and improving extractability of cereal arabinoxylans. *Int J Biol Macromol*. 2018;109:819-31.
177. van Soest Pv, Robertson J, Lewis B. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J Dairy Sci*. 1991;74:3583-97.
178. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes*. 2012;3:289-306.
179. Slominski BA. Recent advances in research on enzymes for poultry diets. *Poul Sci*. 2011;90:2013-23.

180. Kouzounis D, Hageman JA, Soares N, Michiels J, Schols HA. Impact of xylanase and glucanase on oligosaccharide formation, carbohydrate fermentation patterns, and nutrient utilization in the gastrointestinal tract of broilers. *Animals*. 2021;11:1285.
181. Ojha B, Singh PK, Shrivastava N. Enzymes in the animal feed industry. *Enzymes in food biotechnology*: Elsevier; 2019. p. 93-109.
182. Yu R, Scott EV. Oligosaccharide aldonic acids and their topical use. Patent US20080090772A1. 2004.

Summary

Summary

Lytic polysaccharide monooxygenases (LPMOs) have been shown to play an important role in enzymatic conversion of plant cell wall polysaccharides during lignocellulose-based biorefinery, which is considered as a sustainable process for the production of biofuels, biochemicals and biomaterials. In the last decade, advances have been achieved in revealing the three-dimensional structure and catalytic mechanism of LPMOs, however, (bio-)chemical characterization and understanding cleavage profiles is still limited. Assigning such characteristics will contribute to the understanding of biological functions and roles of LPMOs in nature, and to the further implementation of LPMOs for industrial applications. Therefore, this thesis attempted to provide deeper insights into the mode-of-action, regioselectivity, substrate cleavage profiles and specificity of Auxiliary Activity (AA) family 9 LPMOs from the fungus *Myceliophthora thermophila* C1 (*MtLPMOs*), and from the fungus *Neurospora crassa* (*NcLPMOs*).

In **Chapter 1**, the background and objectives of this PhD project are described. The first part of the general introduction starts with the composition and architecture of plant cell walls. Subsequently, structure and composition of the three major plant cell wall components cellulose, hemicellulose and lignin are presented, together with their interactions within the plant cell wall. In addition, plant cell wall polysaccharide degrading enzymes classified in the Carbohydrate-Active enZymes (CAZy) database are described, with focus on cellulases, xylanases and xyloglucanases. In the second part of the general introduction, an overview of published LPMO research is provided till early 2018 (e.g., till the start of this PhD research). This overview includes the discovery of LPMOs, current classification of LPMOs, structural features of LPMO catalytic domains and the often attached carbohydrate-binding modules. Furthermore, the current understanding of O₂- and H₂O₂-dependent catalytic pathways, regioselectivity and substrate specificity is reviewed. We also set out (partially) characterized *MtLPMOs* and *NcLPMOs* till the beginning of this PhD project. Lastly, an overview of analytical approaches and assays to monitor LPMO activity and catalytic performance is provided.

In **Chapter 2**, we investigated product profiles of four fungal AA9 LPMOs (*MtLPMO9B*, *MtLPMO9H*, *MtLPMO9I* and *NcLPMO9M*) during their oxidative cleavage of three types of cellulose: bacterial cellulose (BC), Avicel® PH-101 (AVI), and regenerated amorphous cellulose (RAC). First, we showed that removal of a CBM1 did not alter the substrate specificity of *MtLPMO9B*, but decreased the efficiency of degradation of all three types of cellulose. Subsequently, we quantified the oxidized ends in both soluble and insoluble fractions, and profiled the oxidized cello-oligosaccharide patterns generated by *MtLPMO9B* (with and without a CBM1) from time-course incubations. It was found that *MtLPMO9B* released mainly oxidized cellobiose from BC, while a more evenly in degree of polymerization (DP) distributed mixture of oxidized cello-oligosaccharide (DP2–8) was found for AVI and RAC. Our findings suggest that cellulose specificity and product profiles of AA9 LPMOs are modulated by the type of cellulose rather than being LPMO type dependent.

