Structural and functional characterization of *Nc*LPMO9C, a broad-specificity lytic polysaccharide monooxygenase

Strukturell og funksjonell karakterisering av *Nc*LPMO9C, en lytisk polysakkarid monooksygenase med bred spesifisitet

Philosophiae Doctor (PhD) Thesis

Trine Øye Isaksen

Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food Science

Ås (2017)

Thesis number 2017:93
ISSN 1894-6402
ISBN 978-82-575-1480-8

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The present work was carried out during the period of 2012-2017 in the Protein Engineering and Proteomics (PEP) group, at the Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences (NMBU). The project was funded through a grant from the Research Council of Norway to NMBU and received support from Borregaard, in Sarpsborg, Norway. Travel grants to conferences and courses were generously awarded by BioStruct, a national Norwegian graduate school. I am indebted to many people for their efforts and guidance throughout these years:

First of all, I would like to thank my supervisor Prof. **Vincent Eijsink** for accepting me as a PhD student. Thank you for your endless enthusiasm and encouragement and for always seeing opportunities when all I see are limitations. Most of all, thank you for believing in me and for the amount of work you have put into helping me with achieving today's results. I would also like to express my gratefulness to my co-supervisors Dr. **Jane W. Agger**, Prof. **Svein Jarle Horn** and Dr. **Geir Mathiesen**. Thank you for your helpfulness, patience and all your guidance.

I would also like to thank Dr. **Bjørge Westereng**, Dr. **Anikó Várnai**, Prof. **Morten Sørlie**, Dr. **Åsmund Røhr**, Dr. **Dejan Petrovic** and especially Dr. **Zarah Forsberg** for all your help. Thank you for sharing so much of your time and knowledge with me. Thanks to all other members of the PEP-group, in particular **Ellen** and **Anne-Cath**, and to the groups of Prof. **Dietmar Haltrich** and Dr. **Roland Ludwig** for hosting me and making my stay in Vienna so nice. Thanks also to other Norwegian and foreign collaborators who have made contributions to the papers that are included in this thesis. A big thank you to **Live** for being not only a good colleague, but also such a good friend. Thank you for every laugh, all your help and for always keeping your door open. Coffee breaks with you and **Kasia** have often been a day-saver.

To **all** my friends, in particular **Eirin**, **Anders**, **Idunn and Tor**: Thank you for being so good in distracting me from work! I am truly blessed with having friends like you. As always, I am thankful for having such a great family. Thank you for always being there for me and for all your love. **Håvard**, thank you for your endless love and support. You are one of a kind! My beloved son, **Mikkel**, you have changed all my perspectives. Thank you for showing me what life is really about.

Trine

SUMMARY

Lignocellulose represents a major source of renewable organic matter and is of interest as feedstock for the biorefining industry, not only as a promising strategy for replacing fossil fuels in the transportation sector but also for production of higher value chemicals and animal feed. However, the complex and recalcitrant nature of lignocellulosic biomass puts major challenges to biorefining of this feedstock. Enzymatic saccharification of the polysaccharides in the feedstock, especially cellulose, is considered a crucial and challenging step in biorefining and this step is one of the limiting factors in the transition towards a sustainable bioeconomy. To overcome the recalictrant nature of lignocellulosic biomass, microorganisms have developed intricate enzyme systems, including lytic polysaccharide monooxygenases (LPMOs). LPMOs are copper-dependent enzymes that, in the presence of an electron donor, cleave the glycoside bonds of various polysaccharides using an oxidative mechanism. When working in synergy with cellulases, which are hydrolytic enzymes, LPMOs boost the enzymatic degradation of cellulose. Since their discovery in 2010, the novel catalytic abilities of LPMOs and their great potential in biomass degradation have attracted researchers in academia as well as industry. However, despite considerable research efforts, our knowledge of the action of LPMOs is still limited.

This thesis describes studies of the substrate-degrading properties of LPMOs, which were assessed through a detailed study of an LPMO from the saprotrophic fungus *Neurospora crassa*. *Nc*LPMO9C is a two-domain protein with an N-terminal LPMO domain connected through a linker to a C-terminal carbohydrate-binding domain (CBM1). Prior to the studies presented here, LPMO activity had only been shown for crystalline substrates, namely chitin and cellulose. Paper I describes, for the first time, oxidative action of an LPMO on soluble cello-oligosaccharides. Upon cleavage of cellopentaose, *Nc*LPMO9C generates two major products, native cellotriose and a cellobiose with an oxidation in the non-reducing end. By the use of NMR and MS-analyses it was demonstrated that oxidative cleavage of β–(1-4) glycosidic linkages by *Nc*LPMO9C involves the introduction of oxygen at the C4-position.

Paper II elaborates on the activity of *Nc*LPMO9C towards soluble substrates. Using glycan microarray screening, it was discovered that *Nc*LPMO9C is active on various hemicellulosic substrates and this capability was further studied by studying the degradation of common hemicelluloses such as tamarind xyloglucan. It was shown that *Nc*LPMO9C cleaves β-(1,4)-glycosidic linkages in most hemicelluloses with a β-glucan backbone (i.e. various β-glucans,

glucomannan and xyloglucan) and can tolerate various backbone substitutions. *Nc*LPMO9C seems particularly active on xyloglucan and it was shown that the enzyme is capable of releasing oxidized fragments from xyloglucans occurring in *Arabidopsis thaliana* and *Solanum lycopersicum* (tomato plant).

Paper III describes the crystal structure of the catalytic domain of *Nc*LPMO9C and additional biochemical characterization of this enzyme. Although no structure with a bound substrate was obtained, insights into the binding properties of *Nc*LPMO9C were derived from isothermal titration calorimetry and studies of enzyme reaction rates. The substrate binding affinity and the activity of *Nc*LPMO9C were affected by removal of the C-terminal CBM1-domain for both phosphoric acid-swollen cellulose (PASC) and xyloglucan. While binding constants for these substrates were in the low micromolar range, the binding affinity for cellohexaose was very low, with a K_d close to 1 mM. Structural comparison of LPMO9s with known activities and structures revealed that the oxidative regioselectivity of LPMO action (C1-specific, C4-specific, or mixed specificity) correlates with structural features near the copper-site, as had previously been observed for bacterial LPMOs. EPR analyses revealed, for the first time, that substrate binding induces changes in the copper binding site.

In conclusion, this study presents novel insights into the substrate-degrading properties of LPMOs and the structural basis thereof. The work presented here contributes to current research efforts aimed at understanding the catalytic capabilities of LPMOs and at harnessing these capabilities in industrial biorefining.

SAMMENDRAG

Lignocellulose representerer en viktig kilde til fornybart organisk materiale og er av interesse som råmateriale for bioraffineringsindustrien, ikke bare som en lovende strategi for erstatning av fossile brensler i transportsektoren, men også for produksjon av kjemikalier og dyrefôr. Den komplekse og gjenstridige sammensetningen av lignocellulose gir store utfordringer i bioraffineringen av dette råmaterialet. Enzymatisk omdanning av polysakkaridene i lignocellulosen, spesielt cellulose, regnes som et avgjørende og utfordrende trinn i bioraffineringsprosessen. Dette trinnet representerer også en av hovedbegrensningene i overgangen til en bærekraftig bioøkonomi. For å overvinne den tungt nedbrytbare lignocellulosen har mikroorganismer utviklet intrikate enzymsystemer, som blant annet inneholder lytiske polysakkarid monooksygenaser (LPMOer). LPMOer er kobberavhengige enzymer som, i nærvær av en elektrondonor, bryter glykosidbindingene i forskjellige polysakkarider ved bruk av en oksidativ mekanisme. I synergi med cellulaser, som er hydrolytiske enzymer, øker LPMOer hastigheten til den enzymatiske nedbrytningen av cellulose. Siden oppdagelsen av LPMOer i 2010 har den katalytiske funksjonen til disse enzymene, samt deres potensiale innen nedbryting av biomasse, tiltrukket forskere fra både akademia og industri. Til tross for betydelig forskningsinnsats er kunnskapen om LPMOenes funksjon fortsatt begrenset.

Denne avhandlingen beskriver studier av de substratnedbrytende egenskapene til LPMOer, anskaffet gjennom en detaljert studie av en LPMO fra en sopp, *Neurospora crassa*. *Nc*LPMO9C er et to-domene protein med et N-terminalt LPMO-domene koblet, via en linker, til et C-terminalt karbohydratbindende domene (CBM1). I forkant av studiene presentert her, hadde LPMO-aktivitet kun blitt vist for krystallinske substrater av kitin og cellulose. Artikkel I beskriver for første gang oksidativ aktivitet av en LPMO på løselige cello-oligosakkarider. Ved spaltning av cellopentaose genererer *Nc*LPMO9C to hovedprodukter, en nativ cellotriose og en oksidert cellobiose. Ved bruk av både NMR- og MS-analyser ble det påvist at *Nc*LPMO9C tilfører et oksygenatom til C4-karbonet på den ikke-reduserende enden av cellobiosemolekylet når enzymet spalter β–(1,4) glykosidbindingen.

Artikkel II utdyper aktiviteten til *Nc*LPMO9C mot løselige substrater. Ved bruk av glukan-mikroarray screening ble det oppdaget at *Nc*LPMO9C er aktiv på forskjellige hemicellulose substrater. Denne egenskapen ble studert videre ved å undersøke nedbrytningen av vanlig forekommende hemicellulosepolymerer. Det ble vist at *Nc*LPMO9C kløyver β-(1,4)-

glykosidbindinger i de fleste hemicelluloser med en hovedkjede av β-glukan (som forskjellige β-glukaner, glukomannan og xyloglukan) og kan tolerere forskjellige substitusjoner på hovedkjeden. *Nc*LPMO9C virker spesielt aktiv på xyloglucan, og det ble vist at enzymet er i stand til å frigjøre oksyderte fragmenter fra xyloglukane isolert fra både *Arabidopsis thaliana* og *Solanum lycopersicum* (tomat plante).

Artikkel III beskriver krystallstrukturen for det katalytiske domenet til *Nc*LPMO9C og ytterligere biokjemisk karakterisering av dette enzymet. Selv om ingen struktur med et bundet substrat ble oppnådd, ble innsikt i bindingsegenskapene til *Nc*LPMO9C avledet fra isotermisk titreringskalorimetri og studier av enzymreaksjonshastigheter. Substratbindingsaffiniteten og aktiviteten til *Nc*LPMO9C ble påvirket ved fjerning av det C-terminale CBM1-domenet for både fosforsyre-svellet cellulose (PASC) og xyloglukan. Mens bindingskonstanter for disse substratene var i det lave mikromolare området, var bindingsaffiniteten for cellohexaose meget lav, med en K_d nær 1 mM. Strukturell sammenligning av LPMO9er med kjente aktiviteter og strukturer viste at LPMOens oksidative regioselektivitet (C1-spesifikk, C4-spesifikk eller blandet spesifisitet) korrelerer med strukturelle egenskaper nær kobberbindingssetet, som tidligere også har blitt observert for bakterielle LPMOer. EPR-analyser viste for første gang at substratbinding forårsaker endringer i hvordan enzymet binder kobberatomet.

Til slutt kan det oppsummeres at studiene i avhandlingen presenterer ny innsikt i de substratnedbrytende egenskapene til LPMOer, samt de strukturelle grunnlag derav. Arbeidet bidrar også til dagens forskning som tar sikte på å forstå de katalytiske egenskapene til LPMOer og utnyttelsen av disse egenskapene i industrielle biorefiningsprosesser.

ABBREVIATIONS

AA	Auxiliary Activity
CAZyme	Carbohydrate Active enZyme
CBM	Carbohydrate-Binding Module
CBP	Chitin-Binding Protein
CBH	CelloBioHydrolase
CDH	Cellobiose DeHydrogenase
CE	Carbohydrate Esterase
DP	Degree of Polymerization
EC	Enzyme Commission
EG	EndoGlucanase
EPR	Electron Paramagnetic Resonance
ESI	ElectroSpray Ionization
GH	Glycoside Hydrolase
HPAEC	High-Performance Anion-Exchange Chromatography
ITC	Isothermal Titration Calorimetry
LPMO	Lytic Polysaccharide MonoOxygenase
NMR	Nuclear Magnetic Resonance
MALDI-ToF	Matrix-Assisted Laser Desorption/ Ionization Time off Flight
MS	Mass Spectrometry
PASC	Phosphoric Aid-Swollen Cellulose
PDB	Protein Data Bank
PGC	Porous Graphitized Carbon
PUL	Polysaccharide Utilization Locus
$XG14^{OH}$	$XXXGXXXG^{OH}$ (X indicate a glucose substituted with xylose)

LIST OF PAPERS

Paper I

A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides

Trine Isaksen, Bjørge Westereng, Finn L. Aachmann, Jane W. Agger, Daniel Kracher, Roman Kittl, Roland Ludwig, Dietmar Haltrich, Vincent G. H. Eijsink and Svein J. Horn (2014) *The Journal of Biological Chemistry*, 289:2632–2642.

Paper II

Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation

Jane W. Agger, Trine Isaksen*, Anikó Várnai*, Silvia Vidal-Melgosa, William G.T. Willats, Roland Ludwig, Svein J. Horn, Vincent G. H. Eijsink and Bjørge Westereng (2014) *Proceedings of the National Academy of Sciences of the United States of America*, 111:6287–6292.

*T.I. and A.V. contributed equally to this paper

Paper III

Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity

Anna S. Borisova*, Trine Isaksen*, Maria Dimarogona, Abhishek A. Kognole, Geir Mathiesen, Anikó Várnai, Åsmund K. Røhr, Christina M. Payne, Morten Sørlie, Mats Sandgren and Vincent G. H. Eijsink (2015) *The Journal of Biological Chemistry*, 290: 22955–22969.

*Shared 1st authorship

Other publications by the author

Fungal lytic polysaccharide monooxygenases bind starch and β-cyclodextrin similarly to amylolytic hydrolases

Laura Nekiunaite, Trine Isaksen, Gustav Vaaje-Kolstad, Maher Abou Hachem (2016) *FEBS Letters*, 590:2737-2747.

1 INTRODUCTION

1.1 Towards a sustainable bioeconomy

Since the industrial revolution in the 18th century, and especially since the invention of the combustion engine in the mid-19th century, the world's demand for energy, fuels and chemicals has increased dramatically. Although the production of biomass-derived alcohols and oils was an option at the time, the increased availability of inexpensive petroleum during the late 19th century more or less replaced all renewable alternatives (Van Wyk, 2011). Today, the world economy is based on the petroleum industry. Petroleum and other hydrocarbons in the form of gas and coal have been formed during millions of years, as a result of microbial degradation of biomass, combined with exposure to high temperature and pressure. Increased demands and the depletion of hydrocarbon reservoirs now set the focus on alternative sources. In addition, the negative effects on the global environment from the constant release of carbon dioxide from petroleum-based processes are considered one of the key causes of current global climate changes (Vanholme et al., 2013). Consequently, the demand for sustainable and renewable sources to replace petroleum is increasing. Although the demand for energy can be met by the conversion of solar, wind, geothermal, hydroelectric or wave energy into electricity, none of these alternative energy sources is able to replace the chemicals and fuels that are currently generated from petroleum (Vanholme et al., 2013). Moreover these alternative energy sources are not easily implemented in all of today's transportation infrastructure (Alonso et al., 2010), despite major progress in the development of electricity-driven vehicles. In this regard, liquid biofuels derived from renewable biomasses are of interest, since such fuels are compatible with existing fuel distribution systems and combustion engines (Rubin, 2008).

Conventional, or first-generation, biofuels are made from feedstocks that are easy to process, including food crops such as corn, sugarcane, wheat and beets (Himmel et al., 2007). The sugar, starch or vegetable oil obtained from the crops is converted into ethanol or biodiesel via yeast fermentation and transesterification, respectively. Although this process is economically competitive with the petroleum industry, in a situation with a growing population and an increasing demand for food, first generation biofuels are not a sustainable choice for the future (Vanholme et al., 2013).

As a promising alternative to conventional biofuels, conversion of non-edible lignocellulosic biomass to fuels and chemicals has received great attention in the past decade. Lignocellulose

is a generic term for the dry matter of plant cell wall, which are co-polymeric structures consisting primarily of cellulose, hemicellulose and lignin, and represents the largest source of renewable organic matter on Earth. Lignocellulosic biomasses such as woody crops, agricultural residues or waste plant material from food crops are of high value as feedstocks for the production of second-generation biofuels (Himmel et al., 2007). Notably, the terms third and fourth generation biofuel are sometimes used to indicate fuels generated from algael biomass or photobiological solar fuels and electrofuels, respectively (Aro, 2016).

Production of biofuels from lignocellulose is currently unable to compete economically with either petroleum refineries or conventional biofuels, and one of the bottle-necks is the incomplete utilization of the biomass feedstock (Alonso et al., 2017). One limiting factor lies in the recalcitrant nature of the lignocellulose, making it almost impermeable to solvents or enzymes. To improve processability, e.g. by enzymes, pretreatment of lignocellulose is needed and common pretreatment processes involve heat and/or acidic conditions that are expensive, slow and relatively inefficient (Rubin, 2008). Full utilization of the biomass requires conversion of all polysaccharides by enzymes into their component 5- and 6-carbon sugars. The lignocellulose-derived sugars are then fermented by microorganisms to useful products, such as ethanol, chemicals, single cell protein, or a combination of these. Notably, this process leaves a lignin-rich rest fraction, the valorization of which is not easy, but could potentially contribute to overall process economy.

Although in Nature microbial strategies for degrading lignocellulose are diverse and effective, and despite decades of research, our current understanding of these processes and the enzymes involved is still limited. Great efforts still need to be made to overcome the recalcitrant nature of lignocellulosic biomass and to develop economically sustainable strategies in both pretreatment and enzymatic hydrolysis (Brethauer & Studer, 2015).

1.2 The plant cell wall

Some plants reach heights of more than 100 meters and to do so, they rely on robust cell walls that are capable of withstanding large physical forces (Scheller & Ulvskov, 2010). The plant cell wall is an intricate network of polysaccharides (cellulose, pectin and hemicelluloses) and polyphenols (lignin) interspersed with proteins and various inorganic compounds (Figure 1A). The composition of the cell wall depends on the plant species, age and tissue type. In addition to supplying mechanical strength and support, the cell wall protects the cell against infectious pathogens and functions as energy storage. A plant cell wall is generally described as containing four layers: the middle lamella, the primary cell wall, the secondary cell wall, and warty layers (Figure 1B) (Sjöström, 1993). These layers differ in chemical composition and structure.

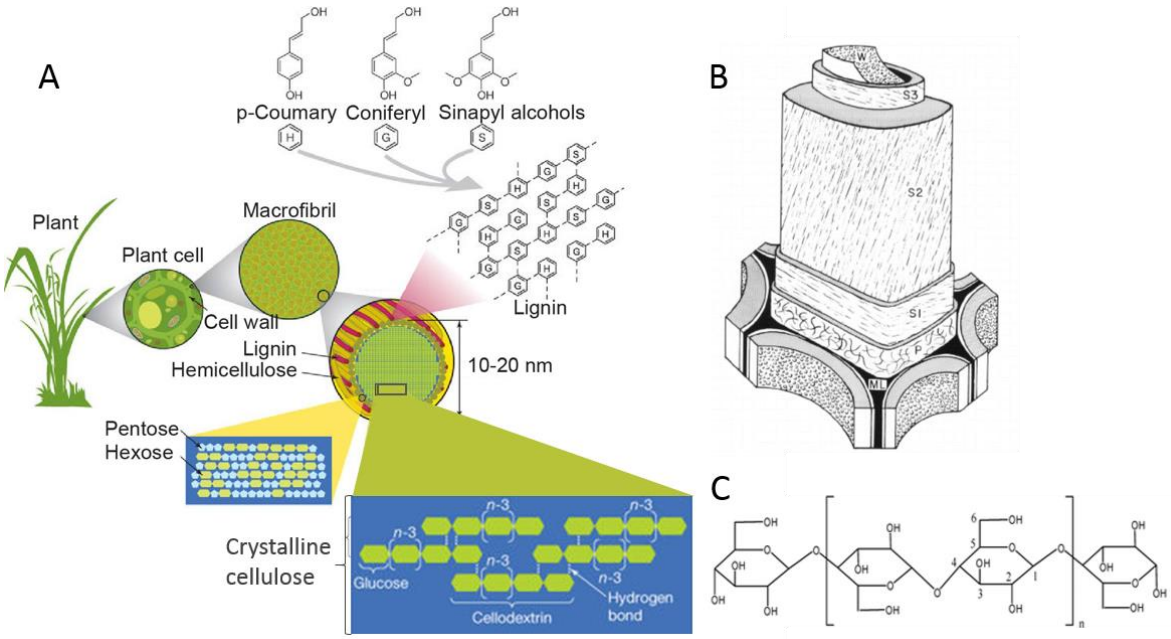

Figure 1: Plant cell wall organization and the structure of lignocellulosic material. Panel (**A**) shows the organization of the three main components of the plant cell wall: cellulose, hemicellulose and lignin. Panel (**B**) visualizes the different layers of plant cell walls: Middle lamella (ML), Primary wall (P), the three secondary walls (S1, S2 and S3) and the Warty layer (W). Panel (**C**) shows the structure of cellulose, a linear chain of β-(1,4)-linked glucose residues, with the non-reducing end of the glucan chain shown to the left and the reducing end to the right. The repeating unit of cellulose is cellobiose as indicated by (n). The pictures were taken from A: (Rubin, 2008), B:(Côtes, 1967) and C: (Shen et al., 2011).

The middle lamella is a pectin layer that forms a unified and continuous layer, cementing adjacent cells together. As the plant is growing, the pectin layer is lignified to reinforce the cell structure (Sjöström, 1993). The primary cell wall is a thin, flexible layer (0.1-1µm) formed while the cell is growing (Cosgrove, 2005). It consists of an irregular network of crystalline cellulose in a matrix of hemicellulose, pectin and protein, and is completely

embedded in lignin. When the cell is fully expanded, a secondary cell wall is deposited onto the primary wall in cells that need extra structural reinforcement (Somerville et al., 2004). This secondary wall is highly organized with parallel fibers of crystalline cellulose interlinked with lignin and hemicellulose. Based on the cellulose orientation with respect to the fiber axis, three distinct sublayers of the secondary cell wall can be visualized in the electron microscope (designated S1, S2, S3 in Figure 1B) (Fujita & Harada, 2001). Layer S2 constitutes the main portion of the cell wall (its thickness varies between 1 and 5 μm), and contains the majority of lignin in the cell. The inner surface of the cell walls of conifers and some hardwoods (Figure 2, see legend) are covered with an amorphous membrane called the warty layer (Sjöström, 1993). The warts are mainly compost of lignin and hemicelluloses and are believed to arise form excess of wall materials deposited onto the S3 layer (Fujita & Harada, 2001).

Although the exact chemical composition of the cell walls varies between plant species, cell types and even between different cell wall layers, typically lignocellulosic feedstocks are made up of 40-50% cellulose, 23-35 % hemicellulose and 15-20% lignin (Alonso et al., 2010).

1.2.1 Cellulose

The main component of lignocellulose, and the most abundant polysaccharide on Earth, is cellulose. Cellulose is primarily found in plant cell walls, with a global production estimated to be 120-140 billion tons annually (Wang et al., 2016), but is also produced by other organisms, including bacteria, algae and some marine animals. The primary structure of cellulose is a linear polymer of β-(1,4)-linked glucans; the consecutive monomers are rotated 180 degrees relative to each other, meaning that the repeating unit is cellobiose (Figure 1C).

In plants, cellulose is synthesized at the plasma membrane by hexameric complexes of transmembrane cellulose synthase, called rosettes. Cellulose synthase is a glycosyltransferase that uses uridine diphosphate (UDP)-glucose (i.e. activated glucose) to synthesize individual glucan chains that further aggregate into cellulose microfibrils as they are expelled from the cell (Morgan et al., 2012), stabilized by inter- and intramolecular hydrogen bonds and van der Waals interactions. It is presumed that the number of glucan chains in each microfibril corresponds to the number of cellulose synthases in each rosette. However, due to the small size of each fibril and the tendency of fibrils to tightly associate with each other in the cell wall, determination of the number of chains per fibril has so far been somewhat inconclusive (Nishiyama, 2009). Suggested models vary between 12 and 36 chains per microfibril (McFarlane et al., 2014) made from a multiple of six cellulose synthases per rosette (Jarvis,

2013). Recent data combining wide-angle X-ray scattering and solid-state nuclear magnetic resonance (NMR) support a model of 18 glucan chains per microfibril (Newman et al., 2013). During microfibril synthesis, the close proximity of the cellulose synthases in one rosette is essential for the correct bundling of the glucan chains. Transmission electron microscopy and computational analyzes predict that one rosette contains 18 cellulose synthases, which correlates well with a model of 18 glucan chains per microfibril (Nixon et al., 2016).

In nature, crystalline cellulose is found in two forms (polymorphs), Iα and Iβ (Wang et al., 2016). While Iα is the dominating cellulose allomorph in bacteria and algae, Iβ, which is more thermodynamically stable, is the major type of cellulose found in plants (O'Sullivan, 1997). Both allomorphs have parallel glucan chains, but differ in terms of inter-chain hydrogen bonding. Recently, solid-state NMR studies of cellulose from the primary wall of several plant species hava suggested the presence of several conformations that differ from Iα and Iβ and that are specific to the position of the cellulose chain in the microfibril network (Wang et al., 2016). Five other polymorphs of cellulose (II, III_I, III_{II}, IV_I, and IV_{II}) can be obtained from pretreatment of native cellulose. Cellulose II can be formed either through dissolution followed by recrystallization, or alkali treatment of cellulose I (Beckham et al., 2011). The glucan chains in cellulose II are arranged in an antiparallel fashion and this is the most stable cellulose polymorph. Treatment of either cellulose I or II with liquid ammonia yields type III_I or III_{II}, respectively, which can further be thermally treated to obtain cellulose type IV_I or IV_{II} (O'Sullivan, 1997). Because cellulose is the main load-bearing polymer, the length, angle and crystallinity of the microfibrils are important determinants of the physical characteristics of the plant cell wall. Details of the morphology of cellulose microfibrils remain unresolved, including detailed knowledge of the different hydrophilic and hydrophobic faces that interact with other cell wall components such as hemicelluloses and lignin (Pereira et al., 2017).

1.2.2 Hemicellulose

Hemicellulose is a generic term for a diverse group of amorphous non-cellulose plant polysaccharides, mainly based on a β-(1,4)-linked backbone. This includes xyloglucans, xylans and mannans, which are all present in the cell walls of terrestrial plants, and mixed-link beta-glucans, which are restricted to Poales, a taxonomic order including cereal grasses (Figure 2), and a few other groups (Scheller & Ulvskov, 2010). Hemicelluloses tend to bind to cellulose and play key roles in energy storage and in securing both the flexibility and the strength of plant cell walls. Hemicelluloses may be modified and/or branched (Figure 3),

which adds variation and structural and functional complexity. Several hypotheses exist to describe the interactions between cellulose and xylan (Busse-Wicher et al., 2014) or xyloglucan (Vincken et al., 1995). However, the exact organization and function of hemicelluloses in the plant cell wall is intricate to study and is to date poorly understood (Busse-Wicher et al., 2016).

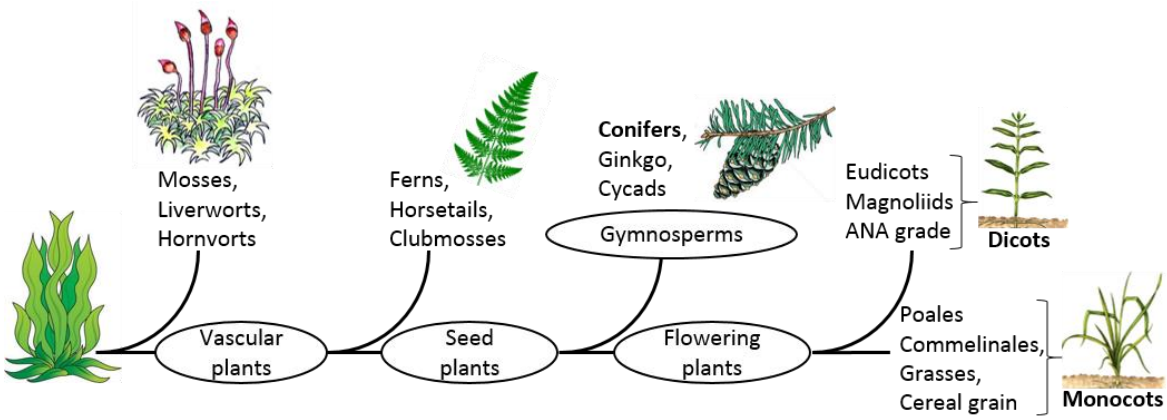

Figure 2: Plant phylogeny. In plant phylogeny, seeding plants are classified as either gymnosperms (e.g. softwood conifers like pine or spruce) or flowering plants. Traditionally, flowering plants are further divided into two groups based on if they have one or two embryonic leaves in their seeds, referred to as monocots (e.g. Poales, including grasses, and hardwood Eudicots like oak, maple, and walnut) and dicots, respectively.

1.2.2.1 Xyloglucan

Xyloglucan is mainly found in the primary cell walls of all land plant species analyzed (Peña et al., 2008; Pauly & Keegstra, 2016). The backbone of β-(1,4)-linked glucose is highly substituted with α-xylosyl residues that are attached to the O6-position in a regular pattern (Hayashi, 1989). The xylosyl residues, and in some cases the backbone glucosyl residues, can be further substituted by arabinose, galactose, galacturonic acid and/or fucose (Pauly & Keegstra, 2016). In addition, acetylation frequently occurs at terminal galactose or arabinose sugars or at unsubstituted backbone glucoses. These substitutions contribute to functional variation of xyloglucan (Scheller & Ulvskov, 2010). To date, 24 unique side-chain compositions have been identified in xyloglucan (see Figures 1 and 2 in (Pauly & Keegstra, 2016), each of which has its own one-letter designation based on an unambiguous nomenclature described in (Fry et al., 1993) and utilized in Paper II of this thesis. Although the xylose substitution patterns differ among plant species, the most common xyloglucans are of the XXXG- or XXGG- type (Vincken et al., 1997).

In the primary cell walls of dicots, the xyloglucan content can be up to 25% of the total polysaccharides, whereas in grasses, the abundance is much lower (2-10%) (Scheller & Ulvskov, 2010). Xyloglucan interacts with cellulose by hydrogen bonding and because of its high affinity for cellulose, it is believed that the main biological function of xyloglycans is to (non-covalently) crosslink cellulose microfibrils (Somerville et al., 2004) to form a load-bearing structure (Pauly & Keegstra, 2016). However, *Arabidopsis*-mutants lacking xyloglucan only display minor phenotypic changes in structure and stiffness of the plant, (Cavalier et al., 2008), thus raising questions whether other polymers, such as pectin, can take over the role of xyloglucan and whether xyloglucan may in fact have other functions in the cell wall (Park & Cosgrove, 2015) such as protection against pathogens (Vincken et al., 1994). Hayashi et al. observed that considerable microfibril swelling, promoted by strong alkali treatment, is required for the extraction of xyloglucan from pea stem and hypothesized that xyloglucan is not only interacting with the cellulose surface, but may be intertwined with the cellulose fibrils (Hayashi, 1989). A possible role of the xyloglucans that cover cellulose and bind to and weave into microfibrils could be prevention of excessive hydrogen bonding between cellulose fibrils, thus avoiding cellulose aggregation and/or ensuring cell wall flexibility during cell expansion (Park & Cosgrove, 2012; Talbott & Ray, 1992). In the seeds of certain plant species, xyloglucan seems to simply act as a storage polysaccharide. In these plants, the xyloglucan is not bound to cellulose and is therefore water soluble and easily accessible for degradation to provide energy for the emerging seedling (Edwards et al., 1985).

1.2.2.2 Xylan

Whereas xyloglucan is the major hemicellulose in the primary cell walls of dicots, xylans are dominating in the secondary cell wall of dicots, as well as in the primary cell wall of grasses (Scheller & Ulvskov, 2010). Xylans represent a diverse group of polysaccharides (Figure 3) with a backbone of β-(1,4)-linked xylose residues. The xyloses may be substituted with residues of (4-O-methyl-) glucuronic acid (in glucuronoxylan), arabinose residues (in arabinoxylan or glucuronoarabinoxylan).

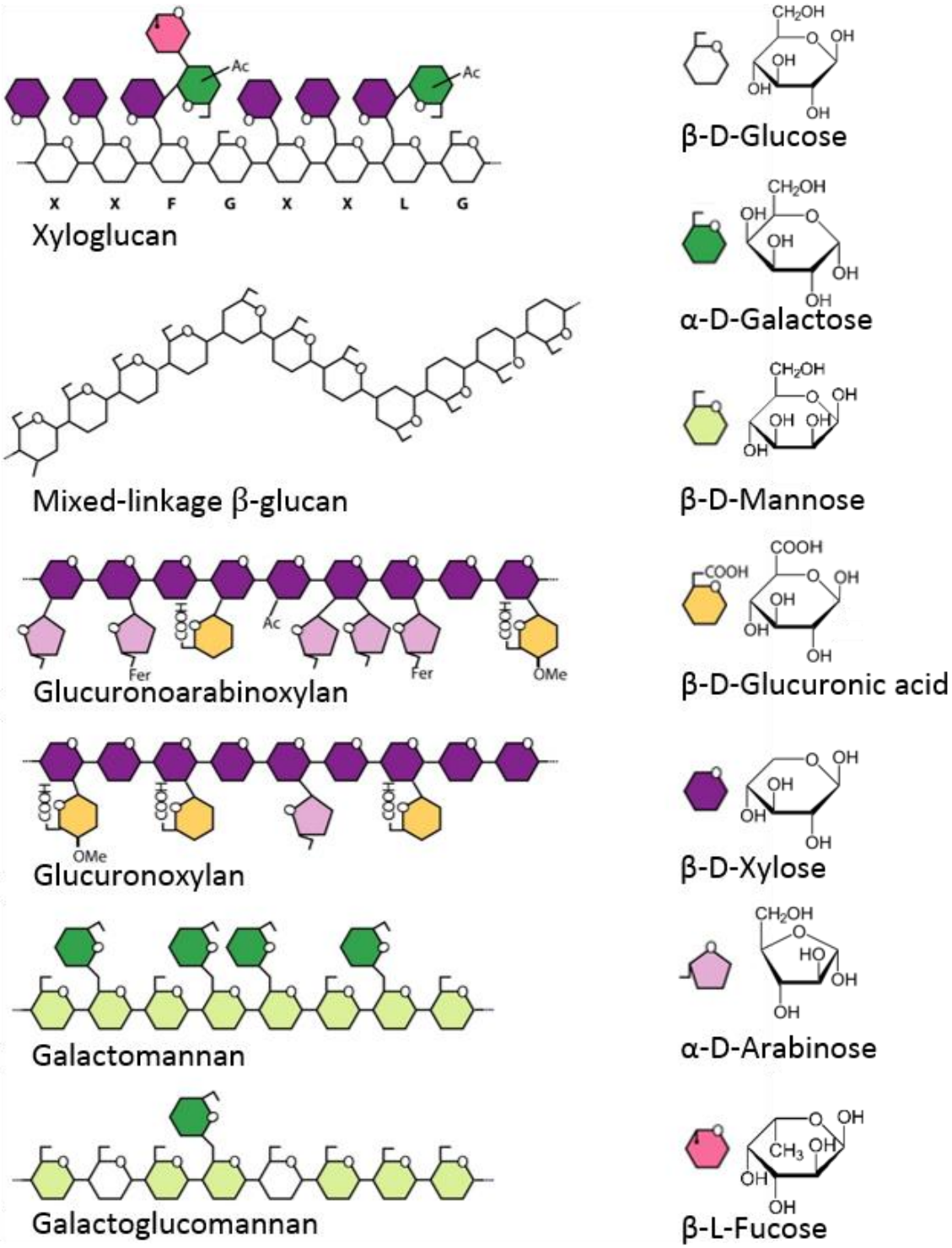

Figure 3: Common hemicelluloses in the plant cell wall. The abundance and structure of the hemicelluloses vary between plant species and tissues, as described in the text. Capital letters under the xyloglucan molecule illustrate the symbols used for the most common substituted glucoses. Other abbreviations: Fer, esterification with ferulic acid; Ac, acetylation; OMe, 4-O-methylation. The figure was adapted from (Scheller & Ulvskov, 2010).

In addition, most xylans may carry acetylations at either the O-3 or the O-2-position of the xylose residues. Studies of various substitution patterns by solid-state NMR conclude that substitution patterns leave one of the faces of the xylan chain totally unsubstituted and available to interact with the hydrophobic face of cellulose microfibrils (Busse-Wicher et al., 2016; Simmons et al., 2016) most likely as a two-fold helical screw (Busse-Wicher et al., 2014). Xylans with β-(1,3)-linked xylose residues, termed homoxylans, are found in some seaweeds (Ebringerová, 2006).

1.2.2.3 Mannans

β-(1,4)-linked polysaccharides containing mannose are widely distributed and their backbones may consist of mannose only (mannan and galactomannan), or of mannose and glucose (glucomannans and galactoglucomannans) (Scheller & Ulvskov, 2010). The backbone can be substituted with α-(1,6)-linked galactose residues and may be acetylated (Moreira & Filho, 2008).

Mannan and (galacto)glucomannan are the major hemicelluloses in the secondary cell wall of softwoods, or conifers (Figure 2), where they are suggested to play a structural role by interacting with cellulose. Galactomannans are mainly found as non-starch energy storage compounds in seeds. Furthermore, their water retaining properties protect the seed proteins from denaturation by drying (Moreira & Filho, 2008).

1.2.2.4 Mixed-linkage β-glucans

Mixed-linkage β-glucans have an unbranched backbone of β-(1,4)-linked glucose interspersed with single β-(1,3)-linkages. They are unique to the Poales where their role in the primary cell wall is linked to cell expansion (Scheller & Ulvskov, 2010). A type of mixed-linkage glucans known as lichenan occurs in some lichens such as *Cetraria islandica*. The frequency of 1,3-linkages in the glucan backbone varies between species, with high frequencies in oat and barley compared to lichenan (Ebringerová, 2006).

1.3 Enzymatic degradation of lignocellulose

Although the complex and rigid nature of the cell wall polysaccharides makes them seemingly recalcitrant and inaccessible to enzymatic degradation, there is a turnover of lignocellulosic biomass in Nature. This is due to the activity of microorganisms that have evolved a specialized enzyme system for deconstruction of polysaccharides and lignin.

Cellulolytic microorganisms are widespread in nature and may be isolated from soils, water sediments, compost and the guts of wood-degrading termites (Warnecke et al., 2007) as well as from the digestive tract of ruminants (Leschine, 1995). Deconstruction of cellulose is achieved by bacteria, fungi, protozoa, and (to some extent) archaea, under a wide range of environmental conditions (Cragg et al., 2015). One major distinction in cellulose degradation mechanisms, which seems related to the presence or absence of oxygen, lies in the arrangement of the cellulolytic enzymes (see below).

Fungi are among the most efficient degraders of lignocellulose (Blanchette, 1991). Wood-degrading fungi are generally classified as either brown or white rot, named after the color on their degradation products. Brown rot fungi are thought to utilize Fenton chemistry to generate hydroxyl radicals and degrade hemicellulose and cellulose by an oxidative mechanism, leaving lignin essentially intact and making the sugars accessible (Arantes & Goodell, 2014; Koenigs, 1974). As a result, the wood gets brown and "crispy". White rot fungi produce a range of enzymes that depolymerize cellulose, hemicellulose and lignin, and the decaying wood appears as white porous filaments. Common to most aerobe lignocellulose-degrading microorganisms, including the white rot fungi, is that they secrete numerous individual enzymes that end up in the extracellular milieu or bound to the outer membrane. Although these enzymes are not physically connected, their combined action on the substrate is favorable compared to their individual actions, a phenomenon termed synergy (Wood & Garcia-Campayo, 1990).

In contrast, anaerobe microorganisms have evolved other enzyme systems with one thing in common: all enzymes and substrate-binding domains are held in close proximity to each other and to the cell surface. One such strategy is based on the assembly of unique extracellular multi-enzyme complexes, called cellulosomes (Gilbert, 2007). The first cellulosome to be described was that of the anaerobic bacterium *Clostridium thermocellum*, which degrades cellulose under thermophilic conditions (Lamed et al., 1983). Cellulosomes tend to be connected to the cell surface through anchoring proteins. They consist of a scaffolding

protein, or scaffoldin, which contains multiple cohesion domains. Different enzymes, each containing a dockerin domain with activity to the cohesins, will then bind to the scaffoldin, thus forming a multi-molecular cellulolytic enzyme complex (Bayer et al., 2004). Binding to the substrate is ensured by a carbohydrate-binding module (CBM). The resulting close proximity of the cell to its substrate permits efficient uptake of the enzyme products (such as cellobiose) in the cell. The combination of enzymes attached to the scaffoldin ensures synergistic actions and hinders non-productive binding (Schwarz, 2001). A second strategy utilized by Gram-negative anaerobic bacteria predominantly found in rumen or gut microbiomes involves gene clusters termed polysaccharides utilization loci (PULs) (Martens et al., 2009). The co-expressed proteins encoded by one PUL include regulatory proteins, binding proteins, glycoside hydrolases (GHs) and sugar transporters, and the hydrolases tend to be tethered to the outer membrane (Pope et al., 2012; Koropatkin et al., 2008). Since PULs are substrate specific, one microorganism will normally have several PULs to enable utilization of multiple substrates like starch (Martens et al., 2009), xyloglucan (Larsbrink et al., 2014), mannan (Cuskin et al., 2015), cellulose (Naas et al., 2014), chitin (Larsbrink et al., 2016) and pectin (Ndeh et al., 2017). Although studies have shown that cellulosomes and free enzyme systems may act in synergy during degradation of cellulose (Resch et al., 2013), the focus of this thesis will be solely on free enzymes, which will be described in the following sections.

Enzymes and accessory proteins involved in the assembly or breakdown of complex carbohydrates are collectively termed carbohydrate-active enzymes (CAZymes). In an attempt to collect all online information on CAZymes across organisms and fields of study, the CAZy database was launched in 1999 (www.cazy.org). The database classification system is based on amino acid sequence similarities that reflect structural similarities and conserved catalytic mechanisms, rather than on enzyme specificity (Cantarel et al., 2009). As a consequence, enzymatic specificities classified according to the enzyme commission (EC) can be found in multiple families and one family may contain multiple enzyme specificities (Lombard et al., 2014). Currently (autumn 2017), the CAZy database covers more than 380 families that are all based on experimentally characterized proteins. The families are divided into six classes based on their mode of action: glycoside hydrolases (GHs), carbohydrate esterases (CEs), auxiliary activities (AA), glycosyl transferases (GTs), polysaccharide lyases (PLs) and the non-catalytic carbohydrate-binding modules (CBMs). It is worth noting that many CAZyme are multi-modular, for example containing a GH and a CBM.

1.3.1 Glycoside hydrolases

The GHs comprise the largest class in CAZy, with 145 families (autumn 2017). The GHs also comprise the best biochemically characterized set of enzymes in the database. GHs catalyze the cleavage of glycosidic linkages and their catalytic centers contain two conserved acidic amino acids, where, depending on the mechanism, one functions as catalytic acid or acid/base and one functions as base or nucleophile (Davies & Henrissat, 1995). In the double displacement mechanism, catalysis happens in two steps and involves the formation of a covalent intermediate, leading to retention of the anomeric carbon configuration. The alternative single displacement mechanism, requires that the two catalytic residues are somewhat further apart and leads to inversion of the anomeric carbon configuration (Koshland, 1953). GHs acting on polymeric substrates can target glycosidic bonds either at one of the chain ends or within the polymer, thus being exo- or endo-acting, respectively. The active site of exo-acting enzymes most often has a pocket or crater topology, which is adapted to binding chain ends, like starch. Alternatively, the substrate-binding site of exo-acting enzymes may be shaped as a deep groove or even tunnel, such as in the exo-acting cellobiohydrolases (Divne et al., 1994; Rouvinen et al., 1990) and see below. Endo-acting enzymes normally have more open active site architectures in the shape of a (shallow) cleft or groove, to allow random binding of several sugar units within non-crystalline polysaccharide chains (Davies & Henrissat, 1995).

Glycoside hydrolases that specifically target cellulose are known as cellulases. Cellulases hydrolyze β-(1,4)-glycosidic linkages in cellulose and are generally referred to as either endoglucanases (EGs), cellobiohydrolases (CBHs) or β-glucosidases. The non-processive endo-acting EGs (EC 3.2.1.4) randomly hydrolyze internal bonds, preferentially in the amorphous regions of the cellulose polymer (Figure 4). Currently, EGs are found in 14 GH families both as single domain enzymes or bound to CBMs. Hydrolysis of cellulose by EGs generates one reducing end and one non-reducing end that may function as accession points for the reducing end (CBH1, EC 3.2.1.176) or non-reducing end (CBH2, EC 3.2.1.91) CBHs, respectively. Both CBH1 and CBH2 have tunnel shaped active sites that enable processive degradation of cellulose by threading a single chain through the tunnel and hydrolyzing every second glycosidic linkage (Divne et al., 1994; Rouvinen et al., 1990). The CBHs are mainly found in family GH6, GH7, GH48, but also in GH5 and GH9. The cellobiose resulting from CBH action is further hydrolyzed by β-glucosidases (EC 3.2.1.21, GH family 1, 2, 3, 5_9,

5_12, 9, 30, 39, 116) to prevent product inhibition of the CBHs and providing glucose for the cell to metabolize (Henrissat et al., 1985).

In plant cell walls, cellulose is found in tight association with the diverse and complex hemicelluloses as shown in Figure 1A and discussed above. Insufficient removal of hemicellulose may limit the efficiency of cellulose saccharification (Saha et al., 2013) and saccharification of hemicellulose itself is of interest to increase overall product yields, explaining why hemicellulases are attracting increased attention (Hemsworth et al., 2016). Hemicellulases differ in their action on distinct substrates and may be endo- or exo-acting, and they are found in many different GH families. Endo-acting hemicellulases include xylanases and mannanases, whereas exo-acting enzymes, some of which primarily acting on substitutions, include arabinofuranosidases, arabinanases, galactosidases, glucuronidases, mannosidases and xylosidases. Figure 5 shows an overview of well-known hemicellulase activities. Clearly, for complete degradation of a complex hemicellulose such as glucuronoarabinoxylan, multiple enzymes are needed.

Most plant polysaccharides (with cellulose as the main exception) are O-acetylated, which may possibly serve as a protection against pathogens (Gille & Pauly, 2012). This explains why also acetyl-esterases are involved in lignocellulosic degradation. Acetyl esterases, as well as ferulic acid esterases, are described briefly in section 1.3.3 and are included in Figure 5.

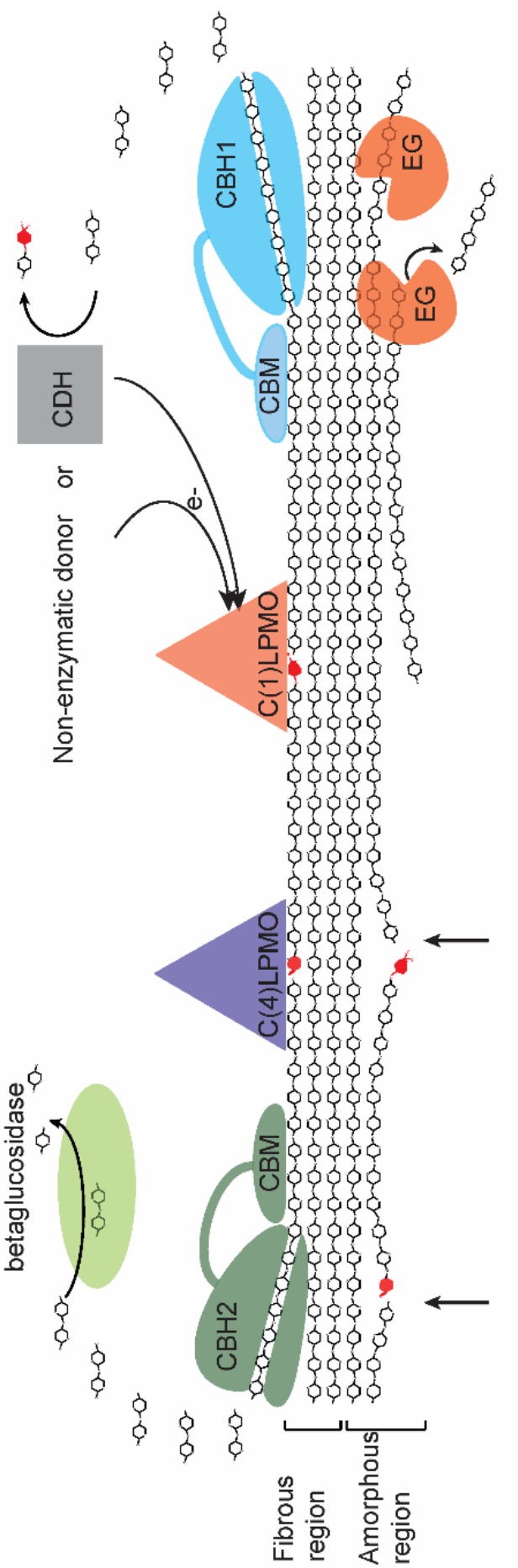

Figure 4: Schematic view of enzymatic cellulose degradation by fungi. Lytic polysaccharide monooxygenases (LPMO, triangles) use an oxidative mechanism to cleave the glycosidic linkages in crystalline cellulose, leaving one of the newly formed chain ends oxidized (oxidized sugars are coloured red), either in the reducing end (C(1), orange LPMO) or the non-reducing end (C(4), purple LPMO). The activity of LPMOs depend on electrons that can either be gained from non-enzymatic donors (e.g. ascorbate, lignin) or enzymatic donors like Cellobiose dehydrogenase (CDH; grey square) which gains its electrons from oxidation of cellobiose. LPMOs and endoglucanases (EGs) cleave randomly in the cellulose chains and generate free chain ends for the cellobiohydrolases (CBH) to attach to. The CBHs are processive enzymes that slide along the cellulose chain, while cleaving every second glycosidic linkage. The CBHs move either from the reducing end (CBH1, blue) or the non-reducing end (CBH2, dark green). The released cello-oligomers are converted by the β-glucosidase (light green) to monomeric glucose. The figure is adapted from (Horn et al., 2012).