In **Chapter 3**, mass spectrometric fragmentation behaviors and patterns of (NaBD₄-reduced) non-, C4- and C1-oxidized cello-oligosaccharides were studied by using hydrophilic interaction chromatography and negative ion mode collision induced

dissociation–mass spectrometry (HILIC-ESI-CID-MS/MS²). We revealed that each type of cello-oligosaccharide studied showed a distinct “signature mass spectrometric fragmentation pattern”: i) non-oxidized showed predominantly C- and A-type cleavages; ii) C4-oxidized underwent B-/Y and X-cleavage close to the oxidized non-reducing end, while C-/Z- and A-fragmentation were predominated closer to the reducing end; iii) C1-oxidized showed extensively A-cleavage, and iv) reduced showed predominant both B-/Y- and C-/Z-, close to the non-reducing end.

In **Chapter 4**, we reported for the first time, identification of a series of LPMO products: C4/C6-double oxidized cello-oligosaccharides aided by NaBD₄ reduction and HILIC-ESI-CID-MS/MS². We found that C4/C6-double oxidized cello-oligosaccharides were generated by C4- and C1/C4-oxidizing LPMOs, but not by C1-oxidizing ones. By using isotope labelled H₂¹⁸O, we confirmed that the C6-gem-diol structure resulted from monooxygenation, though oxidation to a C6-aldehyde, followed by hydration to the C6-gem-diol, could not be excluded.

In **Chapter 5**, five segments around the LPMO active site (Seg1–Seg5) involved in substrate recognition were defined based on a structure-based alignment. Next, we investigated the changes in activity induced by shortening the Seg2 or removing the CBM1 from *NcLPMO9C*. The absence of CBM1 reduced the binding affinity and activity of *NcLPMO9C*, but did not alter its regioselectivity. The linker was found important for the thermal stability of *NcLPMO9C* and the CBM1 improved binding to RAC. The wild-type *NcLPMO9C* exhibited the highest activity and strongest substrate binding. Shortening of Seg2 greatly reduced the activity on RAC and completely abolished the activity on xyloglucan.

In **Chapter 6**, we characterized xyloglucan degradation profiles and elucidated structures of oxidized xylogluco-oligosaccharide formed by *NcLPMO9C* and *NcLPMO9M*. Our detailed structural characterization showed that oxidative cleavage occurred next to glucosyl units substituted with galactosyl and xylosyl substituted units. This chapter further shows our structure-based phylogenetic analysis of AA9 LPMOs. By assigning substrate cleavage specificities of *NcLPMO9C*, *NcLPMO9M* and other published xyloglucan-(in)active LPMOs to this structure-based phylogeny, we indicated a correlation between the configuration of active site segments and xyloglucan specificity of AA9 LPMOs. In brief, LPMOs with +Seg1–Seg2 were found to be tolerant to xyloglucan substitution, while LPMOs with –Seg1+Seg2 were found to be intolerant to xyloglucan substitution. LPMOs with –Seg1–Seg2 or –Seg1+Seg2+Seg3 were not active towards xyloglucan. All three types could oxidatively cleave cellulose.

In **Chapter 7**, a new *MtLPMO9F* and a partially characterized *MtLPMO9H* were studied for their substrate specificity, especially for their xyloglucan specificity. Aided by NaBD₄ reduction and HILIC-ESI-CID-MS/MS² analysis, we found that both *MtLPMOs* released predominately C4-oxidized, and C4/C6-double oxidized xylogluco-oligosaccharides. Further characterization showed that *MtLPMO9F*, having –Seg1+Seg2, generated a “substitution-intolerant” xyloglucan cleavage profile, while for *MtLPMO9H* (+Seg1–Seg2) a “substitution-tolerant” profile was found. The here characterized xyloglucan specificity and substitution (in)tolerance of *MtLPMO9F* and *MtLPMO9H* was as predicted

Summary

according to our previously published phylogenetic grouping of AA9 LPMOs based on structural active site segment configurations (in Chapter 6).