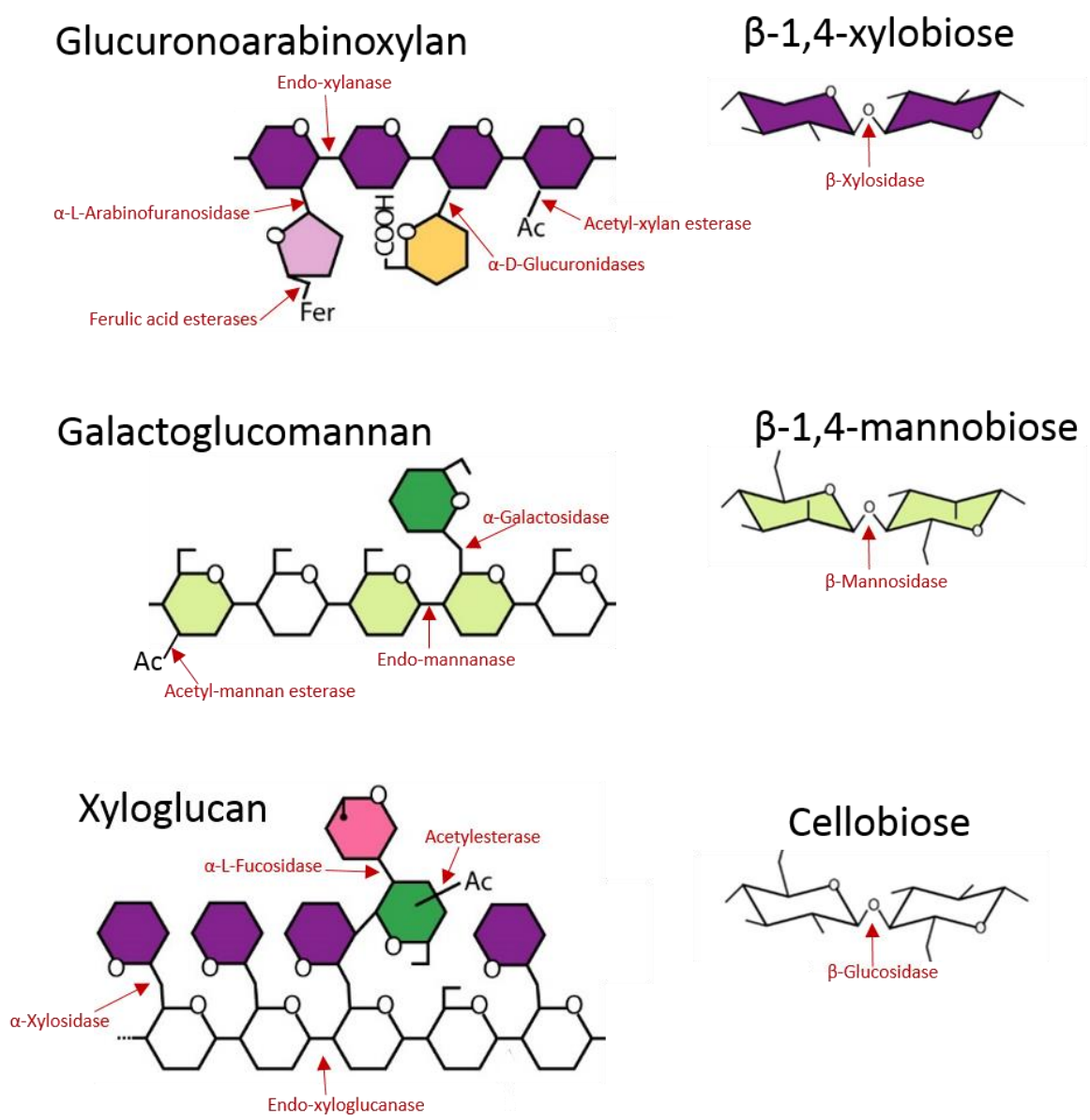

Figure 5: The action of various hemicellulases. Various exo-acting glycosidases remove substitutions, whereas endo-acting enzymes hydrolyze the β-(1,4)-linked backbone residues. Finally, β-xylosidases, β-mannosidases or β-glucosidases convert short oligosaccharides into free sugars. Esterases remove acetylations or ferulic acid from xylose or arabinose moieties, respectively. The Figure is adapted from (Scheller & Ulvskov, 2010). The color-coding of individual sugars is as in Figure 3.

One particular group of hemicellulases of particular relevance for the work decribed in this thesis are the xyloglucanases. Xyloglucan-specific endo-β-1,4-glucanases (EC 3.2.1.151, GH5_4) cleave the xyloglucan backbone at multiple sites while exo-acting xyloglucanases (EC 3.2.1.150 and EC 3.2.1.155, GH74) carry out cleavages at the reducing end (Gilbert et al., 2008). Both enzyme types may have specific preferences when it comes to which substitution pattern they can accept or prefer (Feng et al., 2014), patterns that, as mentioned

earlier, vary between plant species. Several endoglucanases (cellulases) can tolerate some degree of substitution and are therefore able to cleave non-substituted glucosyl-units in xyloglucan, while members of the GH74 family have been frequently reported to have high specificity towards xyloglucan (Feng et al., 2014; Hasper et al., 2002; Enkhbaatar et al., 2012).

1.3.2 Auxiliary activities

The latest addition to the CAZy database covers several different redox enzymes that act in conjunction with CAZymes and are classified as "auxiliary activities" (AA) (Levasseur et al., 2013). The class of AAs encompasses a range of enzyme mechanisms and substrates and was generated as a response to the reclassification of members of family GH61 and CBM33, to lytic polysaccharide monooxygenases (LPMOs) (Vaaje-Kolstad et al., 2010; Horn et al., 2012; Quinlan et al., 2011). The role and action of LPMOs (families AA9, 10, 11 and 13) will be discussed in detail in section 1.4.

Notwithstanding that lignin is not a carbohydrate, ligninolytic enzymes are included in the class of AAs due to their cooperative action with CAZymes in degradation of lignocellulose (Levasseur et al., 2013). Lignin biodegradation is most efficiently achieved by white rot fungi, and involves the action of up to four enzymes from family AA1 and AA2: laccases (EC 1.10.3.2), manganese peroxidases (EC 1.11.1.13), lignin peroxidases (EC1.11.1.14) and the versatile peroxidases (EC. 1.11.1.16). In addition to this, certain microorganisms are able to depolymerize lignin non-enzymatically, by utilizing highly reactive hydroxyl radicals generated from hydrogen peroxide (Fenton chemistry) (Pollegioni et al., 2015).

Family AA3 contains glucose-methanol-choline (GMC)-oxidoreductases, and subfamily AA3_1 mostly encompasses cellobiose dehydrogenases (CDHs, EC 1.1.99.18) which are often associated with the enzymatic conversion of lignocellulose (Kracher et al., 2015). The CDHs are two-domain proteins comprising a flavin adenine dinucleotide (FAD)-binding dehydrogenase (DH) domain coupled to a haem-binding cytochrome *b* (cyt*b*) domain. Oxidation of cellobiose by the DH domain reduces the FAD and a following inter-domain electron transfer shuffles the electrons from the reduced FAD to the haem-group, one at the time. Finally, electrons are transferred to external electron acceptors, such as LPMOs (Tan et al., 2015); see section 1.4 and Figure 4.

1.3.3 Other CAZymes

CEs remove ester-based modifications like O- and N-acetylations present in polysaccharide chains to enable or promote GH action (Cantarel et al., 2009). Hemicellulolytic esterases include acetyl-xylan esterases (EC 3.1.1.72, CE family 1-7, 12 and 15) and ferulic-acid esterases (EC 3.1.1.73, family CE1) that hydrolyze the acetyl substitutions on xylose residues and the ester-linkages between arabinose substitutions and ferulic acid, respectively (Figure 5). The latter ester linkage is involved in covalently linking xylan to lignin (Shallom & Shoham, 2003). Polysaccharide lyases cleave the glyosidic bonds of uronic acid-containing polysaccharides and are currently not exploited in the degradation of lignocellulosic biomass. As the only anabolic class of CAZymes, glycosyltransferases perform the biosynthesis of glycosidic bonds between activated sugars and other saccharides, lipids or proteins (Lairson et al., 2008).

1.3.4 Carbohydrate-binding modules

Non-catalytic CBMs are commonly found as part of modular proteins containing, for example, one or several CAZymes (Cantarel et al., 2009). By binding to specific polysaccharides, CBMs potentiate the activity of its associated enzymes and prolong their interaction with the substrate. Interestingly, the targeting actions of CBMs are not always consistent with the substrate specificities of their appended catalytic domains, which can be explained by a more general proximity effect: since plant cell walls are co-polymeric structure, binding of the CBM to one carbohydrate structure (e.g. a hemicellulose) brings the catalytic domain (e.g. a cellulase) in proximity of its preferred substrate.

In modular enzymes, CBMs are connected to the catalytic domains through linkers that can vary in length and, in fungal enzymes, their degree of glycosylation. Whether or not this linker possesses a biological function besides connecting the domains is unclear. Studies on the thermostability of cellobiohydrolases have reported positive effects of the CBM-linker-domain on overall enzyme stability (Voutilainen et al., 2008; Hall et al., 2011b). Other aspects of the CBMs have been studied, and although quantitative data are sparse, several studies have indicated that some CBMs may have a non-hydrolytic disruptive effect on fibrous cellulose (Hall et al., 2011a; Xiao et al., 2001; Din et al., 1991; Wang et al., 2008). However, these effects are not well established, and it remains to be seen whether such an effect can be exploited in industrial processing of lignocellulosic biomass. Notably, while CBMs increase substrate affinity, they may also delay enzyme efficiency by their stickiness, which reduces

off-rates. Interestingly, a study by Várnai et al. indicated that the presence of CBMs may be unfavorable at high substrate concentrations (20% w/w), thus questioning the benefit of CBMs in an industrial biorefinery setting (Várnai et al., 2013).

1.4 Lytic Polysaccharide Monooxygenases (LPMOs)

1.4.1 The history of LPMOs

For long, the conversion of cellulose into glucose was thought to be performed solely by cellulases (Pringsheim, 1912), even though indications of a more complex system for cellulose degradation had been observed (Seillière, 1907). In 1950, Reese and co-workers studied the liquification of a cellulose derivative (carboxy methyl cellulose; CMC) and found that both cellulosic and non-cellulosic yeast were able to utilize CMC, whereas only the cellulosic organisms were able to grow on crystalline cellulose. From this, it was hypothesized that microorganisms that can utilize native cellulose deploy at least two enzyme systems, named C_1 and C_x (Reese et al., 1950). In this proposed model, the C_1 system would convert the native cellulose into more accessible linear polysaccharides, which would be further degraded and utilized by both cellulolytic and non-cellulolytic organisms through the C_x system. The latter system was thought to comprise hydrolytic enzymes capable of cleaving β-(1,4)-glycosidic linkages between glucose units, whereas the precise action of system C_1 was not clear at that time.

Twenty-four years later, in 1974, Eriksson et al. described a two-fold enhancement in degradation of cellulose by *Sporotrichulum pulverulentum* in the presence of oxygen, compared to a nitrogenous atmosphere (Eriksson et al., 1974). These authors suggested that the fungi utilized an oxidative enzyme mechanism involving incorporation of uronic acid moieties into the cellulose, which would disturb hydrogen bonding between cellulose chains, which in turn would cause swelling of the cellulose and make the crystalline parts more accessible (Eriksson et al., 1974). These observations strengthened the hypothesis put forward by Reese et al. in 1950. In 1992, a novel fungal enzyme, CelI, form *Agaricus bisporus* was cloned and described as a potential cellulose degrading enzyme (Raguz et al., 1992), although no activity was observed. In retrospect, this was probably the first AA9 to be identified (Lo Leggio et al., 2012). Following this, several proteins were characterized and classified in family GH61, despite these enzymes having very low cellulose degrading activity compared

to well-known cellulases [(Beeson et al., 2015) and references within]. In 2007, Merino and Cherry described that addition of a GH61 had a boosting effect on the activity of conventional hydrolytic cellulases in degradation of lignocellulolytic substrates (Merino & Cherry, 2007). Notably, this synergistic effect was only observed when using lignocellulosic substrates and not when using pure cellulose as substrate. No hydrolytic activity was observed upon incubation of only GH61 with lignocellulosic substrates. In 2010, Harris et al. described similar results [(Harris et al., 2010); see below].

In the 1990ies, several bacterial proteins were isolated from chitinolytic microorganisms with high affinity for chitin (Schrempf, 1999; Schnellmann et al., 1994; Kolbe et al., 1998), a polymer of β-(1,4)-linked *N*-acetylglucosamine found as a structural component in the cell walls of fungi and the exoskeletons of crustaceans and insects. Similarly as cellulose, chitin forms crystalline structures that are resistant to enzymatic degradation. One such chitin-binding protein (or CBP) was found covalently bound to a mannanase (Sunna et al., 2000). Since the CBP also contained conserved aromatic amino acids (Zeltins & Schrempf, 1997), which is typical for CBMs, these proteins were classified as CBM family 33 (CBM33). The structure of CBP21 from *Serratia marcescens* (Figure 6A), which was resolved in 2005, revealed an "atypical" CBM, with a conserved polar surface and showed that most aromatic residues were located in the core of the protein (Vaaje-Kolstad et al., 2005b). Shortly after, Vaaje-Kolstad et al. described disruption of crystalline chitin by CBP21, visualized by scanning electron microscopy, and showed that CBP21 had a strong boosting effect on the degradation of chitin by GH18 and GH19 chitinases (Vaaje-Kolstad et al., 2005a).

In 2008, the first crystal structure of a GH61, *Hypocrea jecorina* (*Tricoderma reesei*) Cel61B, was determined (Karkehabadi et al., 2008). The structure of Cel61B revealed a flat surface with a stretch of highly conserved amino acids, unlike the tunnel or cleft active site typically found in cellulases. In addition, the protein lacked the conserved carboxylate residues that catalyze hydrolytic cleavage in GHs. Importantly, the structures of Cel61B and CBP21 were strikingly similar, providing the first clue that fungal GH61s and bacterial CBM33s are related and potentially linking the boosting effects observed in 2005 by Vaaje-Kolstad et al. in chitin degradation to the boosting effects described in 2007 by Merino and Cherry et al. in cellulose degradation. The flat surface of Cel61B contained a nickel ion (from the crystallization buffer), which seemed coordinated by two conserved two conserved histidines, one of which was the N-terminal histidine (Karkehabadi et al., 2008). Structural superpositions, showed that both histidine residues in Cel61 align with the corresponding histidines in CBP21.

In 2010, the structure of GH61E from *Thielavia terrestris* and the boosting effect of this protein on cellulose degradation by cellulases were described by Harris et al. (2010). Importantly, this study showed that the boosting effect of the GH61 depends on a bound divalent metal ion. From the study by Harris et al., it appeared that several divalent metal ions could be accommodated by GH61s, and that removal of such metals (e.g. by metal ion chelators such as EDTA) diminished the effect of the GH61 on cellulase efficiency (Harris et al., 2010). In accordance with the results of mutational studies of chitin-active CBP21 (Vaaje-Kolstad et al., 2005a), Harris et al. observed a loss in the cellulase-boosting activity of GH61E, upon mutating conserved amino acid residues in or near the metal binding site. Later in 2010, Vaaje-Kolstad et al., published a landmark paper showing that CBP21 is an enzyme that, in the presence of oxygen and an electron donor, carries out oxidative cleavage of glycosidic bonds in crystalline chitin to promote further degradation by conventional glycoside hydrolases (Vaaje-Kolstad et al., 2010). Based on the similar features of CBP21 and GH61 proteins described above, Vaaje-Kolstad et al. predicted that GH61s utilize a similar mechanism when acting on cellulosic substrates, thus explaining the observed boosting effects of GH61s on cellulase efficiency. In the following year, oxidative cleavage of cellulose was reported for another bacterial CBM33 (Forsberg et al., 2011) and several fungal GH61s (Langston et al., 2011; Phillips et al., 2011; Quinlan et al., 2011; Westereng et al., 2011). The studies by Quinlan et al. and Phillips et al. led to the discovery that the metal binding site is a type 2 copper site and that binding of one single copper ion is essential for LPMO activity.

The two enzyme families were collectively named lytic polysaccharide monooxygenases (LPMOs) (Horn et al., 2012), based on work by the Marletta-group (Phillips et al., 2011) and suggestions in (Mba Medie et al., 2012). The findings that LPMOs are neither GHs nor CBMs led to a reclassification in 2013 into families of auxiliary activities (AA) (Levasseur et al., 2013). Family AA10 contains CBM33-type LPMOs, whereas GH61-type LPMOs appear in family AA9. According to common practice for CAZymes, the LPMOs discussed in detail above are hereafter referred to as follows: CBP21, *Sm*LPMO10A; Cel61B, *Hj*LPMO9B; GH61E, *Tt*LPMO9E.

1.4.2 Substrate specificity and occurrence of LPMOs

While AA9 enzymes are of fungal origin, the AA10s are found in all domains of life. In 2012 all characterized family AA9 LPMOs (LPMO9s) were active on cellulosic substrates, whereas both chitin-active and cellulose-active AA10 LPMOs (LPMO10s) had been described. Since then, additional activities have been discovered for fungal LPMOs, including activity on soluble cello-oligosaccharides (degree or polymerization, DP >3) [Paper I in this thesis and (Frandsen et al., 2016)] and various hemicelluloses with β-(1,4)-linked glucan or xylan backbones [Paper II in this thesis and (Bennati-Granier et al., 2015; Jagadeeswaran et al., 2016; Frommhagen et al., 2015; Kojima et al., 2016; Nekiunaite et al., 2016)]. Recently, novel fungal LPMOs not belonging to the AA9 or AA10 family have been discovered, which oxidize chitin or starch. These LPMOs are classified as AA11 (Hemsworth et al., 2014) and AA13 (Vu et al., 2014b; Lo Leggio et al., 2015), respectively. Table 1 summarizes the main characteristics of LPMOs with determined crystal structures to date.

Table 1. Key features of LPMOs with known crystal structures. Cell, cellulose; Celloolig, cello-oligosaccharides; XG, xyloglucan; NP, not published; *, C1/C4-oxidizing on cellulose but strictly C1-oxidizing on chitin. Note that not all LPMOs have been tested with all possible substrates, meaning that, likely, several activities of the LPMOs listed below have remained undetected. For some of the xyloglucan active LPMOs, activity on glucomannan and/or β-glucan has also been demonstrated

Enzyme	Organism	AA-class	Substrate preference	Regio-specificity	PDB-code	Structural Reference
*Hj*LPMO9A (EGIV,Egl4,EG4)	*Hypocrea jecorina*	AA9		C1/C4	502W	(Hansson et al., 2017)
*Hj*LPMO9B (EG7, Cel61B)	*Hypocrea jecorina*	AA9	NP	NP	2VTC	(Karkehabadi et al., 2008)
*Ls*LPMO9A (*Ls*AA9A)	*Lentinus similis*	AA9	Cell, Cellooligo, Xylan	C4	5ACH	(Frandsen et al., 2017)
*Nc*LPMO9A (NCU02240)	*Neurospora crassa*	AA9	Cell	C4	5FOH	NP
*Nc*LPMO9C (NCU02916)	*Neurospora crassa*	AA9	Cell, Cellooligo, XG	C4	4D7U	(Borisova et al., 2015)
*Nc*LPMO9D (PMO-2, NCU01050)	*Neurospora crassa*	AA9	Cell, Cellooligo	C4	4EIR	(Li et al., 2012)
*Nc*LPMO9F (NCU03328)	*Neurospora crassa*	AA9	Cell	C1	4QI8	(Tan et al., 2015)
*Nc*LPMO9M (PMO-3, NCU7898)	*Neurospora crassa*	AA9	Cell	C1/C4	4EIS	(Li et al., 2012)
*Pc*LPMO9D (GH61D)	*Phanerochaete chrysosporium*	AA9	Cell	C1	4B5Q	(Wu et al., 2013)
*Ta*LPMO9A (GH61A)	*Thermoascus aurantiacus*	AA9	Cell, XG	C1/C4	2YET	(Quinlan et al., 2011)
*Mt*LPMO (*Mt*PMO3*, MYCTH_92668)	*Thermothelomyces thermophila*	AA9	Cell	C1	5UFV	(Span et al., 2017)
*Tt*LPMO9E (GH61E)	*Thielavia terrestris*	AA9	Cell	C1	3EII	(Harris et al., 2010)
*Ba*LPMO10A (Chb, *Ba*CBM33)	*Bacillus amyloliquefaciens*	AA10	Chitin	C1	2YOX	(Hemsworth et al., 2013b)
*Bp*LPMO10A (*Bp*AA10A)	*Burkholderia pseudomallei*	AA10	NP	NP	3UAM	NP
*Bt*LPMO10A	*Bacillus thuringiensis*	AA10	Chitin	C1	5WSZ	NP
*Cj*LPMO10A	*Cellvibrio japonicus*	AA10	Chitin	C1	5FJQ	(Forsberg et al., 2016)
*Ef*LPMO10A (*Ef*AA10A)	*Enterococcus faecalis*	AA10	Chitin	C1	4A02	(Vaaje-Kolstad et al., 2012)
*Jd*LPMO10A	*Jonesia denitrificans*	AA10	Chitin	C1	5AA7	(Mekasha et al., 2015)
*Lm*LPMO10A (LMRG_01781, Lmo2467)	*Listeria monocytogenes*	AA10	Chitin	C1	5L2V	Unpublished
*Sc*LPMO10C (CelS2, *Sc*AA10C)	*Streptomyces coelicolor*	AA10	Cell	C1	4OY7	(Forsberg et al., 2014a)
*Sc*LPMO10B	*Streptomyces coelicolor*	AA10	Cell, Chitin	C1/C4*	4OY6	(Forsberg et al., 2014a)
*Sl*LPMO10E	*Streptomyces lividans*	AA10	Chitin	C1	5FTZ	(Chaplin et al., 2016)
*Sm*LPMO10A (CBP21)	*Serratia marcescens*	AA10	Chitin	C1	2BEM	(Vaaje-Kolstad et al., 2005b)
*Tf*LPMO10A (E7)	*Thermobifida fusca*	AA10	Cell, Chitin	C1/C4*	4GBO	NP
*Vc*LPMO10B (GbpA,*Vc*AA10B)	*Vibrio cholerae*	AA10	Chitin	C1	2XWX	(Wong et al., 2012)
*Ae*LPMO10virus (ACV034)	*Anomala cuprea entomopoxvirus*	AA10	NP	NP	4YN1	(Chiu et al., 2015)
*We*LPMO10virus	*Wiseana entomopoxvirus*	AA10	NP	NP	4YN2	(Chiu et al., 2015)
*Me*LPMO10virus	*Melolontha entomopoxvirus*	AA10	NP	NP	4OW5	(Chiu et al., 2015)
*Ao*LPMO11 (*Ao*AA11)	*Aspergillus oryzae*	AA11	Chitin	C1	4MAI	(Hemsworth et al., 2014)
*Ao*LPMO13 (*Ao*AA13)	*Aspergillus oryzae*	AA13	Starch	C1	4OPB	(Lo Leggio et al., 2015)

During growth on biomass, white-rot fungi express both lignin- and carbohydrate-active redox enzymes, simultaneously with hydrolytic enzymes (Floudas et al., 2012). While bacterial genomes generally encode less than 10 LPMOs, and often only one, LPMO-encoding genes are abundant in the genomes of many fungi. Fungal genomes may contain more than 30 LPMO-encoding genes and the number of LPMO-encoding genes can be much higher than the number of cellulase encoding genes (Berka et al., 2011; Hori et al., 2014; Tian et al., 2009; Navarro et al., 2014; Lenfant et al., 2017). Growth on various biomasses affects the level and type of LPMOs that are upregulated, indicating specific substrate preferences among LPMOs. As shown in Table 2, the number of LPMO-encoding genes in the genomes of white-rot fungi varies dramatically. Several white-rot fungi also secrete CDH, an extracellular flavochytochrome that can fuel LPMOs with electrons and that is often thought to be associated with LPMO activity (Phillips et al., 2011) (Figure 4). It should be noted though that several other fungal redox enzymes typically found in fungal secretomes are capable of activating LPMOs (Kracher et al., 2016; Garajova et al., 2016).

Although brown-rot fungi are capable of degrading biomass, their genomes suggest that they have an "incomplete" enzymatic system for biomass degradation, lacking key enzymes such as cellobiohydrolases and ligninolytic enzymes (Eastwood et al., 2011; Martinez et al., 2009). Brown-rot fungi also tend to have few LPMOs (Lenfant et al., 2017). For example the genome of the brown-rot fungus *Gleophylum trabeum* encodes only four LPMOs (Table 2), of which one has been functionally characterized (Kojima et al., 2016).

LPMOs have also been found in the secretomes of both cellulolytic and chitinolytic bacteria cultivated on lignocellulosic or chitinous biomass, respectively (Tuveng et al., 2016; Suzuki et al., 1998; Takasuka et al., 2013). Compared to e.g. white-rot fungi, bacteria encode only a few LPMOs in their genome. However, AA10 LPMOs are widely spread and are found in all sorts of living organisms (Floudas et al., 2012; Horn et al., 2012) and some has even been identified as potential virulence factors including the N-terminal domain of GbpA from *Vibrio cholerae* (*Vc*LPMO10B) (Kirn et al., 2005) and *Lm*LPMO10A (Lmo2467), an LPMO from *Listeria monocytogenes* (Paspaliari et al., 2015). Recently, the crystal structures of three LPMO10s from an entomopoxvirus have been resolved (Chiu et al., 2015). So far, no bacterial counterpart of CDH has been identified, which indicates that bacterial LPMOs are activated by other redox enzymes.