In **Chapter 8**, we studied a new AA16 member from *M. thermophila* C1 (*MtAA16A*), which was suggested to act as LPMO. We found that *MtAA16A* did not show any oxidative cleavage of carbohydrate substrates including cellulose, hemicellulose, chitin and starch. However, *MtAA16A* largely boosted all three AA9 *MtLPMOs* in their oxidative cellulose degradation. The same boosting effect was observed for another AA16 enzyme from *Aspergillus nidulans* (*AnAA16A*). However, oxidative cellulose degradation by three *NcLPMOs* was not boosted by the two AA16 oxidoreductases. The boosting effect was found to relate to H_2O_2 production by the AA16s and suggested to relate to a fine-tuned H_2O_2 delivery helped by specific protein-protein interaction.

In the final **Chapter 9**, outcomes and findings of the previous chapters are discussed in context of recent literature (2018-2021). The diversity of AA9 LPMOs with focus on *MtLPMOs* and *NcLPMOs* is presented in two overview tables and discussed with recent published data. In addition, recommendations for further improvements of the developed chromatographic approaches in this thesis are provided. The intriguing dual ability of AA9 LPMOs and AA16 oxidoreductases to produce and/or consume H_2O_2 depending is further highlighted. Connected to this dual ability, in Chapter 9, the accuracy and outcomes of an in literature widely used H_2O_2 -based activity assay to monitor LPMO activity are questioned. Furthermore, new results of a syringol- H_2O_2 driven LPMO reaction are presented, and the potential of using lignin derived compounds to drive LPMOs in degrading lignocellulose is discussed. Lastly, we present current applications and future perspectives of LPMOs in lignocellulosic biomass degradation and valorization.

Acknowledgements

Acknowledgements

The moment that I finished my PhD thesis, I realized that I am close to the end of this wonderful journey. It took me some time to write down this “chapter”, simply because I cannot find enough words to express my appreciation and gratitude to all people who supported me and contributed to this PhD research in the past four years. Without you, I could not have made it.

First of all, I would like to thank my supervisor, Mirjam. Everything started with the MSc Enzymology course in 2016. I found so many interests in your lectures, which let me choose to join your research group without hesitation. After being your student during my master thesis and internship, now I am so proud of being your first PhD student with you as one of my two promotor! You gave me great guidance and supervision, allowing me to gradually become an independent researcher. I am very grateful for your understanding of my difficulty in directly expressing feelings, ideas and opinions in the first period of my research. You were always patient and willing to help me overcome these issues. I have learned a lot from you on how to plan and organize my work, and how to be efficient and productive. I also very much appreciate that you gave me a lot of freedom during my research, and let me try out (sometimes crazy) ideas. You helped me to shape these ideas in more feasible experiments, and I have enjoyed the collaborations you initiated inside and outside the Netherlands. In the beginning of my PhD research, the outcomes of my experiments were not all immediately promising, but you were confident that I could finish my thesis in time. Your trust and support always encouraged me, both in work and in life. Thank you for our collaboration and for the great discussions we had about the still extremely challenging LPMO research.

I would also like to thank my other supervisor and promotor, Willem, the true enzyme expert. We met the first time in a PhD defense party and I was very impressed by your knowledge and experience on oxidases. I am so happy that you are involved in my project, providing us with different ways of thinking and designing the research. Every time when Mirjam and me were not sure about some ideas, we concluded as “let’s ask Willem”. You helped me with focusing on the main research questions and sorting out how to extract an understandable story out of a massive amount of data. I am very grateful for your critical thinking, suggestions and feedback in my PhD research. Especially in the writing period, your advice and comments always further improved the structure, logic and readability of manuscripts. You and Mirjam are amazing supervisors, I enjoyed very much working with you.

Matthias, the “grandfather” of our “lignocellulose family”, I would like to specifically acknowledge you for your involvement in the first part of my research. You shared your knowledge and hands-on experience with me as much as possible, to let me quickly get into the LPMO research. I also much appreciated all your help, for example, you drove with me to the company to collect trays of crude enzyme broths, you helped me with filtering these crude enzyme broths (which took about a week!), you were the “connector” to fix the joint to prevent an explosion :P, etc., too many things. Thanks for all your support and contributions!

I would like to thank my two paranymphs, Dimitris and Katharina, for all their help and support during these years, and now for the preparation of my defense. I really enjoyed and missed our dinner time at each of our places! Dimitris, my buddy, we almost started our PhD research at the same time and finally we made it! You were always there to listen, to care, to help, to motivate and to contaminate me. We have talked and laughed so much at FCH, at our places and at restaurants. I will miss the time that we spent together forever! Katharina, you were so sweet and kind to offer your help and support, and always stood with me and thought from my side. I am so happy that we became such good friends and colleagues!

I would like to thank my thesis committee members, prof. Dr Richard van Kranenburg, prof. Dr Katja Salomon Johansen, prof. Dr Manfred Wuhrer, and prof. Dr Dietmar Haltrich for their time to read and evaluate my PhD thesis.

My great gratitude and thanks to the entire FCH group. It is so great to work here. I not only enjoyed the scientific atmosphere and research environment, but also all the lab-out activities, PhD trips, coffee/lunch breaks and "borrel" sessions.

Jean-Paul, the big boss, you were and are passionate about all researches. I especially enjoyed our discussions about the LPMO-xyloglucan study. I appreciated it that you regularly talked with me to see if everything was going well. Thanks for all your efforts to make FCH better and like a big family!

Henk, the encyclopedia of carbohydrates, thanks for your advice and help on carbohydrate analysis and characterization. I really appreciated your support, patience and kindness whenever I came with questions!

My special thanks go to Jolanda. You were always patient, friendly and helpful with any questions and issues I had. I cannot image how I could focus on the research without your help on all non-scientific matters!

Many thanks to all technicians at FCH. Your expertise ensures all analytical instruments to perform in good conditions. Especially Margaret, Edwin, Rene and Mark, thanks for helping me with protein purification, liquid chromatographic and mass spectrometric analysis. I could not have learned all these analytics without you!

Gijs, thanks for being one of my best friends and colleagues, "father" of the "lignocellulose family", my MSc thesis daily supervisor and office mate. We had good times together, for example in our discussions about lignin, LPMOs, and the bet of LPMO AA class number. I have really enjoyed our regular "pizza nights", "peanut-LPMO breaks" (all with Matthias :D), coffee breaks outside (with Dimitris), funny jokes (I guess not all people can get your jokes :P), plant "research" and slogans ("Orbi et orbi", "It's a trap!"). I have learned a lot from you, such as how to critically perform a research, to work hard, to think out of the box, etc. Whenever I came to you with questions and puzzles, you were, and still are, willing to help me with solutions. I really appreciate all your support and help.

Acknowledgements

Maloe, Melanie, Pim, Romy and Kai, thanks for choosing my topics as part of your MSc and BSc studies. I am very grateful that I got the opportunity to supervise you, and I was proud to be your daily supervisor. It was my pleasure to work with all of you. You really contributed to my PhD research, and each of you is also coauthored on one of my papers! In addition, it was great to see Pim and Romy entering as PhD students at FCH, which means the "family tree" is branching further.

I would like to thank my lab mates in X0225 (previously L1.05), Margaret, Matthias, Hugo, Bianca, Suzanna, Dimitris, Katharina, Pim, Dazhi (刘大志), Krishna for sharing experience, helping each other, organizing lab cleaning-up, arranging apparatus and creating so nice atmosphere. I also would like to thank my office mates, Junfeng, Gijs, Christophe, Dazhi (刘大志), Laura, Anne, Bram, Dounia for nice conversations, fruitful discussions and sharing peanuts and cookies, which I missed a lot during the Corona time!