Table 2. Occurrence of LPMO-encoding genes in the genomes of selected microorganisms. Most of the data shown here were taken from the genome section of CAZy. An * indicates that the genome is not listed in the CAZy database and information is based on single CAZy-entries and in-house genome searches using the JGI database (https://jgi.doe.gov/). Families AA9, AA10, AA11 and AA13 are LPMOs. The AA3_1 subfamily contains cellobiose dehydrogenases. WR, white-rot; BR, brown-rot fungi.

Organism	Phylogeny	AA9	AA10	AA11	AA13	AA3_1
Aspergillus oryzae	Ascomycota	8	0	5	1	2
*Hypocrea jecorina**	Ascomycota	3	0	3	0	0
Neurospora crassa	Ascomycota	14	0	4	1	2
*Phanerochaete chrysosporium**	Basidiomycota (WR)	16	0	0	0	1
*Thermoascus aurantiacus**	Ascomycota	3	0	3	0	0
Thermothelomyces thermophila	Ascomycota	22	0	4	1	3
Thielavia terrestris	Ascomycota	22	0	5	1	3
Gloeophyllum trabeum	Basidiomycota (BR)	4	0	0	0	1
Streptomyces coelicolor	Actinobacteria	0	7	0	0	0
Serratia marcescens	Proteobacteria	0	1	0	0	0

1.4.3 The three-dimensional structure and copper-binding site of LPMOs

Despite the low sequence similarities, the overall fold of LPMOs is conserved. The conserved core structure of LPMOs resembles both fibronectin type III (FnIII) and immunoglobulin (Ig) domains. Both are β-sandwich modules made up of seven to nine anti-parallel β-strands. The strands are connected by loops of varying lengths, some of which may contain short helices. As show in Figures 6 and 7, loop variations are more pronounced at the substrate-recognizing interface of the protein (Span & Marletta, 2015; Vaaje-Kolstad et al., 2017). In particular, the long "L2" region (colored red in Figure 6A) is highly variable and is presumably an important determinant of (variations) in substrate specificity (Li et al., 2012; Wu et al., 2013; Forsberg et al., 2014a). The substrate-interacting surfaces of LPMO9s are further defined by three additional loops, which have been named LS, LC and L3 (Figure 6A) (Wu et al., 2013; Vaaje-Kolstad et al., 2017).

The flat surface of LPMOs accommodates the active site that includes a solvent-exposed type-2 copper center (Quinlan et al., 2011). The copper ion is coordinated by three nitrogen ligands provided by two fully conserved histidine residues, of which one is the N-terminal residue,

which form a T-shaped histidine brace (Figure 6B). In homologously expressed fungal LPMOs, the N-terminal histidine is Nε-methylated; the functional significance of this post-translational modification remains unknown. Additional coordination positions for the copper are usually occupied by water in the equatorial position and in the solvent exposed (distal) axial position, and a tyrosine/phenylalanine in the proximal axial position (Quinlan et al., 2011). However, the presence of water depends on the oxidation state of the bound copper, since the two waters are not typically present in the reduced state (Cu^{+}) (Gudmundsson et al., 2014; Hemsworth et al., 2013b). Notably, most of the LPMO structures available are in the reduced state as a result of photoreduction by the X-ray beam (Beeson et al., 2015; Frandsen & Lo Leggio, 2016).

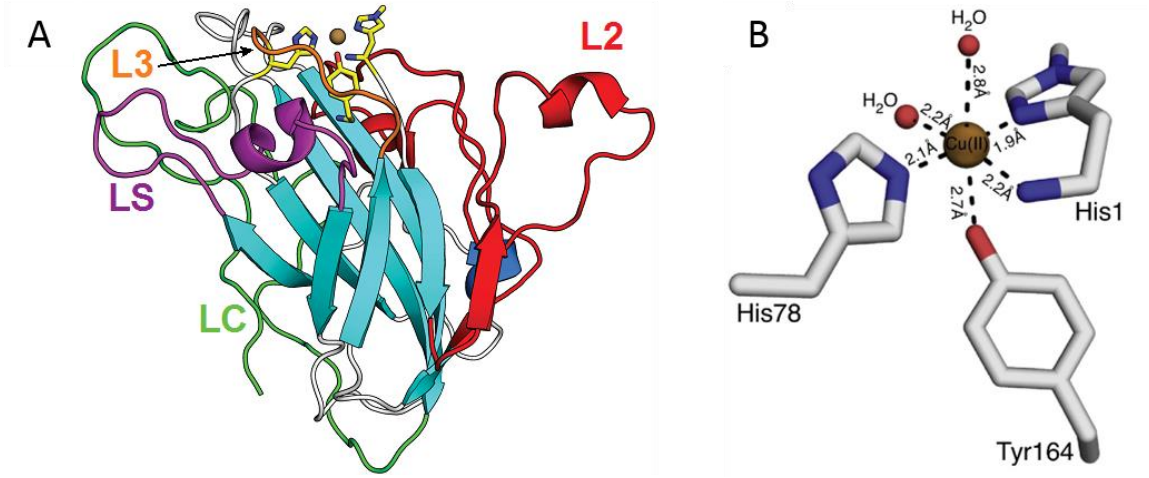

Figure 6: The overall structure and the active site in family AA9 LPMOs. (**A**) The structure of *Nc*LPMO9M (PDB-code: 4EIS). Variable loop regions that form part of the substrate binding surface are indicated (L2, red; LC, green, LS, purple). A subgroup of LPMOs contain an extra insertion in the region marked "L3" in the figure (Borisova et al., 2015; Fanuel et al., 2017). (**B**) The copper active site of *Ls*LPMO9A (PDB-code: 5ACG). The presence of two copper-coordinating water molecules indicates that the copper ion is in the Cu^{2+} state. Both figures are adapted from (Vaaje-Kolstad et al., 2017)

The flat substrate-binding surface of fungal lignocellulose-active LPMOs contains solvent exposed aromatic side chains that are believed to π-stack with the glucose units of cellulose (Li et al., 2012). The corresponding amino acids in the fungal chitin-oxidizing *Ao*LPMO11 have hydrophilic side chains that can possibly form hydrogen bonds with the acetyl moieties of chitin (Hemsworth et al., 2014). The starch oxidizing *Ao*LPMO13 has a shallow groove along the active site surface, which may accommodate helical structures in the starch substrate (Lo Leggio et al., 2015). Bacterial LPMOs, regardless of substrate preference, have largely hydrophilic substrate-recognizing surfaces with the exception of one conserved aromatic acid,

which has its ring oriented parallel to the binding surface and could thus participate in stacking interactions (Forsberg et al., 2014a). A structure-based clustering of all LPMOs, revealed nine clusters, which differ from one another, not only in structure, but also in substrate preferences and oxidative regioselectivity (see below), as shown in Figure 7 (Vaaje-Kolstad et al., 2017). Adding even more variation, in some LPMOs, the C-terminus of the catalytic domain is covalently attached via a flexible linker to a CBM and/or (rarely) to other domains, for example GHs (Vaaje-Kolstad et al., 2017).

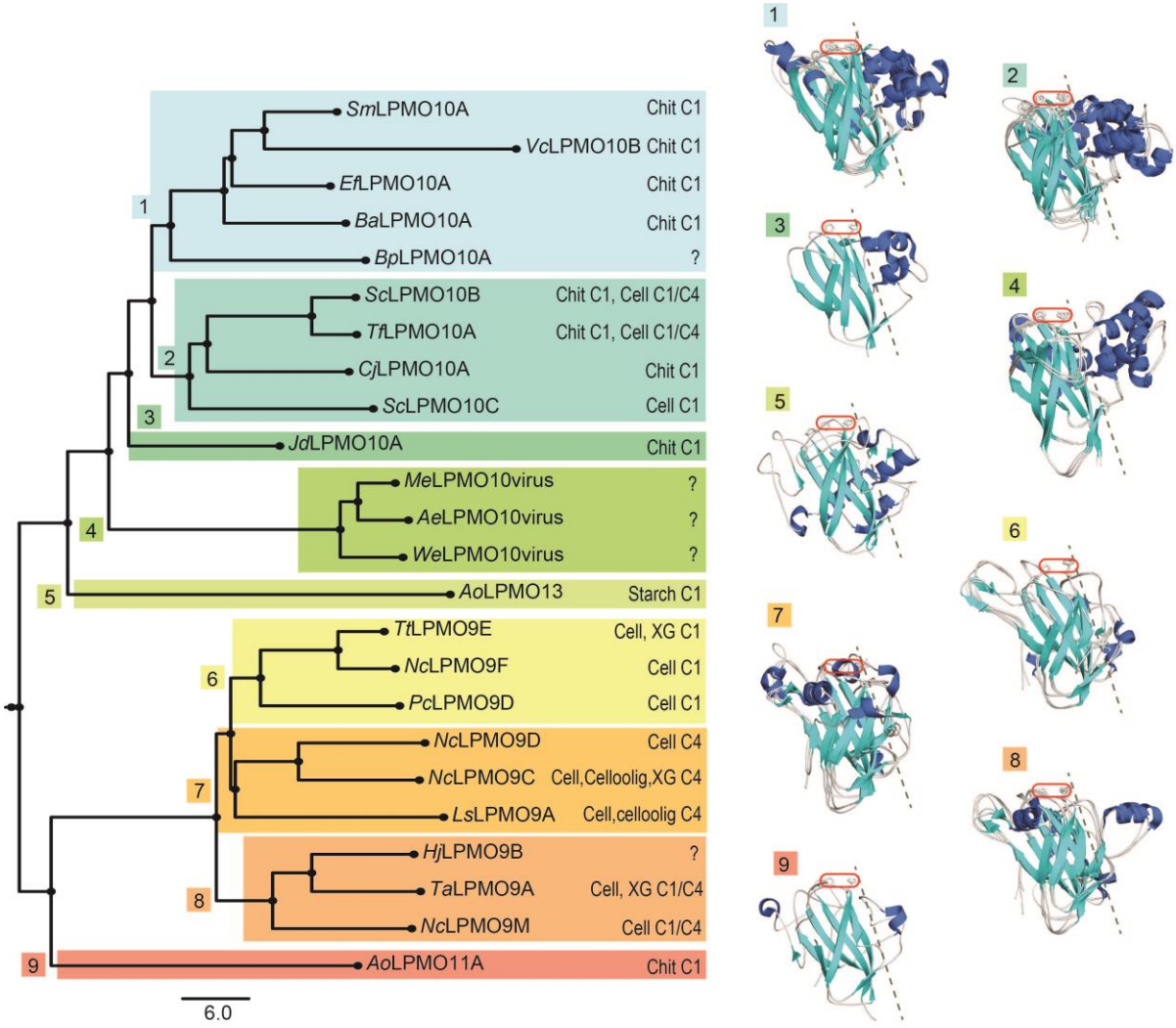

Figure 7: Structural diversity among LPMOs. The dendrogram at the left shows the structural clustering of 24 LPMOs with their substrate preference and oxidative regioselectivity (C1, C4 or C1/C4). The scale indicates the DALI Z-score. The right side of the figure show representative structures for each clade, shown structurally aligned. The two Cu-coordinating histidines are shown in stick representation and marked by a red oval. A dashed line separates the core β-sandwich (left) from the highly variable L2-region (right). Chit, Chitin; Cell, cellulose; Celloolig, cello-oligosaccharides; XG, xyloglucan. PDB-identifiers for each enzyme is listed in Table 1. The figure is taken from (Vaaje-Kolstad et al., 2017).

1.4.4 LPMO mechanism

LPMOs are copper dependent enzymes that, in the presence of an electron donor and molecular oxygen attacks the C1 or C4 in the glycosidic bond of its substrate and breaks the glycosidic linkage (Figure 8 and 9). The abstraction of one hydrogen atom from either the C1 or C4 carbon in the glycosidic bond, and subsequent hydroxylation followed by a spontaneous elimination reaction of the resulting substrate radical results in scission of the bond (Quinlan et al., 2011; Phillips et al., 2011) .

The reaction mechanism of LPMOs is still unclear even though several attempts have been made to understand it, as reviewed by others (Walton & Davies, 2016; Beeson et al., 2015). Mutations of residues in the substrate-binding surface or active site, quantum mechanical calculations and activity measurements under varying conditions have given some insight. In all catalytic scenarios, copper reduction is the starting point, which is somehow followed by generation of a reactive oxygen species that is sufficiently powerful to abstract a hydrogen from the substrate (Walton & Davies, 2016). One remaining question is how the second electron is transferred to the catalytic center. The LPMO can likely only store one electron, by copper reduction, and in the proposed mechanism, the second electron needs to arrive while the copper ion is interacting with the substrate and thus not very accessible.

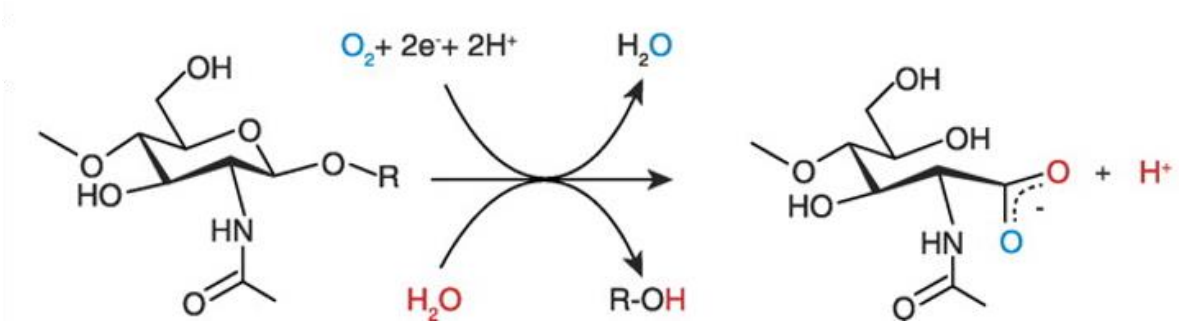

Figure 8: Proposed oxidative cleavage of chitin by *Sm*LPMO10A. The resulting aldonic acid contain one oxygen atom from water (red) and one from molecular oxygen (blue). The figure is taken form (Vaaje-Kolstad et al., 2010) Copyright© (2010) AAAS.

A recent study by Bissaro et al. indicated that LPMOs may not be using molecular oxygen but rather hydrogen peroxide (Bissaro et al., 2017). The authors propose a mechanism, where after initial reduction of the enzyme-bound copper ion, the enzyme carries out multiple catalytic cycles using H_2O_2 ($2H^+ + O_2 + 2e^-$, see Fig 8 for comparison). Importantly, Bissaro et al. point out that although H_2O_2 is a required co-substrate for LPMO catalysis, an excess of H_2O_2 in the reaction will generate reactive oxygen species that can lead to oxidative self-

inactivation of the LPMO. Thus, the level of H_2O_2 must be controlled to ensure stable and optimal LPMO performance (Bissaro et al., 2017). Lack of such control may explain why high quality kinetic data for LPMOs are scarce in the literature.

According to the originally proposed O_2-dependent LPMO reaction mechanism (Figure 8), each catalytic cycle requires the supply of two electrons. Since the discovery of LPMO activity, several possible small molecule electron donors have been identified, e.g. ascorbic acid, glutathione, gallic acid, catechin, hydroquinone and caffeic acid (Vaaje-Kolstad et al., 2010; Quinlan et al., 2011; Isaksen et al., 2014; Westereng et al., 2015; Lo Leggio et al., 2015). Although hints on the electron donating potential of lignin have been present in the literature since Merino and Cherry described that the cellulase-boosting effect of GH61s was most prominent in the presence of lignin (Merino & Cherry, 2007), lignin effects were not properly established until later, as described in e.g. (Westereng et al., 2015; Hu et al., 2014).

Fungal secretomes contain several redox enzymes that can activate LPMOs (Garajova et al., 2016; Kracher et al., 2016). The most well-known of these is CDH (described in section 1.3.2), which notably, does not occur in the secretomes of all LPMO producing fungi (Kracher et al., 2016), is thought to be part of the *in vivo* electron donating system that LPMOs require for activity (Sygmund et al., 2012; Langston et al., 2011; Beeson et al., 2012). Speculations have been made as to how CDH delivers electrons to the LPMO and various electron delivery pathways have been suggested. In a recent study, Courtade et al. used protein NMR to demonstrate that CDH competes with the substrate in binding to the active site of *Nc*LPMO9C to reduce the copper (Courtade et al., 2016). This finding raises questions as to how CDH would deliver the second electron to the copper site during catalysis (see above), but is compatible with the H_2O_2-based mechanism proposed by Bissaro et al. (2017).

1.4.5 LPMO regioselectivity in cellulose oxidation

Except for the starch-active LPMOs in family AA13, all LPMOs characterized so far, catalyze cleavage of the β-(1,4)-glycosidic linkages connecting the backbone glucosyl or xylosyl residues of cellulose and several hemicelluloses, or N-acetylglucosamine residues in the backbones of chitin. As shown in Figure 8 for chitin oxidation and in Figure 9 for cellulose oxidation, two new chain ends are formed; one oxidized chain end and one native chain end. Oxidation of C1 leads to the formation of a native non-reducing end and one oxidized chain end in the form of a lactone. The lactone is in equilibrium with the aldonic acid form, which

is dominating at neutral and higher pH (Westereng et al., 2013). Oxidation of the C4-carbon leads to the formation of a native reducing end whereas the non-reducing end is a 4-ketosugar that, in aqueous solution, is hydrated to a gemdiol (Westereng et al., 2016). Some LPMOs generate both C1- and C4-oxidized products (referred to as "C1/C4-oxidizers"). Compared to cellulose oxidizing LPMOs, chitinolytic LPMOs are seemingly restricted to oxidation of the reducing end, meaning that they only generate C1-oxidized products.

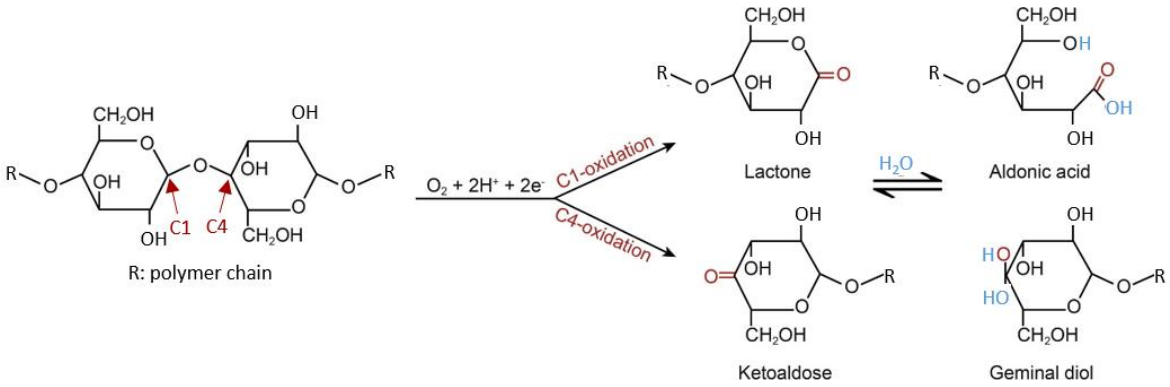

Figure 9: Cellulose oxidation by an LPMO at the C1 or C4 carbon. The figure has been adapted from (Loose et al., 2014).

Bioinformatics have been used in attempts to subgroup LPMOs based on their site of oxidation, also known as regioselectivity. Distinct features of the loop regions, particularly the L2 region close to the N-terminus of the protein (colored red in Figure 6A; see also Figure 7), have been anticipated to discriminate between LPMO regioselectivity (Vu et al., 2014a). In bacterial LPMO10s, the L2 region contributes to at least 50% of the substrate-binding surface (Forsberg et al., 2014a), whereas in LPMO9s, the L2 is generally less extended, with the longest region occurring in C1/C4-oxidizers (Figure 7, clade 8) (Vaaje-Kolstad et al., 2017). In an attempt to relate the extra insertion in the L2 region of C1/C4-oxidizing LPMO9s to the regioselectivity, Vu et al. removed the 11 amino acid insert, including one tyrosine that aligns with the active site surface and may be involved in substrate interactions. The removal of this insertion diminished the C4-oxidizing activity, while keeping the C1 activity intact (Vu et al., 2014a). Further, Vu et al. noted the possible role of the surface tyrosine, but mutation of this residue to a glycine did not alter regioselectivity.

Recent mutational studies by Danneels et al. using *Hj*LPMO9A indicated that both the tyrosine in the L2 region and a tyrosine aligning with the substrate-binding surface on the opposite side of the active site (in the LC-region; Figure 6) are involved in determining the C4/C1-oxidation ratio (Danneels et al., 2017). This latter tyrosine, in the LC loop, is highly conserved among LPMO9s and its importance for substrate binding has been experimentally

shown by both X-ray crystallography and NMR (Frandsen et al., 2016; Courtade et al., 2016). A crystal structure of *Ls*LPMO9A with a cello-oligosaccharide with a degree of polymerization of 6 (DP6) bound, showed that the LC-tyrosine is involved, presumably by CH-π interactions, in binding of cellulosic substrates (Frandsen et al., 2016). In NMR studies, the corresponding LC-tyrosine in *Nc*LPMO9C showed chemical shifts upon addition of xyloglucan (Courtade et al., 2016). Although several attempts have been made to relate the regioselectivity and substrate preference of LPMOs to their amino acid sequence, so far none have succeeded in accurate prediction of substrate specificity or the site of oxidation (Lenfant et al., 2017). Interestingly, structural comparisons have indicated a correlation between regioselectivity and the accessibility of the solvent exposed axial copper-coordination site in LPMO10s (Forsberg et al., 2014a). This correlation also seems to apply to LPMO9s, as shown in Figure 10 and Paper III of this thesis.

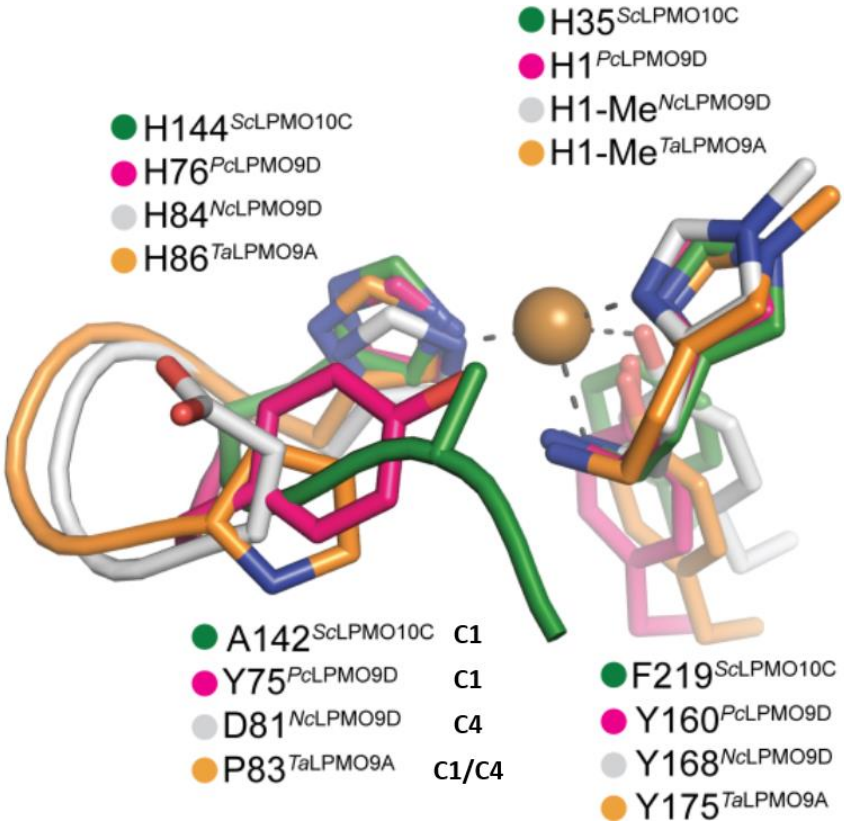

Figure 10. Occupancy of the solvent-exposed axial coordination site in LPMOs with varying regioselectivity. The picture shows high conservation of the active site residues, except the residue that determines accessibility of the axial copper coordination site (A142 in *Sc*LPMO10C; Lower left of the Figure). The regioselectivity (C1, C4 or C1/C4) of the individual LPMOs is indicated. The copper ion is shown as a golden sphere. The presence of a methyl group on H1 in the fungal LPMOs depends on how the crystallized enzyme was produced, in a fungus (with methyl) or in *Pichia pastoris* (no methyl). The figure is adapted from (Forsberg et al., 2014a).

1.4.6 Determination of the product profile of LPMOs

When LPMO activity was first described, the products generated during degradation of chitin by *Sm*LPMO10A were analyzed by hydrophobic interaction liquid chromatography (HILIC) and matrix-assisted laser desorption/ ionization time off flight mass spectrometry (MALDI-ToF MS). In the MS analyses, ion doping and isotope labeling were crucial for detailed and conclusive characterization of the reaction products. The HILIC method is optimal for separation of native chitooligosaccharides and chitoaldonic acids that are both well retained on the column in acetonitrile-water. This is not the case for oxidized cello-oligosaccharides, which require increased ionic strength and pH of the elution buffer. Furthermore, while chitooligosaccharides absorb UV-light, cello-oligosaccharides do not. Thus, product analyses for cellulose-active LPMOs requires different methods for both chromatography and detection.