Eva and Madelon, thanks for helping me with solving HPAEC and HPSEC issues in the early time. Sylvia, Roelant and Wouter, thanks for your hands-on experience and help with UPLC-MS. These instruments did not always do what I wanted. :P Sylvia, many thanks to you for helping me with printing the final thesis book. 感谢刘老师与我分享你在科研和日常工作中的宝贵经验, 让我受益匪浅。

I would like to extend my thanks to FCHinese colleagues, Zhibin (刘志彬), Junfeng (谭俊峰), Dazhi (刘大志), Yuxi (邓羽西), Fangjie (谷方婕), Caifang (文彩芳), Weiwei (孙维维), Jianli (王剑力). I really enjoy the time with you at FCH. I miss our lunch and coffee breaks, and also nice food we had together. 非常享受和怀念与你们一起在食品化学组工作的日子, 在荷兰遇见你们是我的幸运。

I also would like to thank all the people outside WUR who were involved and collaborated in my project. Christophe, Stefan and Roland, thanks for the great collaboration we had in the past four years, which led to five coauthored papers. I hope we can continue the collaboration between WUR and BOKU in the future. Xinxin (李新新), Adiphol, Ronald and Miia, thanks for involving me in the fungal glycoside hydrolase research, which allowed me to obtain further knowledge in this research-field. Marco, David, Sanchari, Kristian and Leila, thanks for your knowledge and contribution for understanding the intriguing AA16 enzymes. I have learned so much from you, and your diverse expertise really broaden my views on biochemistry and biocatalysis research. Susana and Javier, thanks for your contribution on physicochemical characterization of cellulose and biomaterials; Damao (王大毛), we have met the first time in the 2nd LPMO Symposium in Marseille, and surprisingly found out that we studied in the same university in China. Thanks for all your support and advice. I really look forward to seeing you again!

Apart from the work, there is also a life next to it. I am very grateful for so many friends around me, to refresh my mind, to release my stress and to encourage me.

Xu (程旭), thanks for everything you did for me. Whenever I had puzzles, questions, happiness and unhappiness, you were always there to approach. I very much

appreciate your endless help and support. 旭哥，在荷兰求学的日子里，你是我最想感谢的人。无论我有任何问题，困惑，失落，还是开心的事，成功的喜悦，总是想第一个与你分享。你给了我和白溪无微不至的关怀和耐心的指导，让我们能够享受生活，努力工作，鼓足干劲和充满勇气迎接一切困难。你无私地分享你对工作和人生的经验和体会，然我们少走弯路，对我们这个年纪的人来说是无比宝贵的财富。你是我和白溪成长道路上最重要的亲人，朋友和人生导师。我想借此处来深深的表达对你的感谢！

Yongran (季永然), table tennis let us meet and know each other, and be good friends. However, our friendship means much more than table tennis only. Thanks for letting me understand how to study abroad, how to integrate into local life, how to balance between work and life. 老季，在我读博道路上你扮演了太重要的角色。在硕士学习刚开始的时候，是你介绍我去 Shot 打球，参加俱乐部联赛和其它各种当地的活动，让我快速适应了荷兰的生活，可以专注于学业。之后，在我有读博的打算后，你又是全心全意的帮助我，与我分享读博过程中的经验。我们不仅仅是在 Shot 打球，喝酒和聊天，还有与思哥一起我们的聚餐，下馆子，参加啤酒节，等等。虽然因为后来我和白溪搬出了瓦村，再加上疫情的限制，我们的交流活动少了，但从未感到疏远。我感到非常开心和幸运在荷兰认识你还有思哥，感谢你们的陪伴。

刘大志，我的“伴郎”（虽然因为疫情没当成哈哈）。在学校我们一起喝咖啡，买饮料，甚至在你第一次回国前一起买酒在组里喝。我们交换对周边事情的看法，你的真诚和洒脱让我感觉非常的放松自在，舒缓了工作的压力。愿你在剩下的博士学习过程中一切顺利，和嘉宁的生活越来越美好。严志纯 & 罗露，虽然我们认识的比较晚，但很快我们两家便如此亲近。和旭哥一起，我们经常边喝酒边交流读博过程中的经历与感悟，还有每次都期待的你和白溪的辩论哈哈。后来又见证了六一的出生，我和白溪作为你们女儿的干爸干妈，衷心的希望你们一切都好，六一健康成长。王子申 & 董蓓，每次和你们一起聚餐和出行都让我非常放松和开心，感谢你们帮我和白溪照结婚照，搬家等等所做的一切，希望你们也顺利的博士毕业！邓磊，本科都是西农的我们，竟然如此巧合的再博士期间又相聚瓦大。怀念我们一起聚餐和喝咖啡的时光。祝你和宇思生活美满，博士最后阶段一切顺利！除此之外，还有更多的朋友，王杨超，赵弘博，夏靓，刘明泽，马俊楠等；一起打篮球的伙伴们（虽然我去的频率越来越低），韩明钊，俞果，纪磊，王正聪，董力友，胡博涵，刘翔宇等；“中国乒协驻荷兰办事处”的朋友们，旭哥，老季，王强，丁宽，牛牧童，黄一根，徐志姚，侯宣，常进，张宇，乔金文，刘润玮等，感谢你们的陪伴。

I am proud of being a member of the Shot (Tafeltennisvereniging Shot Wageningen) family. Aad, Anatolj, André, Eloy, Eric, Frank, Frans, Hans, Helma, Henk, Johan (Frederiks), Johan (van de Pol), Lorenzo, Marco, Ric, Roland, Sebastian, Sietse, Wim and many other friends, thanks for all the happy and fun moments we had together.

I would like to give many thanks to my mom, my dad and my grandma. 爸爸妈妈，有些话对我来说一直羞于表达，但是借此机会，我想感谢你们长久以来对我的理解和包容，在我成长过程对我的无尽的爱。因为我知道你们会相信我，支持我，这让我可以自信和大胆的去拼搏，去追求我想我爱。无论身在何处，都会收到你们的支持和肯定，让我时刻感受到家的温暖，你们是我坚强的后盾，谢谢你们！还有我的外婆（郭香亭），您从小照顾我长大，在大学前从未离开过您，直到后来不在您身边，才知道对您是多么的想念！愿您一直健康长寿，等我回去！我还想感谢白爸爸王妈妈，虽然我们真正一起待过的时间很少，但是我总能收到你们的关心和支持，你们的真情实意让我也感受到多一份来自父母的温暖与幸福。谢谢你们！我还想感谢大姨，大姨夫，二姨，二姨夫，表姐和表妹，谢谢你们对我一直以来的关心和支持，让我们的大家庭如此温暖。

Acknowledgements

Finally, I would like to express my greatest thanks to my beloved Xi. 白溪，我的妻子，伴侣，家人，最好的朋友，你在我这里有太多太多的重要的角色，无论过去，现在还是未来，你无可替代。感谢你出现在我的生活里，感谢你为我所做的一切。从我们在大学的相识相知，毕业后的异国，再到荷兰追求学业，你和我一路走来，无论我们遇到什么样的困难和挫折，你都不离不弃。在过去的九年里，你改变了我太多，鼓励我努力奋斗，追求梦想；提醒我专心专注，不为身边的嘈杂喧闹所影响；最重要的是，让我知道了如何去爱。愿执子之手，与子偕老！