Although several methods have been established for both quantitative and qualitative assessment of oxidized cello-oligosaccharides, they are far from optimized and have several limitations. There is a main distinction between qualitative and quantitative methods and most often, a combination is needed for proper characterization of the product profile of an LPMO.

1.4.6.1 Qualitative methods

One of the main characteristics that distinguishes cellulose-degrading LPMOs is their oxidative regioselectivity. One complicating factor in distinguishing the two oxidized forms is that the lactone and the 4-ketone have identical masses in both their non-hydrated and hydrated forms. Nevertheless, it is possible to distinguish between the two forms by the use of MALDI-ToF MS. Sample preparation for MALDI-ToF MS leads to considerable dehydration of the gemdiol, so that C4-oxidized products are mainly found in the spectra as the 4-ketone form bound to a single ion (e.g. sodium). The charged aldonic acids, on the other hand, are not dehydrated during sample preparation, and tend to form double adducts with sodium (Westereng et al., 2017). Other types of MS, like electrospray ionization MS (ESI-MS) have been used for more in-depth studies of the composition of the produced oligosaccharides. Based on fragmentation patterns from MS/MS analyses, C1- and C4-oxidized species may be discriminated, as described in depth in paper I of this thesis. Precautions need to be made when analyzing MS-data, as the mass of several of the possible reaction products, and their adducts, have overlapping masses. For example, since the mass difference between a hydrated oxidized oligosaccharide and a native oligosaccharide equals the mass difference between sodium and potassium (m/z =16), the mass of the potassium

adduct of a native sugar is equal to the mass of the sodium adduct of its corresponding aldonic acid or gemdiol (Westereng et al., 2017).

NMR may be applied to confirm product identification. For example, this approach was followed to confirm that C4, and not C6, was the site of oxidation for non-reducing end oxidizing LPMOs, as described in Paper I of this thesis. Chromatographic methods, described in the next section, also give further insights into product profiles.

1.4.6.2 Quantitative methods

Quantitative analysis of native and oxidized products generated from LPMO action on cellulose may be achieved using several chromatographic methods, the most common being high-performance anion-exchange chromatography (HPAEC) with pulsed amperometric detection (PAD). HPAEC is particularly useful for analysis and quantification of C1-oxidized products since the alkaline conditions shift the equilibrium of the C1-oxidized products towards the stable and negatively charged aldonic acid form (Westereng et al., 2013). Thus, the aldonic acids have longer retention times compared to the native sugars and since both aldonic acid and native sugar standards are commercially available, quantification is relatively straight forward.

C4-oxidized products are also retained on the column and peaks that are characteristic for such products elute after the aldonic acids. C4-oxidized products are less stable in alkaline conditions and undergo on-column decomposition, yielding products with additional oxidations as well as native cello-oligosaccharides with one less sugar unit than the original C4-oxidized oligomer (Westereng et al., 2016). Nevertheless, quantification of C4-oxidized dimer has been done by exploiting the ability of some C4-oxidizing LPMOs to cleave short cello-oligosaccharides into two products, one native oligomer and an equal amount of a C4-oxidized oligomer (Müller et al., 2015). Using such LPMOs, standards can be prepared and quantification of the C4-oxidized product is achieved by quantifying the amount of the native degradation product. Recently, efforts have been made to develop an alternative method based on porous graphitized carbon (PGC) chromatography (Westereng et al., 2016). The main advantage of PGC is that the pH of the eluent can be lowered to hinder decomposition of the C4-oxidized products. Furthermore, the combination of a PGC set-up with charged aerosol detection (CAD) yields high sensitivity. When using low ion-strength eluents, the PGC-CAD method can even be coupled to MS-detection for discrimination of co-eluting species (Westereng et al., 2016). Due to the (too) high affinity of PGC-columns high affinity for cello-

oligosaccharides with a DP>5, HPAEC remains favorable for analysis of longer oligomeric products.

Although progress is being made, there are still limitations in the analytical methods used for LPMO characterization and obtaining good kinetic data remains a major challenge. Recently, LPMO kinetics have been assessed by using a dinitrosalicylic acid (DNS)-assay to measure formation of reducing ends by a C4-oxidizing LPMO (Agger et al., 2014; Paper II of this thesis). The activity of an LPMO acting on soluble substrates has been assessed using a double-labeled oligomeric substrate that allowed measurement of fluorescence resonance energy transfer (FRET) (Frandsen et al., 2016). Although both methods are relatively easy and fast, they have great limitations in their applicability, since they are limited in both the nature of the substrate used and the regioselectivity of the LPMO.

Kinetic analysis of LPMOs is complicated by the complexity of LPMO catalysis, which implies that the overall performance of an LPMO depends on several non-independent factors. One complication concerns that stability issues, meaning that obtaining linear progress curves is not always straightforward. Another issue concerns the multiple impacts of typical reaction parameters such as pH and temperature, which will not only affect the catalytic machinery of the LPMO, but also the redox properties of the reductant and the stability/reactivity of H_2O_2 (Westereng et al., 2017).

2 OUTLINE AND PURPOSE OF THE THESIS

Lignocellulosic biomass is a renewable resource that can substitute fossil resources for the production of fuels, chemicals, and materials. To harness this potential, industry needs to overcome the recalcitrant nature of the lignocellulose and its resistance to microbial degradation. Great efforts are being made to study natural lignocellulose-degrading enzyme systems. However, despite recent progress, the enzymatic conversion of lignocellulosic biomass remains inefficient and a major cost driver in industrial biomass processing. The discovery that LPMOs boost the hydrolytic degradation of cellulose by classical cellulases has made these enzymes a major focus of current research.

The overall goal of the research described in this thesis was to identify and characterize novel LPMOs with potential use in industrial applications for lignocellulose degradation. In particular, this research aimed for obtaining a better understanding of how LPMOs function to degrade their substrates. During growth on plant materials, increased expression level was observed for 11 of the 14 predicted LPMO9s in the genome of *N. crassa* (Tian et al., 2009), several of which had been cloned and characterized to various extents, at the start of the work described in this thesis. Based on its high expression during growth on lignocellulosic substrates and its rare ability to release short oxidized oligosaccharides upon cleavage of cellulose, one of these LPMOs, *Nc*LPMO9C, was chosen for in-depth characterization. *Nc*LPMO9C is a two-domain protein, with a catalytic AA9-domain, which is connected to a CBM1-domain via a threonine-rich linker.

In Paper I, the activity of *Nc*LPMO9C on cellulosic substrates is described. The most interesting result described in Paper I is the discovery of LPMO activity on soluble cello-oligosaccharides. This activity, which allowed the production of relatively simple product mixtures (with few components), enabled detailed studies of the substrate-degrading activity of LPMOs. In particular, MS- and NMR-based studies of products showed that oxidative cleavage of the glycosidic linkage by *Nc*LPMO9C is caused by hydroxylation at the C4-position, as opposed to the C6-position in the non-reducing end glucosyl, which also was considered a possible oxidation target at the time.

Considering the ability of *Nc*LPMO9C to cleave on soluble cello-oligosaccharides, the study described in Paper II aimed at investigating whether this LPMO was active on additional substrates. As a result LPMO activity on hemicellulose, in particular xyloglucan, was demonstrated for the first time. *Nc*LPMO9C is able to cleave the glycosidic linkages of

various hemicelluloses with a β-glucan backbone and can tolerate several, but not all, of the substitutions occurring in β-glucan hemicelluloses, both in the main chain itself (e.g. mannose units in glucomannan) and side-chain substations, as in xyloglucan. *Nc*LPMO9C is particularly active on xyloglucan and activity on naturally occurring xyloglucans from *Arabidopsis thaliana* and *Solanum lycopersicum* (tomato stem) was demonstrated.

Paper III describes the crystal structure of the catalytic domain of *Nc*LPMO9C as well as a more in-depth analyzes of the properties of this enzyme. Paper III also discusses attempts to gain more structural insight in substrate binding. Co-crystallization with cello-, xyloglucan-, or xylo-oligosaccharides was attempted, but unsuccessful. Computational docking studies were not able to reveal a specific binding of cellotetraose or xyloglucan fragments. Structural comparison of LPMO9s with known activity provided some insight into possible structural characteristics underlying the activity of *Nc*LPMO9C on xyloglucan and revealed that the oxidative regioselectivity of LPMO9s correlates with certain structural features near the copper-site. EPR data revealed, for the first time, that substrate binding affects the fine structure of the LPMO copper-site and isothermal titration calorimetry (ITC) was used to measure the binding constants for binding of xyloglucan, PASC and cellohexaose to *Nc*LPMO9C, with and without its CBM1-domain.

3 MAIN RESULTS AND DISCUSSION

3.1 A C4-oxidizing LPMO cleaving both cellulose and cello-oligosaccharides (Paper I)

At the time the work presented in this thesis was initiated, the discovery of LPMOs (active on chitin) was rather recent (Vaaje-Kolstad et al., 2010) and even more recent was the disclosure of LPMO activity on cellulose (Forsberg et al., 2011; Langston et al., 2011; Quinlan et al., 2011). Most of the LPMOs characterized at the time, including all reported bacterial LPMOs, released C1-oxidized products, i.e. aldonic acids, when acting on crystalline cellulosic substrates. Other oxidized products had been observed, but had not been identified due to analytical limitations. These alternative oxidized products were thought to be oxidized in the C4- or C6-carbon of the non-reducing end glucose residue, and their nature had been debated (Quinlan et al., 2011; Phillips et al., 2011; Bey et al., 2013).

The filamentous ascomycete *N. crassa* is a well-studied model organism that proficiently degrades plant cell walls. The genome of *N. crassa* encodes 14 predicted AA9s, four predicted AA11 and one AA13 (Table 2). Of the 14 predicted AA9s, 11 are upregulated together with several hydrolytic enzymes when the fungi is grown on lignocellulosic substrates (Tian et al., 2009). Only three of these 11 AA9s had been previously characterized by HPLC and MS analyzes of product mixtures. *Nc*LPMO9E (also called NCU08760) was reported to be strict C1-oxidizing, *Nc*LPMO9D (NCU01050) had been shown to release non-reducing end oxidized products, and *Nc*LPMO9M (NCU07898) was thought to oxidize both the reducing and the non-reducing end (Phillips et al., 2011; Li et al., 2012). Three additional *Nc*LPMO9s (*Nc*LPMO9C, *Nc*LPMO9F and *Nc*LPMO9J; NCU02916, NCU03328 and NCU01867 respectively) were shown to degrade cellulose in the presence of CDH as electron donor, but no attempts had been made to assess their regioselectivity and substrate preferences (Kittl et al., 2012).

When screening oxidative activity on various polymeric substrates (Avicel, PASC and steam-exploded spruce), using HPAEC for product analysis, *Nc*LPMO9C showed release of soluble oxidized cello-oligosaccharides with DP2 and DP3 (Figure 2A, Paper I). This product profile contrast with the product profiles of previously reported LPMOs, which predominantly generate longer oxidized cello-oligosaccharides (Quinlan et al., 2011; Forsberg et al., 2011; Westereng et al., 2011; Beeson et al., 2012). After ruling out possible contaminations by

glucanases and possible side effects from the chemical reducing agents used, we were able to show that *Nc*LPMO9C indeed is active on soluble oligosaccharides as short as DP4 (Figure 2B). The enzyme showed the highest activity on substrates containing five or more sugar units. Upon oxidation of cellopentaose, *Nc*LPMO9C almost exclusively released oxidized cellobiose and native cellotriose. Both oxidized cellobiose and oxidized cellotriose were generated from cellohexaose. Product profiles observed by HPAEC were confirmed by MS.

Since no aldonic acids were observed in the chromatograms we supposed that the oxidation occurred strictly in the non-reducing end. As described in Section 1.4.6 of this thesis, the identification of oxidized products by MS is difficult, since the masses of possible products overlap. However, the low DP of the oxidized products released by *Nc*LPMO9C, and the relatively simple nature of the generated product mixes, enabled us to use both MS/MS and NMR to evaluate their nature. To our knowledge, none of these methods had previously been used in product determination of oxidized cello-oligosaccharides and major efforts were made in method optimization for Paper I. Results from both MS/MS and NMR experiments clearly demonstrated that the non-reducing end oxidation was in the C4-position and not in the C6-position (Figure 4 and 6 in Paper I). Although enzymes exist that oxidize the C6-carbon of carbohydrates, such as D-galactose oxidase (Avigad et al., 1962), a C6-oxidation would not cleave the glycosidic linkage and thus not release oligosaccharide products.

To further study the nature of the products, a time course of a reaction containing *Nc*LPMO9C, a chemical reducing agent and cellopentaose in isotope labeled water ($H_2{}^{18}O$) was studied. The results revealed that the oxidized product primarily exists as a gemdiol (i.e. incorporation of two oxygen atoms) in solution, rather than the keto form. Thus, as a result of LPMO action on the glycosidic linkage, the C4-oxidized product acquires one oxygen atom from oxygen and one from water, as indicated in Figure 3B in Paper I. One interesting result, the implications of which we did not fully understand at the time, is shown in Figure 7 of Paper I. The figure shows measured H_2O_2 levels in reaction mixtures containing *Nc*LPMO9C, a reductant (ascorbic acid) and various soluble substrates. At the time, H_2O_2-release was thought to be part of a futile side reaction that occurred when the LPMO was not bound to a substrate (Kittl et al., 2012), and Figure 7 shows that, indeed, H_2O_2 accumulate when the LPMO is incubated with a substrate that it cannot degrade. Low or negligible amounts of H_2O_2, indicating productive binding to a substrate, correlated well with the relative activity of the LPMO on those substrates. While Figure 7 supports the overall conclusions of Paper I, the interpretation of this Figure needs some adjustment in light of the recent discovery that

H_2O_2, and not oxygen, is the co-substrate of LPMOs in hydroxylation of glycosidic linkage (Bissaro et al., 2017). In the presence of an electron donor, LPMOs are able to produce H_2O_2 from O_2, which is further utilized by the LPMO for its oxidative action. When the LPMO is bound to a substrate, perhaps reinforced by the presence of its co-substrate (Courtade et al., 2016; Frandsen et al., 2016), the produced H_2O_2 is consumed and the substrate is cleaved. Thus, the lack of H_2O_2 accumulation in reactions with substrates on which the LPMO can act (as shown in Figure 7 of Paper I), likely does not reflect a lack of H_2O_2 production, but rather consumption, by substrate-bound LPMOs, of H_2O_2 generated by non-substrate bound LPMOs.

Based on the crystal structure of another known non-reducing end oxidizing LPMO, *Nc*LPMO9D (Phillips et al., 2011; Li et al., 2012), a model was built for *Nc*LPMO9C. It was concluded that, compared to *Nc*LPMO9D and C1-oxidizing *Pc*LPMO9D (Wu et al., 2013; Westereng et al., 2011), which only have activity towards crystalline cellulose, *Nc*LPMO9C has less aromatic amino acids on its surface (Figure 8B in Paper I). Paper I further states that even if one considers the potential contribution from polar surface residues, the binding surface of *Nc*LPMO9C seems less extended and possibly more adapted to accommodate soluble substrates. In retrospect, describing the substrate-binding surface of *Nc*LPMO9C as "less extended" was not correct, as becomes evident when reading Papers II and III of this thesis, and it seems more likely that the special features of *Nc*LPMO9C relate to the balance between aromatic and polar residues on the surface.

Binding of polysaccharides by aromatic-ring stacking is well known and has also been proposed for LPMOs (Li et al., 2012; Wu et al., 2013). In particular, a conserved tyrosine (Tyr204 in *Nc*LPMO9C) in LPMO9s has been proposed to interact with cellulose (Wu et al., 2013) and section 1.4.5. Assuming that the non-reducing end glucan of a cellopentaose aligns on top of this tyrosine with the rest of the glucan chain extended across the copper-binding center, the reducing end sugar would be positioned near a cluster of three surface exposed asparagines located in L2 (Asn25-Asn26-Asn27 in *Nc*LPMO9C, Figure 8B Paper I). Such an orientation would be in accordance with the products generated from cellopentaose by *Nc*LPMO9C. Indeed, this orientation, i.e. with a reducing end glucose interacting with one of these asparagines, was observed in a structural study of C4-oxidizing *Ls*LPMO9A bound to cello-oligosaccharides of various lengths (Frandsen et al., 2016). The asparagine-cluster described for *Nc*LPMO9C is conserved in *Ls*LPMO9A (Asp27-Asp28-Asp29) and in the structure presented by Frandsen et al., the hydroxyl group of the C2-carbon of the reducing

end sugar interacts with one of these asparagines (Asn28 in *Ls*LPMO9A, which structurally aligns to Asn26 in *Nc*LPMO9C; Figure 11). Interestingly, mutation of a structurally equivalent asparagine in a C1/C4-oxidizing LPMO10 affected the C1:C4 ratio of oxidation (Forsberg et al., 2017). This latter result may be taken to confirm that this asparagine is important for binding and positioning of the substrate.

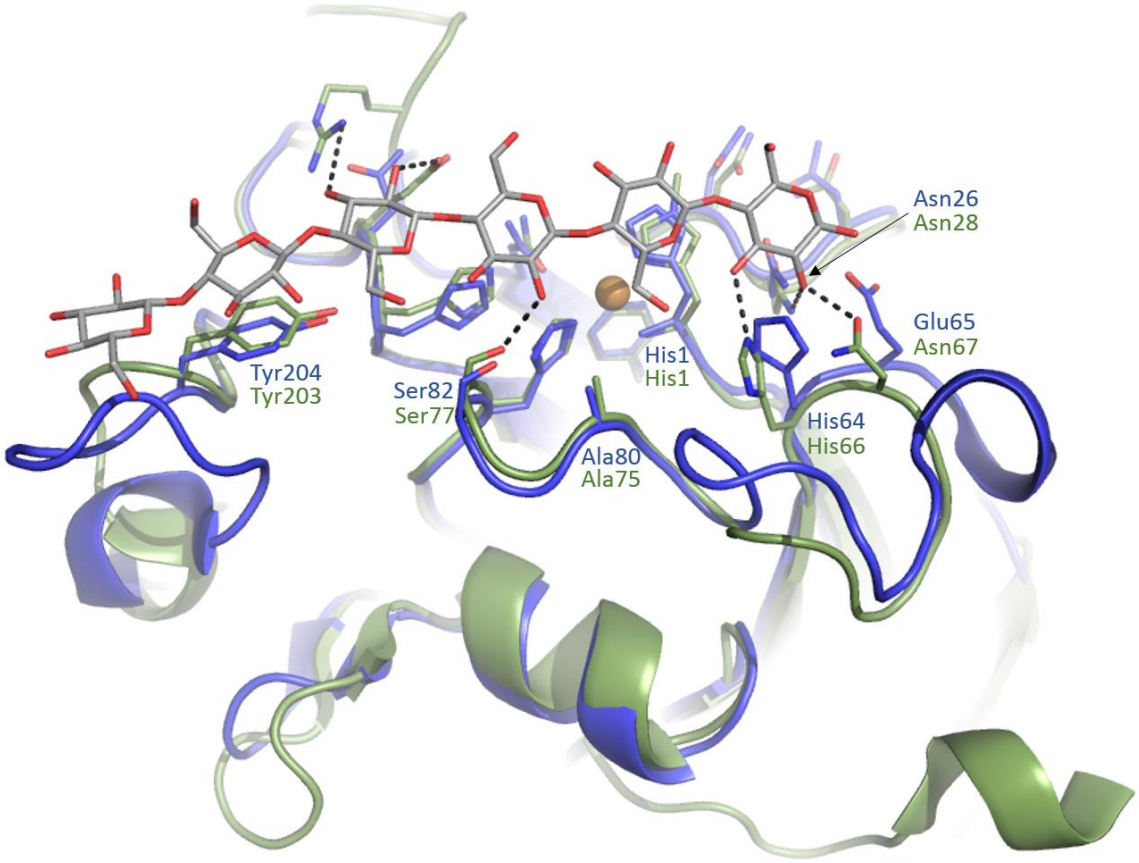

Figure 11. Structural comparison of *Nc*LPMO9C and *Ls*LPMO9A with a bound cellohexaose. Relevant amino acids are labelled and colored according to the structure; *Nc*LPMO9C (PDB code: 4D7U) is colored blue, *Ls*LPMO9A (PDB code: 5ACI) is colored green and the bound cellohexaose in the structure of *Ls*LPMO9A is colored light gray. Dashed lines indicate interactions between amino acids and cellohexaose, and the orange sphere indicates the copper ion in the structure of *Ls*LPMO9A.

3.2 Mapping the broader substrate specificity of *Nc*LPMO9C (Paper II)

During the work described in Paper I, the activity of *Nc*LPMO9C was tested on several other soluble substrates including oligosaccharides of mannan, xylan, chitin and maltodextrin, but no activity was detected. The use of glycan microarray substrate screening system enabled a broad search for possible activities of *Nc*LPMO9C on various polymeric substrates (Figure 1 in Paper II). Interestingly, and totally novel at the time, activity was detected on various hemicelluloses with β-(1,4)-linked glucan backbones, including xyloglucan, which is a β-(1,4)-linked glucan substituted with xylose and other sugars, as described in section 1.2.2.1. Other substrates that were found to be cleaved by *Nc*LPMO9C were mixed-linked β-glucans of oat, barley and lichens, and glucomannan.

Because of the heteropolymeric nature of xyloglucan and its various substitution patterns, the identification of xyloglucan-derived products is more challenging than identification of products generated from cellulose. However, based on known xyloglucan substitution patterns and by using MS-analyses to compare *Nc*LPMO9C-generated product profile to product profiles generated by two endoglucanases, we managed to determine that *Nc*LPMO9C is able to handle a variety of substitutions as long as there is an unsubstituted backbone glucose unit for the LPMO to access. This became even clearer after analyzing the degradation of $XXXGXXXG^{OH}$ ($XG14^{OH}$, i.e. a reduced 14-mer of xyloglucan where six of the eight backbone glucose-units are substituted with xylose, indicated by X) by *Nc*LPMO9C. The generated product profiles showed that the enzyme primarily targets the non-substituted internal glucosyl unit that is oxidized in the non-reducing end, meaning that the only observed native product corresponded to XXX (Figure 3 in Paper II). Analyses of products generated from tamarind xyloglucan (Paper II, Figure 2) and from hemicellulose in cell wall material from *Arabidopsis thaliana* and tomato stem (*Solanum lycopersicum*) (Figure 4 in Paper II) confirmed the capability of this LPMO to accept various substitution patterns and its preference for cleaving at an unsubstituted glucose.

Products generated from konjac glucomannan, a linear β-(1,4)-linked mannan interspersed with randomly distributed glucose, generated characteristic product clusters containing the keto-, gemdiol and native form of hexose oligosaccharides, as well as acetylated species (Paper II; Fig. 2B). Since products were rather long, it seems that *Nc*LPMO9C cannot cleave at any position in the chain and this is likely due to the occurrence of consecutive mannoses, since Figure 1 shows that galactomannan is not cleaved by this enzyme. Products generated

from barley β-glucan showed dominance of products increasing by three sugars in size (DP4, DP7, DP10; Paper II, Figure 2C). This product profile fits well with the repetitive structure of barley β-glucan, in which every third linkage is a β-(1,3)-linkage (Ajithkumar et al., 2006). Products generated from lichenan did not show this repetitive pattern, which may be taken to indicate that the structure of lichenan (i.e. β-glucan from moss) is less repetitive.

All in all, the data presented in Figures 2-4 of Paper II show that activity of *Nc*LPMO9C depends on the presence of a β-(1,4)-linked glucose in the substrate backbone and that the enzyme is quite flexible when it comes to accepting substitutions near this single glucose. Importantly, Paper II also shows a comparison of the LPMO-generated product profiles with the product profiles generated by endo-glucanases acting on the same substrates. These comparisons show clear differences, suggesting that the two enzyme types have different, and perhaps complementary, or even synergistic, roles in Nature.

In attempt to assess what would be the preferred substrates of *Nc*LPMO9C and the natural function of the enzyme, we also assessed the initial degradation rates for cellopentaose, $XG14^{OH}$, polymeric xyloglucan and PASC. The results, shown in Supplementary Figure 5 of Paper II, indicated that the enzyme has a similar rate on all substrates, varying from 0.11 s^{-1} for the two polymeric substrates to 0.06 s^{-1} and 0.03 s^{-1} for $XG14^{OH}$ and cellopentaose, respectively. Importantly, it is likely that these rates reflect the rate of the generation of H_2O_2 in the reaction mixtures rather than the intrinsic catalytic rate of the LPMO (Bissaro et al., 2017). It would be truly interesting to re-determine these rates in an experimental set-up where the LPMO reaction is driven by externally provided H_2O_2.

As described in section 1.2.2 of this thesis, a main function of hemicellulose in plant cell walls is to strengthen the cell wall by interlinking cellulose microfibrils and, in the case of xyloglucan, even weaving into the cellulose microfibrils (Hayashi, 1989). It has also been proposed that coating of cellulose by xyloglucan or xylan could provide protection against cell wall deconstruction by microorganisms (Vincken et al., 1995). While hemicelluloses are generally considered to be more easily degradable than cellulose, this may change if they interact strongly with a cellulose fibril. Furthermore, the varying substitution patterns of xyloglucan may be a challenge for xyloglucanases, i.e. hydrolytic enzymes acting on xyloglucan. Interestingly, as outlined below, some xyloglucan-active LPMOs described in the literature after the discovery of the xyloglucan activity of *Nc*LPMO9C, are even less sensitive to substitutons (Kojima et al., 2016; Nekiunaite et al., 2016). The genomes of some biomass

degrading microorganism contain more than 40 LPMOs (Horn et al., 2012; Lenfant et al., 2017), which are upregulated to various extents in response to growth on various substrates (Tian et al., 2009). It is likely that certain LPMOs may be targeted to certain recalcitrant substructures in plant cell walls. Interestingly, expression levels of the gene encoding *Nc*LPMO9C are five times higher when the *N. crassa* is grown on (xyloglucan-containing) *Miscanthus* stems compared to pure crystalline cellulose (Avicel) (Tian et al., 2009).