Peicheng Sun

孙培成

About the author

Curriculum vitae

Peicheng Sun was born on April 4, 1991, in Shandong, China. From 2010, he majored in Food Science and Technology during his Bachelor study in Northwest Agricultural and Forestry University in Yangling, China. After obtaining his Bachelor degree in 2014, he worked in Shandong Highland Co., Ltd. for a year. In 2015, he started his master study of Food Technology with the specialization Ingredient Functionality in Wageningen University & Research, The Netherlands. He performed his master thesis entitled "Isolation of lignin from wheat straw and characterization by Py-GC/MS" at the Laboratory of Food Chemistry, and an industrial internship entitled "Purification and characterization of a starch-active lytic polysaccharide monooxygenase" at DuPont Industrial Biosciences, The Netherlands. He received his Master degree in 2017 and directly started as a PhD candidate at the Laboratory of Food Chemistry under the supervision of Prof. Dr Willem van Berkel and Dr Mirjam Kabel. The results of his PhD research are described in this thesis. Peicheng continues to work at the Laboratory of Food Chemistry as a postdoctoral researcher.



✉ peicheng.sun@outlook.com



List of publications

Peicheng Sun, Sanchari Banerjee, Marco A.S. Kadowaki, Romy J. Veersma, Silvia Magri, Roelant Hilgers, Christophe V.F.P. Laurent, Roland Ludwig, David Cannella, Leila Lo Leggio, Willem J.H. van Berkel, and Mirjam A. Kabel. Copper-dependent AA16 oxidoreductases boost cellulose active AA9 LPMOs from *Myceliophthora thermophila* C1. *Manuscript in preparation*.

Peicheng Sun, Melanie de Munnik, Willem J.H. van Berkel, and Mirjam A. Kabel. Extending the diversity of *Myceliophthora thermophila* lytic polysaccharide monooxygenases: two different xyloglucan cleavage profiles. *Carbohydrate Polymers*. Accepted for publication.

Peicheng Sun, Christophe V.F.P. Laurent, Vincent J.P. Boerkamp, Gijs van Erven, Willem J.H. van Berkel, Roland Ludwig, and Mirjam A. Kabel. Regioselective C4 and C6 double oxidation of cellulose by lytic polysaccharide monooxygenases. *ChemSusChem*. 2022;15:e202102203.

***Peicheng Sun**, *Xinxin Li, Adiphol Dilokpimol, Bernard Henrissat, Ronald P. de Vries, Mirjam A. Kabel, and Miia R. Mäkelä. Fungal glycoside hydrolase family 44 xyloglucanases are restricted to the phylum Basidiomycota and show a distinct xyloglucan cleavage pattern. *iScience*. 2022;25:103666.

Peicheng Sun, Susana V. Valenzuela, Pimvisuth Chunkrua, Francisco. I. Javier Pastor, Christophe V.F.P. Laurent, Roland Ludwig, Willem J.H. van Berkel, and Mirjam A. Kabel. Oxidized product profiles of AA9 lytic polysaccharide monooxygenases depend on the type of cellulose. *ACS Sustainable Chemistry & Engineering*. 2021;9:14124-14133.

Peicheng Sun, Matthias Frommhagen, Maloe Kleine Haar, Gijs van Erven, Edwin J. Bakx, Willem J.H. van Berkel, and Mirjam A. Kabel. Mass spectrometric fragmentation patterns discriminate C1- and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. *Carbohydrate Polymers*. 2020;234:115917.

***Peicheng Sun**, *Christophe V.F.P. Laurent, Stefan Scheiblbrandner, Matthias Frommhagen, Dimitrios Kouzounis, Mark G. Sanders, Willem J.H. van Berkel, Roland Ludwig, and Mirjam A. Kabel. Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. *Biotechnology for Biofuels*. 2020;13:95.

*Christophe V.F.P. Laurent, ***Peicheng Sun**, Stefan Scheiblbrandner, Florian Csarman, Pietro Cannazza, Matthias Frommhagen, Willem J.H. van Berkel, Chris Oostenbrink, Mirjam A. Kabel, and Roland Ludwig. Influence of lytic polysaccharide monooxygenase active site segments on activity and affinity. *International Journal of Molecular Sciences*. 2019;20:6219.

About the author

Dimitrios Kouzounis, **Peicheng Sun**, Edwin J. Bakx, Henk A. Schols, and Mirjam A. Kabel. Strategy to identify reduced arabinoxylo-oligosaccharides by HILIC-MSⁿ. *Carbohydrate Polymers*. Accepted for publication.

Gijs van Erven, Jianli Wang, **Peicheng Sun**, Pieter de Waard, Jacinta van der Putten, Guus E. Frissen, Richard J.A. Gosselink, Grigory Zinovyev, Antje Potthast, Willem J.H. van Berkel, and Mirjam A. Kabel. Structural motifs of wheat straw lignin differ in susceptibility to degradation by the white-rot fungus *Ceriporiopsis subvermispota*. *ACS Sustainable Chemistry & Engineering*. 2019;7:20032-20042.

Mirjam A. Kabel, Matthias Frommhagen, **Peicheng Sun**, Henk A. Schols. Modification of plant carbohydrates using fungal enzymes. *Encyclopedia of Mycology*: Elsevier; 2021. p.370-384.

Damao Wang, Denis Kao, Jing Li, **Peicheng Sun**, Qijun Meng, Anisha Vyas, Pi-Hui Liang, Yane-Shih Wang, and Yves S.Y. Hsieh. A novel integrated process in cellulose degradation: hematite α -Fe₂O₃-nanoparticle-mediated photocatalysis by lytic polysaccharide monooxygenase and subsequent Fenton reaction. *Submitted*.

*These authors contributed equally to this work.

Overview of completed training activities

Discipline specific activities

Conferences

2 nd LPMO Symposium ^a	INRA, Marseille, 2018
XV Plant Cell Wall Meeting ^a	CU, Cambridge, 2019
CHEMistry As INnovating Science (CHAINS) ^a	NWO, Veldhoven, 2019
Netherlands Biotechnology Conference ^b	NBV, Ede, 2019
Dutch Biocatalysis Symposium ^b	NBV, online, 2020
Carbohydrate Competence Center Symposium	CCC, online, 2021

Courses

Summer Course Glycosciences ^a	VLAG, Wageningen, 2018
Applied Biocatalysis ^a	VLAG, Wageningen, 2019
Ultrafast Sample Treatment for Proteomics	BIOSCOPE, online, 2020
Advanced Organic Chemistry	ORC, online, 2021
Applied Enzymology ^{a,b}	VLAG, Wageningen, 2022

General courses

VLAG PhD week	VLAG, Baarlo, 2018
Scientific Writing	WGS, Wageningen, 2018
Posters & Pitching	WGS, Wageningen, 2018
Searching and Organising Literature	WGS, Wageningen, 2019
Career Perspectives	WGS, Wageningen, 2020

Optional

PhD research proposal	FCH, Wageningen, 2017
PhD trip ^{a,b}	FCH, Italy and Austria, 2018
PhD trip ^{a,b}	FCH, Groningen, 2021
BSc and MSc thesis student supervisions, presentations and colloquia	FCH, Wageningen, 2017-2020
PhD presentations ^b	FCH, Wageningen, 2017-2021

^aPoster presentation. ^bOral presentation.

Abbreviations: INRA, Institut National de la Recherche Agronomique; CU, Cambridge University; NWO, Netherlands Organization for Scientific Research; NBV, Dutch Biotechnology Association; CCC, Carbohydrate Competence Center; VLAG, Graduate school for Advanced Studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences; BIOSCOPE, BIOSCOPE Research Group; ORC, Laboratory of Organic Chemistry; WGS, Wageningen Graduate Schools; FCH, Laboratory of Food Chemistry.

The work described in this thesis was performed at the Laboratory of Food Chemistry of Wageningen University & Research, The Netherlands. This project was sponsored by the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Financial support from Wageningen University and the Graduate School VLAG for printing this thesis is gratefully acknowledged.

Cover design by Xiangming Wang (王湘铭), Xi Bai and Peicheng Sun

This thesis was printed by ProefschriftMaken, The Netherlands

Edition: 200 copies

Peicheng Sun, 2022