The study described in Paper II has contributed to a broadening of LPMO research, including also activity on soluble hemicellulosic substrates. Following the discovery of activity on xyloglucan, LPMO activity towards other substrates has been described, in particular xylan (Frommhagen et al., 2015) and starch (Lo Leggio et al., 2015; Vu et al., 2014b). These discoveries are not only important for our understanding of the complex enzymatic systems utilized by microorganisms for biomass degradation, but are also potentially of great importance for industrial applications in enzymatic saccharification of biomasses. One of the limiting factors of enzymatic degradation of lignocellulose is the insufficient removal of hemicelluloses (Saha et al., 2013), and perhaps certain hemicellulose-active LPMOs contribute to overall process efficiency. Notably, whereas cellulose is homogeneous and relatively similar in all types of plants, hemicelluloses vary a lot in composition and amount, and optimal degradation of hemicellulose could thus require substrate-specific enzyme mixtures. In this regard, the ability of *Nc*LPMO9C to tolerate several substitutions of its β-glucan substrate could make this enzyme particularly attractive for industrial applications.

Subsequent to the work described in Paper II, an LPMO from *Podospora anserina* (*Pa*LPMO9H) was described, showing activity towards both soluble cello-oligosaccharides with DP>4 and β-(1,4)-linked hemicellulose polysaccharides like xyloglucan, lichenan and glucomannan (Bennati-Granier et al., 2015). When cleaving cello-oligosaccharides, *Pa*LPMO9H generated both C1- and C4-oxidized products, however, the product profiles upon cleavage of xylogucan was inconclusive in regard of regioselectivity. Regardless of the regioselectivity, the analogous substrate specificity to *Nc*LPMO9C enabled a sequence comparison. The three asparagine residues (Asn25-Asn26-Asn27) in *Nc*LPMO9C that were discussed above (Paper I) are substituted with Ser25-Asn26-Phe27 in *Pa*LPMO9H. Thus, only the more generally conserved Asn26 is, indeed, conserved. Bennati-Granier et al. discuss whether the presence of a phenylalanine at position 27, possibly binding stronger to a sugar compared to an asparagine, could explain the higher activity of *Pa*LPMO9H on cellotetraose compared to *Nc*LPMO9C (Bennati-Granier et al., 2015).

Interestingly, *Nc*LPMO9A (NCU02240) and *Nc*LPMO9D (NCU01050), two C4-oxidizing LPMOs for which we could not detect activity on xyloglucan, contains the sequence Tyr25-Asp26-Gly27 and Tyr25-Asn26-Gly27, respectively. The structures of *Nc*LPMO9A and *Nc*LPMO9D indicate that mutation of Asn27 to glycine is needed to avoid steric hindrance of the tyrosine at position 25. We have generated and produced the N25Y-N27G mutant of *Nc*LPMO9C and a preliminary characterization of the double mutant showed wild-type like characteristics in terms of substrate preference and activity towards shorter oligosaccharides (Isaksen & Eijsink, unpublished observations).

3.3 Further biochemical characterization and crystal structure of *Nc*LPMO9C (Paper III)

The research described in Paper III aimed at gaining further insights into the substrate binding and catalytical properties of *Nc*LPMO9C. The activity of *Nc*LPMO9C on soluble substrates encouraged us in trying to obtain a crystal structure of an LPMO with a bound substrate. *Nc*LPMO9C-N, a truncated-variant of *Nc*LPMO9C comprising only the N-terminal AA9-domain and lacking the inter-domain linker and the CBM1-domain, was successfully expressed and purified. The truncated enzyme remained its catalytic activity towards soluble cello-oligosaccharides, crystalline cellulose and xyloglucan. Numerous attempts were made to co-crystallize *Nc*LPMO9C-N or soak crystals of *Nc*LPMO9C-N with a variety of oligomeric substrates, such as native and thio-linked cello- and xylo-oligosaccharides, xylogluco-oligosaccharides, β-glucan and low molecular lignin. Unfortunately, these attempts did not yield any crystal with bound substrate. Computational docking studies did neither reveal specific binding of cellotetraose or xyloglucan fragments at any point along the protein surface.

The crystal structure of the catalytic domain of *Nc*LPMO9C resembles that of other known LPMOs (described in section 1.4.3), showing the typical β-sandwich fold and a flat substrate-binding surface with a copper binding site. The copper ion is coordinated by the histidine brace, which includes the N-terminal amino group, the Nδ of His1 and the Nε of His83. Figure 1 in Paper III shows two structures of *Nc*LPMO9C-N, one with one copper bound in the active site (a-c) and one structure obtained after soaking with zinc that shows two additional bound zinc ions next to a zinc ion bound in the copper site (d). Although it is well established that LPMO activity depends on a single bound copper and although ITC experiments yielded data

that fitted well to a single-site metal-binding model, one may still speculate whether the two extra metals bound in the zinc-soaked structure could be of importance. The residues involved in binding the additional zinc ions are part of an insertion, termed L3 (see section 1.4.3). These metal sites seem too far away from the copper site to affect the catalytic action of the LPMO, however, a possible role in enzyme-substrate interactions cannot be excluded. Recently, an NMR study of substrate-binding revealed that one of the zinc-binding residues (His64) in L3 is among the residues that are most affected by binding of either cellohexaose or xyloglucan to *Nc*LPMO9C (Courtade et al., 2016). In the structure of *Ls*LPMO9A with cello-oligosaccharide bound, His66 (structurally aligning to His 64 in *Nc*LPMO9C) is one of three amino acids interacting directly with the reducing end sugar (Figure 11). These observations do by no means establish a functional role for the zinc ions seen in the structure of *Nc*LPMO9C but it is worth noting that the metal ions interact with residues involved in substrate-binding and could just modulate substrate affinity. Although the structural analysis of substrate binding was unsuccessful, ITC experiments allowed us to obtain an indication of the substrate binding affinities of LPMOs. As seen in Figure 6 in Paper III, full-length *Nc*LPMO9C displayed two binding modes for binding to PASC with K_d values of 0.013 ± 0.004 µM and 0.54 ± 0.16 µM. In the truncated form, *Nc*LPMO9C-N, only the lower affinity binding-mode was retained, clearly showing that the CBM1 is important for high affinity binding of cellulose, and perhaps providing an additional explanation for the lack of success of co-crystallization studies, which were done with a smaller substrate and the truncated protein.

Although only one binding mode was observed for binding of full length *Nc*LPMO9C to polymeric xyloglucan, removal of the CBM1 reduced the affinity for xyloglucan from a K_d of 2.3 ± 0.5 µM to a K_d of approximately 24 µM. Notably, the K_d values derived from the ITC experiments indicate that binding to xyloglucan is much weaker than binding to cellulose. Still, the affinity for xyloglucan was much higher than the affinity for cellohexaose for which the K_d was 0.81 ± 0.08 mM. The differences in substrate affinity were hardly reflected in the catalytic activity. Enzyme rates towards xyloglucan and PASC were in the order of 0.04 - 0.1 /s, i.e. very similar to the rates described in Paper II, and almost substrate-independent. This observation supports the suggestion made above, in the discussion of Paper II that in the reaction setup used here; substrate binding is not a rate-limiting factor. Clearly, rate determinations should be redone bearing in mind, and implementing, the recent findings concerning the role of H_2O_2 in LPMO catalysis (Bissaro et al., 2017)

Paper III also describes the copper affinity and the redox potential of *Nc*LPMO9C, and both resemble that of other LPMOs (Forsberg et al., 2014a; Aachmann et al., 2012; Hemsworth et al., 2013a). The EPR spectrum of the Cu^{2+}-loaded enzyme resembled that of other cellulose-oxidizing LPMO9s and LPMO10s (Quinlan et al., 2011; Hemsworth et al., 2013b; Forsberg et al., 2014a, 2014b), being indicative of a mononuclear type 2-copper site. Importantly, Paper III provides the first experimental insight into the effect of substrate binding on the copper site. When adding the soluble substrates cellohexaose or xyloglucan to Cu^{2+}-loaded *Nc*LPMO9C a clear change was observed in the EPR spectra. The changes were similar for both substrates, which suggests that the copper-site interacts with the common β-glucan part of the substrates, and that these interactions induce the same conformational changes at the Cu^{2+} site. In a subsequent study, Frandsen et al. showed similar EPR experimental data (Frandsen et al., 2016).

A comparison among the seven LPMO9 structures available at the time, revealed differences in the accessibility of the solvent-facing axial position that correlate with the regioselectivity of the oxidation reaction (Figure 3 c and d in Paper III and Figure 12). In the case of the two C1-oxidizers, access to the axial-position is blocked by the hydroxyl group of a tyrosine. As discussed previously and shown in Figure 10 of this thesis, in C1-oxidizing LPMO10s an alanine is occupying this position (Forsberg et al., 2014a; Hemsworth et al., 2013b). In the strictly C4-oxidizing LPMO9s, access to this same position seems unrestricted whereas LPMO9s that oxidize C1 and C4 show an intermediate situation with a proline in this position (Figure 3a and d in Paper III and Figure 12). Originally, we speculated whether this variation in the copper coordination sphere could affect either the binding and positioning of oxygen or the positioning of the substrate itself (relative to the copper site). In both situations, one could envisage consequences for the oxidative regioselectivity of the enzyme. Recent structural data suggest that oxygen binds in an equatorial position (as opposed to the solvent-exposed axial position) (Bacik et al., 2017) and that the substrate occupies the solvent-exposed axial position upon binding, probably not leaving room for oxygen (Frandsen et al., 2016). Combined, these recent observations suggest that, if the correlation described above has any impact on the oxidative regioselectivity, this is probably due to an effect on the positioning of the substrate relative to the copper site. Interestingly, in the NMR study of substrate binding by Courtade et al., Ala80, in the solvent-exposed axial coordination position (Figure 12) is one of the amino acids that is affected when *Nc*LPMO9C is incubated with xyloglucan or cellohexaose (Courtade et al., 2016).

Towards the end of the experimental work described in this thesis and inspired by the correlation described above, we set out to explore the structural basis of the oxidative regioselectivity of LPMOs. In addition to *Nc*LPMO9C-N, two LPMO9s from *Myriococcum thermophilum* were included in a mutation study: *Mt*4260-N, the AA9-domain of a two-domain homologue of *Nc*LPMO9C, which also is a strict C4-oxidizer and active on both soluble cello-oligosaccharides and hemicelluloses, and *Mt*358, a C1-oxidizing cellulose-active LPMO sharing high sequence similarity (76%) with *Nc*LPMO9F (NCU03328) which was shown by Vu et al. to be a C1-oxidizer(Vu et al., 2014a). Using *Pichia pastoris* for expression, two mutants of *Nc*LPMO9C-N were generated in which Ala80 (in the solvent-exposed axial coordination position; Figure 12) was mutated to proline (A80P) or to aspartate (A80D). Both mutants were obtained, however the A80P mutant did not express well. The A80D mutant was expressed and purified and preliminary characterization of its properties indicated decreased activity on cello-oligosaccharides, PASC and xyloglucan. However, the product profile and the strict C4-oxidizing activity were unaltered. For *Mt*4260, we succeeded in expressing the A80P-mutant, however in low amounts, and the purified enzyme did not show altered or reduced enzyme activity. In *Mt*358, the tyrosine potentially blocking the axial site was mutated either to a phenylalanine (Y67F) or to an alanine (Y67A). Neither mutation altered the enzyme's regioselectivity or substrate preference. However, the latter mutant showed reduced activity on PASC, while the activity of the former mutant seemed unaffected.

The preliminary data of the mutagenesis studies described above raise questions as to the relevance of the correlation between the solvent-exposed axial coordination site of the copper and oxidative regioselectivity. Considering the number of protruding amino acids on the LPMO substrate, it is possible that several mutations are needed to alter enzyme properties. While we did not succeed in altering the oxidative regioselectivity, some of the mutations affected enzyme activity, as one would expect considering the involvement of the mutated residue in interactions with the substrate, as suggested by the structure of the enzyme-substrate complex in Frandsen et al. (2016). It is important to note that, in light of recent findings on LPMO activity and stability (Bissaro et al., 2017), as discussed above, it would be worth reassessing the catalytic properties of these mutants under conditions that yield stable LPMO kinetics and avoid enzyme inactivation. For example, it is not clear whether the mutational effects observed in the preliminary characterization of the mutants reflect true changes in enzyme activity, or rather changes in enzyme stability.

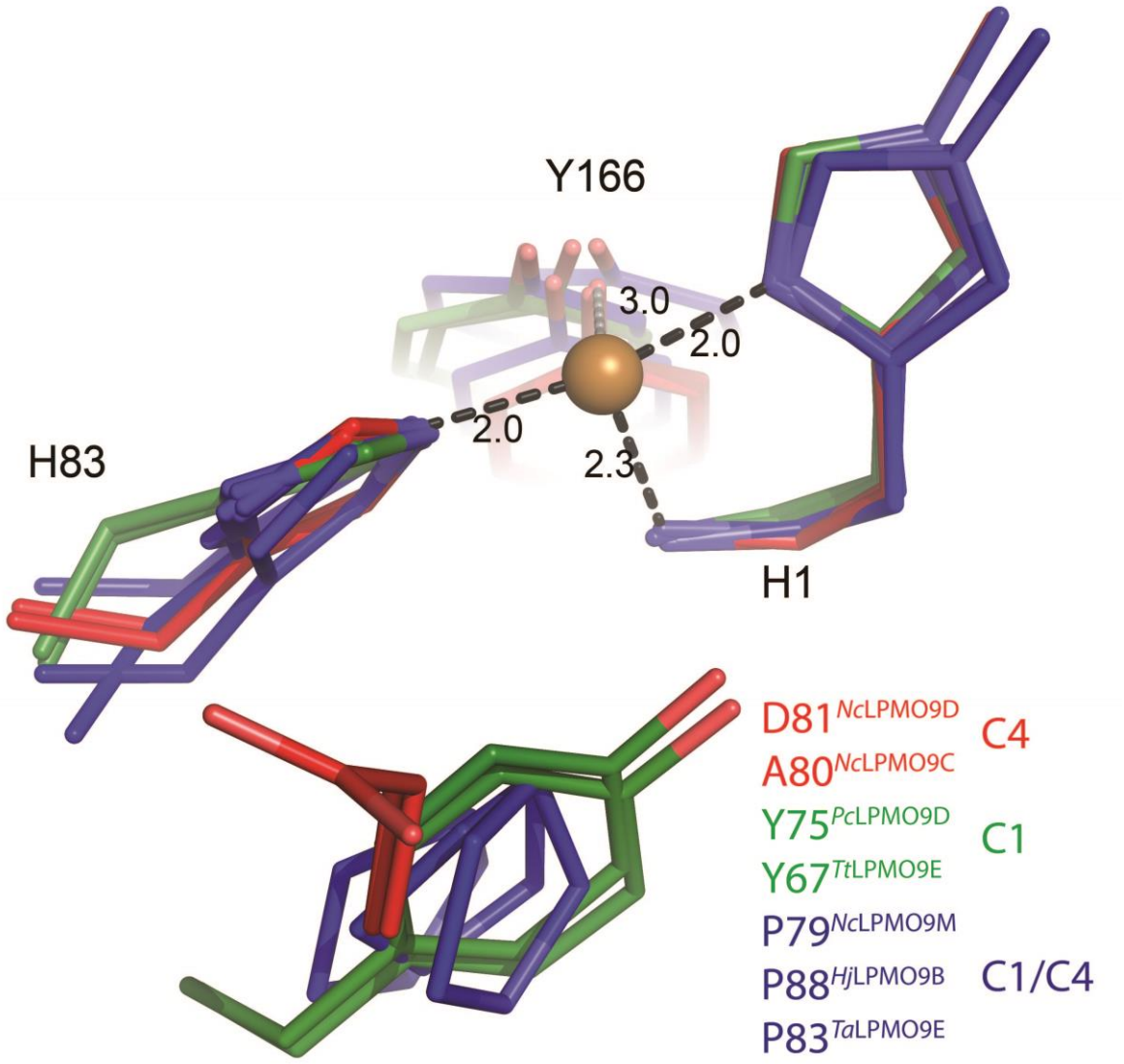

Figure 12. Possible determinants of regioselectivity among LPMOs. Superposition of the active sites of seven known LPMO9s with known structure and activity; C4, C1 and C1/C4 oxidizers are colored red, green and blue, respectively. Residue numbers and distances in black refers to *Nc*LPMO9C.

4 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The discovery of the oxidative degradation of polysaccharides by LPMOs has added an extra dimension to our understanding of how microorganisms degrade plant biomass in Nature. LPMOs boost the activity of conventional cellulases and their potential as additives in enzymatic mixtures used for industrial biomass processing is an attractive field of research.

The work presented in this thesis has given novel insight into the substrate preferences and catalytic properties of LPMOs. Paper I and II describe the broad substrate specificity of *Nc*LPMO9C and are the first papers in the field to describe the LPMO activity on soluble cello-oligosaccharides and on a variety of hemicelluloses with a β-glucan backbone. At the time, there was still some discussion about the oxidative regioselectivity of non-C1-oxidizing LPMOs and this issue was settled in Paper I, showing that *Nc*LPMO9C acts on the C4-position. Notably, the results described in Paper I also helped in the interpretation of product profiles that are typically obtained when using HPAEC for product analysis. Although *Nc*LPMO9C is able to cleave various mixed-linked β-glucans and glucomannan, the data shown in Paper II indicate that this LPMO is particularly active on xyloglucan. It was shown that the enzyme could handle a variety of substitutions on the xyloglucan backbone, as long as there is one unsubstituted backbone glucose unit for the LPMO to access.

Subsequent to the study described in Paper II, several other xyloglucan-active LPMOs have been described (Kojima et al., 2016; Nekiunaite et al., 2016; Jagadeeswaran et al., 2016; Bennati-Granier et al., 2015), including enzymes that, in contrast to *Nc*LPMO9C, seem capable of cleaving *any* position in the β-glucan backbone and that thus seem more versatile than known xyloglucanases. Interestingly, these more versatile LPMOs showed C1/C4-oxidizing activity on cellulose, and in one case, both C1- and C4-oxidized products were observed when cleaving the β-glucan backbone of xyloglucan (Nekiunaite et al., 2016). Sequence comparisons indicate that these more versatile LPMOs do not have the L3-insertion, but rather an extended L2-region (Kojima et al., 2016), which is common for LPMOs with C1/C4-oxidizing activity on cellulose (see Figure 6A, and clade 8 in Figure 7) (Vaaje-Kolstad et al., 2017). Recently, Fanuel et al. published a detailed study of the regioselectivity of *Pa*LPMO9H, indicating that also this LPMO generates both C1 and C4 oxidized products when cleaving the β-(1,4)-glucan backbone of glucomannan and xyloglucan (Fanuel et al., 2017). In contrast to the other xyloglucan active LPMOs with a C1/C4-regioselectivity, *Pa*LPMO9H does have the L3 insertion and lacks the extension of L2. All in all, while NMR

studies have shown that the L2 and L3 regions are affected by binding of xyloglucan-substrates (Courtade et al., 2016), and despite the structural information presented in Paper III and in Frandsen et al., (2016), the experimental data accumulated so far do not provide a structural explanation for variations in substrate specificity. As to the ability to cleave xyloglucan, the L3-insertion discussed in Paper III does not seem to be a prerequisite for xyloglucan activity. Furthermore, the structural determinants of the ability to cleave xyloglucan at any backbone position remain unknown.

The differences in oxidative regioselectivity and substrate preference among LPMOs may explain the multiplicity of LPMO-encoding genes in the genomes of biomass degrading fungi. It is conceivable that further variation in specificities will be discovered in the future. For example, one could imagine that certain LPMOs act on certain substructures in the plant cell wall, for example regions where hemicellulose and cellulose are entangled in specific manners. Considering the various origins of known LPMOs, their observed and possible expected broad range of substrate preferences and the varying oxidative regioselectivities, it is remarkable that the copper binding sites and, by inference, most likely, the reaction mechanism seem highly conserved. The EPR data presented in Paper III and in (Frandsen et al., 2016) indicate that the changes in the copper coordination sphere upon substrate-binding are independent on the nature of the substrate. However, in the final stage of preparing this thesis (October 2017), a study describing the activity of *Ls*LPMO9A on xylan was published (Simmons et al., 2017), where, based on EPR-analyses, the authors claim that binding of xylan affects the active-site copper in a different way than binding of a cello-oligoaccharide. The observed differences could be due to the absence of a C6 in the xylan backbone, since structures have shown (Frandsen et al., 2016) that one of the C6 in a glucan backbone comes very close to the copper site upon substrate binding.

Although there is a correlation between access to the solvent-facing axial copper coordination site and the oxidative regioselectivity of LPMOs (Figure 10 and 12), as discussed by Forsberg et al. (2014) and in Paper III of this thesis, preliminary site-directed mutagenesis experiments did not lead to mutants with an observable change in regioselectivity (unpublished results described in section 3.3). This is somewhat surprising, since structural data (Frandsen et al., 2016) indicate that mutations at this position (Ala75 in *Ls*LPMO9A, Figure 11) to a larger residue, such as e.g. proline, would affect the LPMO-substrate interaction and likely affect how the scissile glycosidic bond is oriented vis-à-vis the copper site.

In light of the recent findings on the role of H_2O_2 and self-inactivation of LPMOs, it would be interesting to repeat the characterization of the mutants and, in general, all kinetic experiments described in Papers II and III. As described by Bissaro et al. (2017), full and stable activity of LPMOs in action depend on controlled feeding or controlled generation of H_2O_2 in appropriate amounts. Too little H_2O_2 in the reaction will hamper LPMO action, whereas too high concentrations of H_2O_2 will lead to irreversible enzyme inactivation (Bissaro et al., 2017). Interestingly, Supplementary Figure 15 from Bissaro et al. (2017) indicates that productive binding of a substrate protects the LPMO from inactivation from accumulated H_2O_2 in the reaction. As indicated in the preceding paragraph, it is likely that mutation of the solvent exposed copper coordination residue, i.e. Ala80 in *Nc*LPMO9C (Figure 11 and 12), leads to weakened substrate binding, since this residue is known to interact with the bound substrate (Courtade et al., 2016). Based on the results of Bissaro et al. (2017), lowered affinity for the substrate would affect the LPMOs stability in the reaction. Furthermore, as mentioned in section 3.2, quantification of LPMO activity should preferably be done in conditions with controlled H_2O_2 supply. In the conditions that have routinely been used so far, including kinetic experiments in Paper II and III, the rate-limiting factor has likely been the generation of the co-substrate, H_2O_2, meaning that the experiments did not provide a true impression of LPMO catalytic activity.

The recalcitrant nature and crystalline packing of cellulose have long been considered the main bottlenecks in efficient saccharification of lignocellulosic biomass. However, in lignocellulose, co-polymeric structures of cellulose and hemicelluloses like xylan, glucomannan or xyloglucan, may create additional barriers for enzymatic degradation. Studies of the complex interactions between cellulose and hemicellulose are intricate, because of the lack of well-characterized models of cellulose microfibrils and the many possible side chain substitutions of hemicelluloses (Pereira et al., 2017; Park & Cosgrove, 2015). Interestingly, recent studies have shown that while some LPMOs, including *Nc*LPMO9C, preferentially act on free xyloglucan, compared to cellulose-bound xyloglucan (Kojima et al., 2016) others act equally well on both forms (Nekiunaite et al., 2016). In both cases, the cellulose-oxidizing activity of the LPMO was hampered by mixing xyloglucan with cellulose, indicating that the LPMOs are active on the cellulose-surface that the xyloglucan attaches to. Such interaction sites between cellulose and hemicellulose may be difficult to access for cellulases (Vincken et al., 1994) while LPMOs may be more suited to target such sites. Despite the discovery of a series of hemicellulose-active LPMOs, it remains to be studied if

and how the hemicellulose-cleaving abilities of these enzymes can be exploited in biomass processing.

In conclusion, the work described in this thesis has contributed to increasing our understanding of LPMOs and their potential. However, despite much progress and considerable research efforts worldwide, important questions concerning LPMO functionality remain unanswered and need further research. Important topics for future research include the development of proper kinetic assays, unravelling the structural basis of substrate specificity and enzyme efficiency, and assessing the roles of a wide variety of LPMOs in Nature and industrial processing of lignocellulosic biomass.

5 REFERENCES

Aachmann, F.L., Sørlie, M., Skjåk-Bræk, G., Eijsink, V.G.H. & Vaaje-Kolstad, G. (2012). NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 109 (46). pp. 18779–18784.

Agger, J.W., Isaksen, T., Várnai, A., Vidal-Melgosa, S., Willats, W.G.T., Ludwig, R., Horn, S.J., Eijsink, V.G.H. & Westereng, B. (2014). Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (17). pp. 6287–6292.

Ajithkumar, A., Andersson, R., Siika-aho, M., Tenkanen, M. & Åman, P. (2006). Isolation of cellotriosyl blocks from barley β-glucan with endo-1,4-β-glucanase from Trichoderma reesei. *Carbohydrate Polymers*. 64 (2). pp. 233–238.

Alonso, D.M., Bond, J.Q. & Dumesic, J.A. (2010). Catalytic conversion of biomass to biofuels. *Green Chemistry*. 12. pp. 1493–1513.

Alonso, D.M., Hakim, S.H., Zhou, S., Won, W., Hosseinaei, O., Tao, J., Garcia-Negron, V., Motagamwala, A.H., Mellmer, M. a, Huang, K., Houtman, C.J., Labbé, N., Harper, D.P., Maravelias, C., Runge, T. & Dumesic, J. a (2017). Increasing the revenue from lignocellulosic biomass: Maximizing feedstock utilization. *Science Advances*. 3 (5). pp. 1–7.

Arantes, V. & Goodell, B. (2014). *Current Understanding of Brown-Rot Fungal Biodegradation Mechanisms: A Review*. [Online].

Aro, E.-M. (2016). From first generation biofuels to advanced solar biofuels. *Ambio*. 45. pp. 24–31.

Avigad, G., Amaral, D., Asensio, C. & Horecker, B.L. (1962). The D-galactose oxidase of Polyporus circinatus. *The Journal of Biological Chemistry*. 237 (9). pp. 2736–2743.

Bacik, J.-P., Mekasha, S., Forsberg, Z., Kovalevsky, A.Y., Vaaje-Kolstad, G., Eijsink, V.G.H., Nix, J.C., Coates, L., Cuneo, M.J., Unkefer, C.J. & Chen, J.C.-H. (2017). Neutron and atomic resolution X-ray structures of a lytic polysaccharide monooxygenase reveal copper-mediated dioxygen binding and evidence for N-terminal deprotonation. *Biochemistry*. 56. pp. 2529–2532.

Bayer, E.A., Belaich, J.-P., Shoham, Y. & Lamed, R. (2004). The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology*. 58 (1). pp. 521–554.

Beckham, G.T., Matthews, J.F., Peters, B., Bomble, Y.J., Himmel, M.E. & Crowley, M.F. (2011). Molecular-level origins of biomass recalcitrance: Decrystallization free energies for four common cellulose polymorphs. *Journal of Physical Chemistry B*. 115 (14). pp. 4118–4127.

Beeson, W.T., Phillips, C.M., Cate, J.H.D. & Marletta, M.A. (2012). Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *Journal of the American Chemical Society*. 134 (2). pp. 890–892.

Beeson, W.T., Vu, V. V, Span, E.A., Phillips, C.M. & Marletta, M.A. (2015). Cellulose degradation by polysaccharide monooxygenases. *Annual Review of Biochemistry*. 84. pp. 923–946.

Bennati-Granier, C., Garajova, S., Champion, C., Grisel, S., Haon, M. & Zhou, S. (2015). Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by Podospora anserina. *Biotechnology for Biofuels*. 8. pp. 1–14.

Berka, R.M., Grigoriev, I. V, Otillar, R., Salamov, A., Grimwood, J., Reid, I., Ishmael, N., John, T., Darmond, C., Moisan, M.-C., Henrissat, B., Coutinho, P.M., Lombard, V., Natvig, D.O., Lindquist, E., Schmutz, J., Lucas, S., Harris, P., Powlowski, J., Bellemare, A., Taylor, D., Butler, G., de Vries, R.P., Allijn, I.E., van den Brink, J., Ushinsky, S., Storms, R., Powell, A.J., Paulsen, I.T., Elbourne, L.D.H., Baker, S.E., Magnuson, J., Laboissiere, S., Clutterbuck, A.J., Martinez, D., Wogulis, M., de Leon, A.L., Rey, M.W. & Tsang, A. (2011). Comparative genomic analysis of the thermophilic biomass-degrading fungi Myceliophthora thermophila and Thielavia terrestris. *Nature Biotechnology*. 29 (10). pp. 922–927.

Bey, M., Zhou, S., Poidevin, L., Henrissat, B., Coutinho, P.M., Berrin, J.G. & Sigoillot, J.C. (2013). Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from Podospora anserina. *Applied and Environmental Microbiology*. 79. pp. 488–496.

Bissaro, B., Røhr, Å.K., Müller, G., Chylenski, P., Skaugen, M., Forsberg, Z., Horn, S.J., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2017). Oxidative cleavage of polysaccharides by monocopper enzymes depends on H2O2. *Nature Chemical Biology*. 13 (10). pp. 1123–1128.

Blanchette, R.A. (1991). Delignification by wood-decay fungi. *Annual Review of Phytopathology*. 29 (1). pp. 381–398.

Borisova, A.S., Isaksen, T., Dimarogona, M., Kognole, A.A., Mathiesen, G., Várnai, A., Røhr, Å.K., Payne, C.M., Sørlie, M., Sandgren, M. & Eijsink, V.G.H. (2015). Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *The Journal of Biological Chemistry*. 290 (38). pp. 22955–22969.

Brethauer, S. & Studer, M.H. (2015). Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals – a review. *CHIMIA International Journal for Chemistry*. 69 (10). pp. 572–581.

Busse-Wicher, M., Gomes, T.C.F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N.J., Bolam, D.N., Skaf, M.S. & Dupree, P. (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *The Plant Journal*. 79 (3). pp. 492–506.

Busse-Wicher, M., Grantham, N.J., Lyczakowski, J.J., Nikolovski, N. & Dupree, P. (2016). Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochemical Society Transactions*. 44 (1). pp. 74–78.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*. 37 (Database issue). pp. 233–238.

Cavalier, D.M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., Zabotina, O.A., Hahn, M.G., Burgert, I., Pauly, M., Raikhel, N. V. & Keegstra, K. (2008). Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell*. 20 (6). pp. 1519–1537.

Chaplin, A.K., Wilson, M.T., Hough, M.A., Svistunenko, D.A., Hemsworth, G.R., Walton,

P.H., Vijgenboom, E. & Worrall, J.A.R. (2016). Heterogeneity in the histidine-brace copper coordination sphere in auxiliary activity family 10 (AA10) lytic polysaccharide monooxygenases. *The Journal of Biological Chemistry*. 291 (24). pp. 12838–12850.

Chiu, E., Hijnen, M., Bunker, R.D., Boudes, M., Rajendran, C., Aizel, K., Oliéric, V., Schulze-Briese, C., Mitsuhashi, W., Young, V., Ward, V.K., Bergoin, M., Metcalf, P. & Coulibaly, F. (2015). Structural basis for the enhancement of virulence by viral spindles and their in vivo crystallization. *Proceedings of the National Academy of Sciences of the United States of America*. 112 (13). pp. 3973–3978.

Cosgrove, D.J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*. 6 (11). pp. 850–861.

Côtes, W.A. (1967). *Wood Ultrastructure- an atlas of electron micrographs*. University of Washington Press, Seattle.

Courtade, G., Wimmer, R., Røhr, Å.K., Preims, M., Felice, A.K.G., Dimarogona, M., Vaaje-Kolstad, G., Sørlie, M., Sandgren, M., Ludwig, R., Eijsink, V.G.H. & Aachmann, F.L. (2016). Interactions of a fungal lytic polysaccharide monooxygenase with β-glucan substrates and cellobiose dehydrogenase. *Proceedings of the National Academy of Sciences of the United States of America*. 113 (21). pp. 5922–5927.

Cragg, S.M., Beckham, G.T., Bruce, N.C., Bugg, T.D., Distel, D.L., Dupree, P., Etxabe, A.G., Goodell, B.S., Jellison, J., McGeehan, J.E., McQueen-Mason, S.J., Schnorr, K., Walton, P.H., Watts, J.E. & Zimmer, M. (2015). Lignocellulose degradation mechanisms across the Tree of Life. *Current Opinion in Chemical Biology*. 29. pp. 108–119.

Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., Cartmell, A., Rogowski, A., Hamilton, B.S., Chen, R., Tolbert, T.J., Piens, K., Bracke, D., Vervecken, W., Hakki, Z., Speciale, G., Munōz-Munōz, J.L., Day, A., Peña, M.J., McLean, R., Suits, M.D., Boraston, A.B., Atherly, T., Ziemer, C.J., Williams, S.J., Davies, G.J., Abbott, D.W., Martens, E.C. & Gilbert, H.J. (2015). Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature*. 517 (7533). pp. 165–169.

Danneels, B., Tanghe, M., Joosten, H.-J., Gundinger, T., Spadiut, O., Stals, I. & Desmet, T. (2017). A quantitative indicator diagram for lytic polysaccharide monooxygenases reveals the role of aromatic surface residues in HjLPMO9A regioselectivity J.-G. Berrin (ed.). *PLOS ONE*. 12 (5). pp. 1–15.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*. 3 (9). pp. 853–859.

Din, N., Gilkes, N.R., Tekant, B., Miller, R.C., Warren, R.A.J. & Kilburn, D.G. (1991). Non–hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Nature Biotechnology*. 9 (11). pp. 1096–1099.

Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J.K., Teeri, T.T. & Jones, T.A. (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from Trichoderma reesei. *Science*. 265 (5171). pp. 524–528.

Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., Asiegbu, F.O., Baker, S.E., Barry, K., Bendiksby, M., Blumentritt, M., Coutinho, P.M., Cullen, D., de Vries, R.P., Gathman, A., Goodell, B., Henrissat, B., Ihrmark, K., Kauserud, H., Kohler, A., LaButti, K., Lapidus, A., Lavin, J.L., Lee, Y.-H., Lindquist, E., Lilly, W., Lucas, S., Morin, E., Murat, C., Oguiza, J.A., Park, J., Pisabarro, A.G., Riley, R., Rosling, A., Salamov, A., Schmidt, O., Schmutz, J., Skrede, I., Stenlid, J., Wiebenga,

A., Xie, X., Kües, U., Hibbett, D.S., Hoffmeister, D., Högberg, N., Martin, F., Grigoriev, I. V & Watkinson, S.C. (2011). The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science*. 333 (6043). pp. 762–765.

Ebringerová, A. (2006). Structural diversity and application potential of hemicelluloses. *Macromolecular Symposia*. 232 (333). pp. 1–12.

Edwards, M., Dea, I.C.M., Bulpin, P. V & Reid, J.S.G. (1985). Xyloglucan (amyloid) mobilisation in the cotyledons of Tropaeolum majus L. seeds following germination. *Planta*. 163 (1). pp. 133–140.

Enkhbaatar, B., Temuujin, U., Lim, J.-H., Chi, W.-J., Chang, Y.-K. & Hong, S.-K. (2012). Identification and characterization of a xyloglucan-specific family 74 glycosyl hydrolase from Streptomyces coelicolor A3(2). *Applied and Environmental Microbiology*. 78 (2). pp. 607–611.

Eriksson, K.-E., Pettersson, B. & Westermark, U. (1974). Oxidation: an important enzyme reaction in fungal degradation of cellulose. *FEBS Letters*. 49 (2). pp. 282–285.

Fanuel, M., Garajova, S., Ropartz, D., McGregor, N., Brumer, H., Rogniaux, H. & Berrin, J.-G. (2017). The Podospora anserina lytic polysaccharide monooxygenase PaLPMO9H catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnology for Biofuels*. 10 (63). pp. 1–10.

Feng, T., Yan, K.-P., Mikkelsen, M.D., Meyer, A.S., Schols, H.A., Westereng, B. & Mikkelsen, J.D. (2014). Characterisation of a novel endo-xyloglucanase (XcXGHA) from Xanthomonas that accommodates a xylosyl-substituted glucose at subsite −1. *Applied Microbiology and Biotechnology*. 98 (23). pp. 9667–9679.

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R.A. & Henrissat, B. (2012). The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science*. 336. pp. 1715–1719.

Forsberg, Z., Gullesen, J., Dalhus, B., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2017). Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *Submitted for publication*.

Forsberg, Z., Mackenzie, a. K., Sørlie, M., Røhr, a. K., Helland, R., Arvai, a. S., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2014a). Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences*. 111 (23). pp. 8446–8451.

Forsberg, Z., Nelson, C.E., Dalhus, B., Mekasha, S., Loose, J.S.M., Crouch, L.I., Røhr, Å.K., Gardner, J.G., Eijsink, V.G.H. & Vaaje-Kolstad, G. (2016). Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in Cellvibrio japonicus. *The Journal of biological chemistry*. 291. pp. 7300–7312.

Forsberg, Z., Røhr, A.K., Mekasha, S., Andersson, K.K., Eijsink, V.G.H., Vaaje-Kolstad, G. & Sørlie, M. (2014b). Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry*. 53. pp. 1647–1656.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A.C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S.J. & Eijsink, V.G.H. (2011). Cleavage of cellulose by a CBM33 protein. *Protein Science*. 20 (9). pp. 1479–1483.

Frandsen, K.E.H. & Lo Leggio, L. (2016). Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes. *IUCrJ*. 3. pp. 448–467.

Frandsen, K.E.H., Poulsen, J.-C.N., Tandrup, T. & Lo Leggio, L. (2017). Unliganded and substrate bound structures of the cellooligosaccharide active lytic polysaccharide monooxygenase LsAA9A at low pH. *Carbohydrate Research*. 448. pp. 187–190.

Frandsen, K.E.H., Simmons, T.J., Dupree, P., Poulsen, J.N., Hemsworth, G.R., Ciano, L., Johnston, E.M., Tovborg, M., Johansen, K.S., von Freiesleben, P., Marmuse, L., Fort, S., Cottaz, S., Driguez, H., Henrissat, B., Lenfant, N., Tuna, F., Baldansuren, A., Davies, G.J., Lo Leggio, L. & Walton, P.H. (2016). The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nature Chemical Biology*. 12 (4). pp. 298–303.

Frommhagen, M., Sforza, S., Westphal, A.H., Visser, J., Hinz, S.W.A., Koetsier, M.J., van Berkel, W.J.H., Gruppen, H. & Kabel, M.A. (2015). Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnology for Biofuels*. 8 (101). pp. 1–12.

Fry, S.C., York, W.S., Albersheim, P., Darvill, A., Hayashi, T., Joseleau, J.-P., Kato, Y., Lorences, E.P., Maclachlan, G.A., McNeil, M., Mort, A.J., Grant Reid, J.S., Seitz, H.U., Selvendran, R.R., Voragen, A.G.J. & White, A.R. (1993). An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum*. 89 (1). pp. 1–3.

Fujita, M. & Harada, H. (2001). Ultrastructure and formation of wood cell wall. In: D. N. S. Hon & N. Shiraishi (eds.). *Wood and cellulosic chemistry*. [Online]. New York: Mercel Dekker, Inc.

Garajova, S., Mathieu, Y., Beccia, M.R., Bennati-Granier, C., Biaso, F., Fanuel, M., Ropartz, D., Guigliarelli, B., Record, E., Rogniaux, H., Henrissat, B. & Berrin, J.-G. (2016). Single-domain flavoenzymes trigger lytic polysaccharide monooxygenases for oxidative degradation of cellulose. *Scientific Reports*. 6 (28276). pp. 1–9.

Gilbert, H.J. (2007). Cellulosomes: microbial nanomachines that display plasticity in quaternary structure. *Molecular Microbiology*. 63 (6). pp. 1568–1576.

Gilbert, H.J., Stålbrand, H. & Brumer, H. (2008). How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Current Opinion in Plant Biology*. 11 (3). pp. 338–348.

Gille, S. & Pauly, M. (2012). O-acetylation of plant cell wall polysaccharides. *Frontiers in Plant Science*. 3. pp. 1–7.

Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M.H., Vaaje-Kolstad, G., Lundberg, D., Royant, A., Ståhlberg, J., Eijsink, V.G.H., Beckham, G.T. & Sandgren, M. (2014). Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by X-ray photoreduction. *The Journal of Biological Chemistry*. 289 (27). pp. 18782–18792.

Hall, M., Bansal, P., Lee, J.H., Realff, M. & Bommarius, A.S. (2011a). Biological pretreatment of cellulose: enhancing enzymatic hydrolysis rate using cellulose-binding domains from cellulases. *Bioresource Technology*. 102 (3). pp. 2910–2915.

Hall, M., Rubin, J., Behrens, S.H. & Bommarius, A.S. (2011b). The cellulose-binding domain of cellobiohydrolase Cel7A from Trichoderma reesei is also a thermostabilizing domain. *Journal of Biotechnology*. 155 (4). pp. 370–376.

Hansson, H., Karkehabadi, S., Mikkelsen, N., Douglas, N.R., Kim, S., Lam, A., Kaper, T., Kelemen, B., Meier, K.K., Jones, S.M., Solomon, E.I. & Sandgren, M. (2017). High-resolution structure of a lytic polysaccharide monooxygenase from Hypocrea jecorina

reveals a predicted linker as an integral part of the catalytic domain. *Journal of Biological Chemistry*. p. jbc.M117.799767.

Harris, P. V, Welner, D., McFarland, K.C., Re, E., Navarro Poulsen, J.-C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S. & Lo Leggio, L. (2010). Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*. 49 (15). pp. 3305–3316.

Hasper, A.A., Dekkers, E., van Mil, M., van de Vondervoort, P.J.I. & de Graaff, L.H. (2002). EglC, a new endoglucanase from Aspergillus niger with major activity towards xyloglucan. *Applied and Environmental Microbiology*. 68 (4). pp. 1556–1560.

Hayashi, T. (1989). Xyloglucans in the primary cell wall. *Annual Review of Plant Physiology and Plant Molecular Biology*. 40. pp. 139–168.

Hemsworth, G.R., Davies, G.J. & Walton, P.H. (2013a). Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Current Opinion in Structural Biology*. 23 (5). pp. 660–668.

Hemsworth, G.R., Déjean, G., Davies, G.J. & Brumer, H. (2016). Learning from microbial strategies for polysaccharide degradation. *Biochemical Society Transactions*. 44 (1). pp. 94–108.

Hemsworth, G.R., Henrissat, B., Davies, G.J. & Walton, P.H. (2014). Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nature Chemical Biology*. 10 (2). pp. 122–126.

Hemsworth, G.R., Taylor, E.J., Kim, R.Q., Gregory, R.C., Lewis, S.J., Turkenburg, J.P., Parkin, A., Davies, G.J. & Walton, P.H. (2013b). The copper active site of CBM33 polysaccharide oxygenases. *Journal of the American Chemical Society*. 135 (16). pp. 6069–6077.

Henrissat, B., Driguez, H., Viet, C. & Schülein, M. (1985). Synergism of cellulases from Trichoderma reesei in the degradation of cellulose. *Nature Biotechnology*. 3. pp. 722–726.

Himmel, M.E., Ding, S.-Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W. & Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*. 315 (5813). pp. 804–807.

Hori, C., Ishida, T., Igarashi, K., Samejima, M., Suzuki, H., Master, E., Ferreira, P., Ruiz-Dueñas, F.J., Held, B., Canessa, P., Larrondo, L.F., Schmoll, M., Druzhinina, I.S., Kubicek, C.P., Gaskell, J.A., Kersten, P., St. John, F., Glasner, J., Sabat, G., Splinter BonDurant, S., Syed, K., Yadav, J., Mgbeahuruike, A.C., Kovalchuk, A., Asiegbu, F.O., Lackner, G., Hoffmeister, D., Rencoret, J., Gutiérrez, A., Sun, H., Lindquist, E., Barry, K., Riley, R., Grigoriev, I. V., Henrissat, B., Kües, U., Berka, R.M., Martínez, A.T., Covert, S.F., Blanchette, R.A. & Cullen, D. (2014). Analysis of the Phlebiopsis gigantea genome, transcriptome and secretome provides insight into its pioneer colonization strategies of wood G. P. Copenhaver (ed.). *PLoS Genetics*. 10 (12). pp. 1–20.

Horn, S.J., Vaaje-Kolstad, G., Westereng, B. & Eijsink, V.G.H. (2012). Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels*. 5 (45). pp. 1–12.

Hu, J., Arantes, V., Pribowo, A., Gourlay, K. & Saddler, J. (2014). Substrate factors that influence the synergistic interaction of AA9 and cellulases during the enzymatic hydrolysis of biomass. *Energy & Environmental Science*.

Isaksen, T., Westereng, B., Aachmann, F.L., Agger, J.W., Kracher, D., Kittl, R., Ludwig, R.,

Haltrich, D., Eijsink, V.G.H. & Horn, S.J. (2014). A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *The Journal of Biological Chemistry*. 289 (5). pp. 2632–2642.

Jagadeeswaran, G., Gainey, L., Prade, R. & Mort, A.J. (2016). 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. *Applied Microbiology and Biotechnology*. 100 (10). pp. 4535–4547.

Jarvis, M.C. (2013). Cellulose biosynthesis: counting the chains. *Plant Physiology*. 163 (4). pp. 1485–1486.

Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C. & Sandgren, M. (2008). The first structure of a glycoside hydrolase family 61 member, Cel61B from Hypocrea jecorina, at 1.6 Å resolution. *Journal of Molecular Biology*. 31. pp. 144–154.

Kirn, T.J., Jude, B.A. & Taylor, R.K. (2005). A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature*. 438 (7069). pp. 863–866.

Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. & Ludwig, R. (2012). Production of four Neurospora crassa lytic polysaccharide monooxygenases in Pichia pastoris monitored by a fluorimetric assay. *Biotechnology for Biofuels*. 5 (79). pp. 1–13.

Koenigs, J.W. (1974). Production of hydrogen peroxide by wood-rotting fungi in wood and its correlation with weight loss, depolymerization, and pH changes. *Archives of Microbiology*. 99 (1). pp. 129–145.

Kojima, Y., Várnai, A., Ishida, T., Sunagawa, N., Petrovic, D.M., Igarashi, K., Jellison, J., Goodell, B., Alfredsen, G., Westereng, B., Eijsink, V.G.H. & Yoshida, M. (2016). Characterization of an LPMO from the brown-rot fungus Gloephyllum trabeum with broad xyloglucan specificity, and its action on cellulose-xyloglucan complexes. *Applied and Environmental Microbiology*. 82 (22). pp. 6557–6572.

Kolbe, S., Fischer, S., Becirevic, A., Hinz, P. & Schrempf, H. (1998). The Streptomyces reticuli alpha-chitin-binding protein CHB2 and its gene. *Microbiology*. 144 (5). pp. 1291–1297.

Koropatkin, N.M., Martens, E.C., Gordon, J.I. & Smith, T.J. (2008). Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure*. 16 (7). pp. 1105–1115.

Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews*. 28 (4). pp. 416–436.

Kracher, D., Scheiblbrandner, S., Felice, A.K.G., Breslmayr, E., Preims, M., Ludwicka, K., Haltrich, D., Eijsink, V.G.H. & Ludwig, R. (2016). Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science*. 352 (6289). pp. 1098–1101.

Kracher, D., Zahma, K., Schulz, C., Sygmund, C., Gorton, L. & Ludwig, R. (2015). Interdomain electron transfer in cellobiose dehydrogenase: modulation by pH and divalent cations. *FEBS Journal*. 282 (16). pp. 3136–3148.

Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annual Review of Biochemistry*. 77 (1). pp. 521–555.

Lamed, R., Setter, E. & Bayer, E.A. (1983). Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum. *Journal of Bacteriology*. 156 (2). pp. 828–836.

Langston, J.A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. & Sweeney, M.D. (2011). Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Applied and Environmental Microbiology*. 77 (19). pp. 7007–7015.

Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tauzin, A.S., Spadiut, O., Klinter, S., Pudlo, N. a., Urs, K., Koropatkin, N.M., Creagh, a. L., Haynes, C.A., Kelly, A.G., Cederholm, S.N., Davies, G.J., Martens, E.C. & Brumer, H. (2014). A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature*. 506 (7489). pp. 498–502.

Larsbrink, J., Zhu, Y., Kharade, S.S., Kwiatkowski, K.J., Eijsink, V.G.H., Koropatkin, N.M., McBride, M.J. & Pope, P.B. (2016). A polysaccharide utilization locus from Flavobacterium johnsoniae enables conversion of recalcitrant chitin. *Biotechnology for Biofuels*. 9 (260). pp. 1–16.

Lo Leggio, L., Simmons, T.J., Poulsen, J.-C.N., Frandsen, K.E.H., Hemsworth, G.R., Stringer, M.A., von Freiesleben, P., Tovborg, M., Johansen, K.S., De Maria, L., Harris, P. V., Soong, C.-L., Dupree, P., Tryfona, T., Lenfant, N., Henrissat, B., Davies, G.J. & Walton, P.H. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nature Communications*. 6 (5961). pp. 1–9.

Lo Leggio, L., Welner, D. & De Maria, L. (2012). A structural overview of GH61 proteins - fungal cellulose degrading polysaccharide monooxygenases. *Computational and Structural Biotechnology Journal*. 2. pp. 1–8.

Lenfant, N., Hainaut, M., Terrapon, N., Drula, E., Lombard, V. & Henrissat, B. (2017). A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydrate Research*. 448. pp. 166–174.

Leschine, S.B. (1995). Cellulose degradation in anaerobic environments. *Annual Review of Microbiology*. 49 (1). pp. 399–426.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M. & Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels*. 6 (41). pp. 1–14.

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

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*. 42. pp. 490–495.

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

Martens, E.C., Koropatkin, N.M., Smith, T.J. & Gordon, J.I. (2009). Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *Journal of Biological Chemistry*. 284 (37). pp. 24673–24677.

Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C.P., Ferreira, P., Ruiz-Duenas, F.J., Martinez, A.T., Kersten, P., Hammel, K.E., Vanden Wymelenberg, A., Gaskell, J., Lindquist, E., Sabat, G., Bondurant, S.S., Larrondo, L.F., Canessa, P., Vicuna, R., Yadav, J., Doddapaneni, H., Subramanian, V., Pisabarro, A.G.,

Lavín, J.L., Oguiza, J.A., Master, E., Henrissat, B., Coutinho, P.M., Harris, P., Magnuson, J.K., Baker, S.E., Bruno, K., Kenealy, W., Hoegger, P.J., Kües, U., Ramaiya, P., Lucas, S., Salamov, A., Shapiro, H., Tu, H., Chee, C.L., Misra, M., Xie, G., Teter, S., Yaver, D., James, T., Mokrejs, M., Pospisek, M., Grigoriev, I. V, Brettin, T., Rokhsar, D., Berka, R. & Cullen, D. (2009). Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (6). pp. 1954–1959.

Mba Medie, F., Davies, G.J., Drancourt, M. & Henrissat, B. (2012). Genome analyses highlight the different biological roles of cellulases. *Nature Reviews Microbiology*. 10. pp. 227–234.

McFarlane, H.E., Döring, A. & Persson, S. (2014). The Cell Biology of Cellulose Synthesis. *Annual Review of Plant Biology*. 65 (1). pp. 69–94.

Mekasha, S., Forsberg, Z., Dalhus, B., Bacik, J.-P., Choudhary, S., Schmidt-Dannert, C., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2015). Structural and functional characterization of a small chitin-active lytic polysaccharide monooxygenase domain of a multi-modular chitinase from Jonesia denitrificans. *FEBS Letters*. 590. pp. 34–42.

Merino, S.T. & Cherry, J. (2007). Progress and challenges in enzyme development for biomass utilization. *Advances in Biochemical Engineering/ Biotechnology*. 108. pp. 95–120.

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

Morgan, J.L.W., Strumillo, J. & Zimmer, J. (2012). Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature*. 493 (7431). pp. 181–186.

Müller, G., Várnai, A., Johansen, K.S., Eijsink, V.G.H. & Horn, S.J. (2015). Harnessing the potential of LPMO-containing cellulase cocktails poses new demands on processing conditions. *Biotechnology for Biofuels*. 8 (187). pp. 1–9.

Navarro, D., Rosso, M.-N., Haon, M., Olivé, C., Bonnin, E., Lesage-Meessen, L., Chevret, D., Coutinho, P.M., Henrissat, B. & Berrin, J.-G. (2014). Fast solubilization of recalcitrant cellulosic biomass by the basidiomycete fungus Laetisaria arvalis involves successive secretion of oxidative and hydrolytic enzymes. *Biotechnology for Biofuels*. 7 (143). pp. 1–14.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R.A., Zhu, Y., O'Neill, M.A., Urbanowicz, B.R., York, W.S., Davies, G.J., Abbott, D.W., Ralet, M.-C., Martens, E.C., Henrissat, B. & Gilbert, H.J. (2017). Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*. 544 (7648). pp. 65–70.

Nekiunaite, L., Petrović, D.M., Westereng, B., Vaaje-Kolstad, G., Hachem, M.A., Várnai, A. & Eijsink, V.G.H. (2016). FgLPMO9A from Fusarium graminearum cleaves xyloglucan independently of the backbone substitution pattern. *FEBS Letters*. 590. pp. 3346–3356.

Newman, R.H., Hill, S.J. & Harris, P.J. (2013). Wide-angle X-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiology*. 163 (4). pp. 1558–1567.

Nishiyama, Y. (2009). Structure and properties of the cellulose microfibril. *Journal of Wood*

Science. 55 (4). pp. 241–249.

Nixon, B.T., Mansouri, K., Singh, A., Du, J., Davis, J.K., Lee, J.-G., Slabaugh, E., Vandavasi, V.G., O'Neill, H., Roberts, E.M., Roberts, A.W., Yingling, Y.G. & Haigler, C.H. (2016). Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. *Scientific Reports*. 6 (1). pp. 1–14.

Naas, A.E., Mackenzie, A.K., Mravec, J., Schuckel, J., Willats, W.G.T., Eijsink, V.G.H. & Pope, P.B. (2014). Do rumen Bacteroidetes utilize an alternative mechanism for cellulose degradation? *mBio*. 5 (4). pp. 1–6.

O'Sullivan, A.C. (1997). Cellulose: the structure slowly unravels. *Cellulose*. 4 (3). pp. 173–207.

Park, Y.B. & Cosgrove, D.J. (2015). Xyloglucan and its interactions with other components of the growing cell wall. *Plant and Cell Physiology*. 56 (2). pp. 180–194.

Park, Y.B. & Cosgrove, D.J. (2012). A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiology*. 158 (4). pp. 1933–1943.

Paspaliari, D.K., Loose, J.S.M., Larsen, M.H. & Vaaje-Kolstad, G. (2015). *Listeria monocytogenes* has a functional chitinolytic system and an active lytic polysaccharide monooxygenase. *FEBS Journal*. 282 (5). pp. 921–936.

Pauly, M. & Keegstra, K. (2016). Biosynthesis of the plant cell wall matrix polysaccharide xyloglucan. *Annual Review of Plant Biology*. 67 (1). pp. 235–259.

Peña, M.J., Darvill, A.G., Eberhard, S., York, W.S. & O'Neill, M.A. (2008). Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology*. 18 (11). pp. 891–904.

Pereira, C.S., Silveira, R.L., Dupree, P. & Skaf, M.S. (2017). Effects of xylan side-chain substitutions on xylan-cellulose interactions and implications for thermal pretreatment of cellulosic biomass. *Biomacromolecules*. 18 (4). pp. 1311–1321.

Phillips, C.M., Beeson, W.T., Cate, J.H. & Marletta, M.A. (2011). Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by Neurospora crassa. *ACS Chemical Biology*. 6 (12). pp. 1399–1406.

Pollegioni, L., Tonin, F. & Rosini, E. (2015). Lignin-degrading enzymes. *FEBS Journal*. 282 (7). pp. 1190–1213.

Pope, P.B., Mackenzie, A.K., Gregor, I., Smith, W., Sundset, M.A., McHardy, A.C., Morrison, M. & Eijsink, V.G.H. (2012). Metagenomics of the Svalbard reindeer rumen microbiome reveals abundance of polysaccharide utilization loci. *PloS ONE*. 7 (6). pp. 1–10.

Pringsheim, H. (1912). Über der fermentativen Abbau der Zellulose. *Angewandte Chemie*. 78. pp. 266–291.

Quinlan, R.J., Sweeney, M.D., Lo Leggio, L., Otten, H., Poulsen, J.-C.N., Johansen, K.S., Krogh, K.B.R.M., Jørgensen, C.I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C.P., Dupree, P., Xu, F., Davies, G.J. & Walton, P.H. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (37). pp. 15079–15084.

Raguz, S., Yagüe, E., Wood, D.A. & Thurston, C.F. (1992). Isolation and characterization of

a cellulose-growth-specific gene from Agaricus bisporus. *Gene*. 119 (2). pp. 183–190.

Reese, E.T., Siu, R.G.H. & Levinson, H.S. (1950). The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *Journal of bacteriology*. 59 (4). pp. 485–497.

Resch, M.G., Donohoe, B.S., Baker, J.O., Decker, S.R., Bayer, E.A., Beckham, G.T. & Himmel, M.E. (2013). Fungal cellulases and complexed cellulosomal enzymes exhibit synergistic mechanisms in cellulose deconstruction. *Energy & Environmental Science*. 6. pp. 1858–1867.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K. & Jones, T.A. (1990). Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei. *Science*. 249 (4967). pp. 380–386.

Rubin, E.M. (2008). Genomics of cellulosic biofuels. *Nature*. 454 (7206). pp. 841–845.

Saha, B.C., Yoshida, T., Cotta, M. a. & Sonomoto, K. (2013). Hydrothermal pretreatment and enzymatic saccharification of corn stover for efficient ethanol production. *Industrial Crops and Products*. 44. pp. 367–372.

Scheller, H.V. & Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*. 61. pp. 263–289.

Schnellmann, J., Zeltins, A., Blaak, H. & Schrempf, H. (1994). The novel lectin-like protein CHB1 is encoded by a chitin-inducible Streptomyces olivaceoviridis gene and binds specifically to crystalline alpha-chitin of fungi and other organisms. *Molecular Microbiology*. 13 (5). pp. 807–819.

Schrempf, H. (1999). Characteristics of chitin-binding proteins from Streptomycetes. *EXS*. 87. pp. 99–108.

Schwarz, W.H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology*. 56 (5–6). pp. 634–649.

Seillière, G.C. (1907). Remarques sur l'hydrolyse disatasique de la cellulose du coton et de quelques autres polysaccharides. *Comp. Rend. Soc. Biol.* 63. pp. 515–517.

Shallom, D. & Shoham, Y. (2003). Microbial hemicellulases. *Current Opinion in Microbiology*. 6 (3). pp. 219–228.

Shen, D., Xiao, R., Gu, S. & Luo, K. (2011). The pyrolytic behavior of cellulose in lignocellulosic biomass: a review. *RSC Advances*. 1 (9). pp. 1641–1660.

Simmons, T.J., Frandsen, K.E.H., Ciano, L., Tryfona, T., Lenfant, N., Poulsen, J.C., Wilson, L.F.L., Tandrup, T., Tovborg, M., Schnorr, K., Johansen, K.S., Henrissat, B., Walton, P.H., Lo Leggio, L. & Dupree, P. (2017). Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nature Communications*. 8 (1064). pp. 1–12.

Simmons, T.J., Mortimer, J.C., Bernardinelli, O.D., Pöppler, A.-C., Brown, S.P., DeAzevedo, E.R., Dupree, R. & Dupree, P. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications*. 7 (13902). pp. 1–9.

Sjöström, E. (1993). *Wood chemistry: Fundamentals and applications*. Academic Press Inc.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S. & Youngs, H. (2004). Toward a systems approach to understanding plant cell walls. *Science*. 306. pp. 2206–2211.

Span, E.A. & Marletta, M.A. (2015). The framework of polysaccharide monooxygenase

structure and chemistry. *Current Opinion in Structural Biology*. 35. pp. 93–99.

Span, E. a., Suess, D.L.M., Deller, M.C., Britt, R.D. & Marletta, M.A. (2017). The role of the secondary coordination sphere in a fungal polysaccharide monooxygenase. *ACS Chemical Biology*. 12. pp. 1095–1103.

Sunna, A., Gibbs, M.D., Chin, C.W., Nelson, P.J. & Bergquist, P.L. (2000). A gene encoding a novel multidomain beta-1,4-mannanase from Caldibacillus cellulovorans and action of the recombinant enzyme on kraft pulp. *Applied and Environmental Microbiology*. 66 (2). pp. 664–670.

Suzuki, K., Suzuki, M., Taiyoji, M., Nikaibou, N. & Watanabe, T. (1998). Chitin binding protein (CBP21) in the culture supernatant of Serratia marcescens 2170. *Bioscience, Biotechnology, and Biochemistry*. 62 (1). pp. 128–135.

Sygmund, C., Kracher, D., Scheiblbrandner, S., Zahma, K., Felice, A.K.G., Harreither, W., Kittl, R. & Ludwig, R. (2012). Characterization of the two Neurospora crassa cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Applied and Environmental Microbiology*. 78 (17). pp. 6161–6171.

Takasuka, T.E., Book, A.J., Lewin, G.R., Currie, C.R. & Fox, B.G. (2013). Aerobic deconstruction of cellulosic biomass by an insect-associated Streptomyces. *Scientific Reports*. 3. pp. 1–10.

Talbott, L.D. & Ray, P.M. (1992). Changes in molecular size of previously deposited and newly synthesized pea cell wall matrix polysaccharides. *Plant Physiology*. 98 (1). pp. 369–379.

Tan, T.-C., Kracher, D., Gandini, R., Sygmund, C., Kittl, R., Haltrich, D., Hällberg, B.M., Ludwig, R. & Divne, C. (2015). Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nature Communications*. 6 (7542). pp. 1–11.

Tian, C., Beeson, W.T., Iavarone, A.T., Sun, J., Marletta, M.A., Cate, J.H.D. & Glass, N.L. (2009). Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. *Proceedings of the National Academy of Sciences of the United States of America*. 106. pp. 22157–22162.

Tuveng, T.R., Arntzen, M.Ø., Bengtsson, O., Gardner, J.G., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2016). Proteomic investigation of the secretome of Cellvibrio japonicus during growth on chitin. *Proteomics*. 16 (13). pp. 1904–1914.

Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., Soetaert, W. & Boerjan, W. (2013). Towards a carbon-negative sustainable bio-based economy. *Frontiers in Plant Scienceant science*. 4 (174). pp. 1–17.

Várnai, A., Siika-aho, M. & Viikari, L. (2013). Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnology for Biofuels*. 6 (30). pp. 1–11.

Vincken, J.P., Beldman, G. & Voragen, A.G.J. (1994). The effect of xyloglucans on the degradation of cell-wall-embedded cellulose by the combined action of cellobiohydrolase and endoglucanases from Trichoderma viride. *Plant Physiology*. 104 (1). pp. 99–107.

Vincken, J.P., de Keizer, A., Beldman, G. & Voragen, A.G.J. (1995). Fractionation of xyloglucan fragments and their interaction with cellulose. *Plant Physiology*. 108 (4). pp. 1579–1585.

Vincken, J.P., York, W.S., Beldman, G. & Voragen, A.G.J. (1997). Two general branching

patterns of xyloglucan, XXXG and XXGG. *Plant Physiology*. 114 (1). pp. 9–13.

Voutilainen, S.P., Puranen, T., Siika-aho, M., Lappalainen, A., Alapuranen, M., Kallio, J., Hooman, S., Viikari, L., Vehmaanperä, J. & Koivula, A. (2008). Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. *Biotechnology and Bioengineering*. 101 (3). pp. 515–528.

Vu, V. V, Beeson, W.T., Phillips, C.M., Cate, J.H.D. & Marletta, M.A. (2014a). Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *Journal of the American Chemical Society*. 136 (2). pp. 562–565.

Vu, V. V, Beeson, W.T., Span, E.A., Farquhar, E.R. & Marletta, M.A. (2014b). A family of starch-active polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (38). pp. 13822–7.

Vaaje-Kolstad, G., Bøhle, L.A., Gåseidnes, S., Dalhus, B., Bjørås, M., Mathiesen, G. & Eijsink, V.G.H. (2012). Characterization of the chitinolytic machinery of Enterococcus faecalis V583 and high-resolution structure of its oxidative CBM33 enzyme. *Journal of Molecular Biology*. 416 (2). pp. 239–254.

Vaaje-Kolstad, G., Forsberg, Z., Loose, J.S.M., Bissaro, B. & Eijsink, V.G.H. (2017). Structural diversity of lytic polysaccharide monooxygenases. *Current Opinion in Structural Biology*. 44. pp. 67–76.

Vaaje-Kolstad, G., Horn, S.J., van Aalten, D.M.F., Synstad, B. & Eijsink, V.G.H. (2005a). The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. *Journal of Biology Chemistry*. 280. pp. 28492–28497.

Vaaje-Kolstad, G., Houston, D.R., Riemen, A.H.K., Eijsink, V.G.H. & van Aalten, D.M.F. (2005b). Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. *The Journal of Biological Chemistry*. 280 (12). pp. 11313–11319.

Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., Sørlie, M. & Eijsink, V.G.H. (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 330 (6001). pp. 219–222.

Walton, P.H. & Davies, G.J. (2016). On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Current Opinion in Chemical Biology*. 31. pp. 195–207.

Wang, L., Zhang, Y. & Gao, P. (2008). A novel function for the cellulose binding module of cellobiohydrolase I. *Science in China Series C: Life Sciences*. 51 (7). pp. 620–629.

Wang, T., Yang, H., Kubicki, J.D. & Hong, M. (2016). Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules*. 17 (6). pp. 2210–2222.

Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., Cayouette, M., McHardy, A.C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S.G., Podar, M., Martin, H.G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N.C., Matson, E.G., Ottesen, E.A., Zhang, X., Hernández, M., Murillo, C., Acosta, L.G., Rigoutsos, I., Tamayo, G., Green, B.D., Chang, C., Rubin, E.M., Mathur, E.J., Robertson, D.E., Hugenholtz, P. & Leadbetter, J.R. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*. 450 (7169). pp. 560–565.

Westereng, B., Cannella, D., Wittrup Agger, J., Jørgensen, H., Larsen Andersen, M., Eijsink, V.G.H. & Felby, C. (2015). Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Scientific Reports*. 5 (18561). pp. 1–9.

Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V.G.H., Igarashi, K., Samejima, M., Ståhlberg, J., Horn, S.J. & Sandgren, M. (2011). The putative endoglucanase PcGH61D from Phanerochaete chrysosporium is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE*. 6 (11). pp. 1–11.

Westereng, B., Wittrup, J., Horn, S.J. & Vaaje-kolstad, G. (2013). Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *Journal of Chromatography A*. 1271 (1). pp. 144–152.

Westereng, B., Arntzen, M.T., Aachmann, F.L., Várnai, A., Eijsink, V.G.H. & Agger, J.W. (2016). Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monooxygenases acting on cellulose. *Journal of Chromatography A*. 1445. pp. 46–54.

Westereng, B., Arntzen, M.Ø., Agger, J.W., Vaaje-Kolstad, G. & Eijsink, V.G.H. (2017). Analyzing Activities of Lytic Polysaccharide Monooxygenases by Liquid Chromatography and Mass Spectrometry. In: D. W. Abbott & A. van Bueren (eds.). *Protein-Carbohydrate Interactions: Methods and Protocols*. [Online]. New York, NY: Springer New York, pp. 71–92.

Wong, E., Vaaje-Kolstad, G., Ghosh, A., Hurtado-Guerrero, R., Konarev, P. V, Ibrahim, A.F.M., Svergun, D.I., Eijsink, V.G.H., Chatterjee, N.S. & van Aalten, D.M.F. (2012). The Vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS Pathogens*. 8 (1). pp. 1–12.

Wood, T.M. & Garcia-Campayo, V. (1990). Enzymology of cellulose degradation. *Biodegradation*. 1 (2–3). pp. 147–161.

Wu, M., Beckham, G.T., Larsson, A.M., Ishida, T., Kim, S., Payne, C.M., Himmel, M.E., Crowley, M.F., Horn, S.J., Westereng, B., Igarashi, K., Samejima, M., Ståhlberg, J., Eijsink, V.G.H. & Sandgren, M. (2013). Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the basidiomycota fungus Phanerochaete chrysosporium. *The Journal of Biological Chemistry*. 288 (18). pp. 12828–12839.

Van Wyk, J.P.H. (2011). Biowaste as a resource for bioproduct development. *Environmental Earth Sciences*. 19 (5). pp. 875–883.

Xiao, Z., Gao, P., Qu, Y. & Wang, T. (2001). Cellulose-binding domain of endoglucanase III from Trichoderma reesei disrupting the structure of cellulose. *Biotechnology Letters*. 23 (9). pp. 711–715.

Zeltins, A. & Schrempf, H. (1997). Specific interaction of the Streptomyces chitin-binding protein CHB1 with alpha-chitin -the role of individual tryptophan residues. *European Journal of Biochemistry*. 246 (2). pp. 557–564.

Paper I

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 5, pp. 2632–2642, January 31, 2014

A C4-oxidizing Lytic Polysaccharide Monooxygenase Cleaving Both Cellulose and Cello-oligosaccharides*

Received for publication, October 28, 2013, and in revised form, December 3, 2013 Published, JBC Papers in Press, December 9, 2013, DOI 10.1074/jbc.M113.530196

Trine Isaksen[‡], Bjørge Westereng[‡§], Finn L. Aachmann[¶], Jane W. Agger[‡], Daniel Kracher[‖], Roman Kittl[‖], Roland Ludwig[‖], Dietmar Haltrich[‖], Vincent G. H. Eijsink[‡1], and Svein J. Horn[‡]

From the [‡]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway, [§]Forest and Landscape, University of Copenhagen, DK-1958 Frederiksberg C, Denmark, the [¶]Department of Biotechnology, NOBIPOL, Norwegian University of Science and Technology, N-7491 Trondheim, Norway, and the [‖]Department of Food Science and Technology, BOKU, University of Natural Resources and Life Sciences, A-1190 Vienna, Austria

Background: Lytic polysaccharide monooxygenases (LPMOs) are recently discovered enzymes that cleave polysaccharides.
Results: We describe a novel LPMO and use a range of analytical methods to characterize its activity.
Conclusion: Cellulose and cello-oligosaccharides are cleaved by oxidizing the sugar at the nonreducing end in the C4 position.
Significance: This study provides unequivocal evidence for C4 oxidation of the nonreducing end sugar and demonstrates a novel LPMO substrate specificity.

Lignocellulosic biomass is a renewable resource that significantly can substitute fossil resources for the production of fuels, chemicals, and materials. Efficient saccharification of this biomass to fermentable sugars will be a key technology in future biorefineries. Traditionally, saccharification was thought to be accomplished by mixtures of hydrolytic enzymes. However, recently it has been shown that lytic polysaccharide monooxygenases (LPMOs) contribute to this process by catalyzing oxidative cleavage of insoluble polysaccharides utilizing a mechanism involving molecular oxygen and an electron donor. These enzymes thus represent novel tools for the saccharification of plant biomass. Most characterized LPMOs, including all reported bacterial LPMOs, form aldonic acids, *i.e.*, products oxidized in the C1 position of the terminal sugar. Oxidation at other positions has been observed, and there has been some debate concerning the nature of this position (C4 or C6). In this study, we have characterized an LPMO from *Neurospora crassa* (*Nc*LPMO9C; also known as NCU02916 and *Nc*GH61–3). Remarkably, and in contrast to all previously characterized LPMOs, which are active only on polysaccharides, *Nc*LPMO9C is able to cleave soluble cello-oligosaccharides as short as a tetramer, a property that allowed detailed product analysis. Using mass spectrometry and NMR, we show that the cello-oligosaccharide products released by this enzyme contain a C4 gemdiol/keto group at the nonreducing end.

In the emerging bio-economy, plant biomass will gradually substitute fossil resources for the production of fuels, chemicals, and materials. One of the main bottlenecks in such biorefining processes is the depolymerization of cellulose, a major constituent of the plant cell wall, to fermentable sugars. In nature this process is catalyzed by cellulases and the recently discovered lytic polysaccharide monooxygenases (LPMOs)[2] (1). Enzymes and binding domains interacting with polysaccharides are categorized in the CAZy database, which comprises families of structurally related carbohydrate-active enzymes, such as glycoside hydrolases (GH), and carbohydrate-binding modules (CBMs) (2). LPMOs were originally classified as CBM33 (family 33 carbohydrate-binding module) or GH61 (family 61 glycoside hydrolase). However, CAZy has recently been revised, and GH61 and CBM33 are now named LPMOs and classified under the heading "auxiliary activities" (AA) as families AA9 and AA10, respectively (3).

The enzyme activities of LPMOs were first discovered in 2010 for an AA10 protein (CBP21) acting on chitin (4). Following this study, cellulose active LPMOs were found in both the AA10 (5) and AA9 (6–8) families. These copper-dependent enzymes carry out oxidative cleavage of the β-1,4-glycosidic bonds in polysaccharides, using molecular oxygen and an electron donor (1). Electrons may be supplied by small molecule reductants such as ascorbic acid and gallic acid (4, 6) or by enzymes, such as cellobiose dehydrogenase (CDH), that are co-

* This work was supported by Norwegian Research Council Projects 193817, 203402, 214613, 216162, and 217708 and by the European Commission through the FP7-KBBE-2013-7-613549 project INDOX.

[1] To whom correspondence should be addressed: Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences, 1432 Ås, Norway. Tel.: +47 64965892; Fax: +47 64965901; E-mail: vincent.eijsink@nmbu.no.

[2] The abbreviations used are (full IUPAC abbreviations in bold type for clarity): LPMO, lytic polysaccharide monooxygenase; CBM, carbohydrate-binding module; GH, glycoside hydrolase; AA, auxiliary activities; DP, degree of polymerization; CDH, cellobiose dehydrogenase; PASC, phosphoric acid swollen cellulose; HPAEC, high performance anion exchange chromatography; DQF-COSY, double quantum filter correlation spectroscopy; IP-COSY, in-phase COSY; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single quantum coherence; HMBC, heteronuclear multibond correlation; Glc4gemGlc, 4-hydroxy-β-D-*xylo*-hexopyranosyl-(1→4)-β-D-glucopyranosyl, **4-hydroxy-β-D-*xylo*-Hex*p*-(1→4)-β-D-Glc*p***; Glc4KGlc$_2$, β-D-*xylo*-hexos-4-ulopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl, **β-D-*xylo*-Hex4ulo*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p***; Glc4gemGlc$_2$, 4-hydroxy-β-D-*xylo*-hexopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-β-D-glucopyranosyl, **4-hydroxy-β-D-*xylo*-Hex*p*-(1→4)-β-D-Glc*p*-(1→4)-β-D-Glc*p***; Glc4KGlcGlc1A, β-D-*xylo*-hexos-4-ulopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-D-gluconic acid, **β-D-*xylo*-Hex4ulop-(1→4)-β-D-Glc*p*-(1→4)-D-Glc1A**; Glc4gemGlcGlc1A, 4-hydroxy-β-D-*xylo*-hexopyranosyl-(1→4)-β-D-glucopyranosyl-(1→4)-D-gluconic acid, **4-hydroxy-β-D-*xylo*-Hex*p*-(1→4)-β-D-Glc*p*-(1→4)-D-Glc1A**.

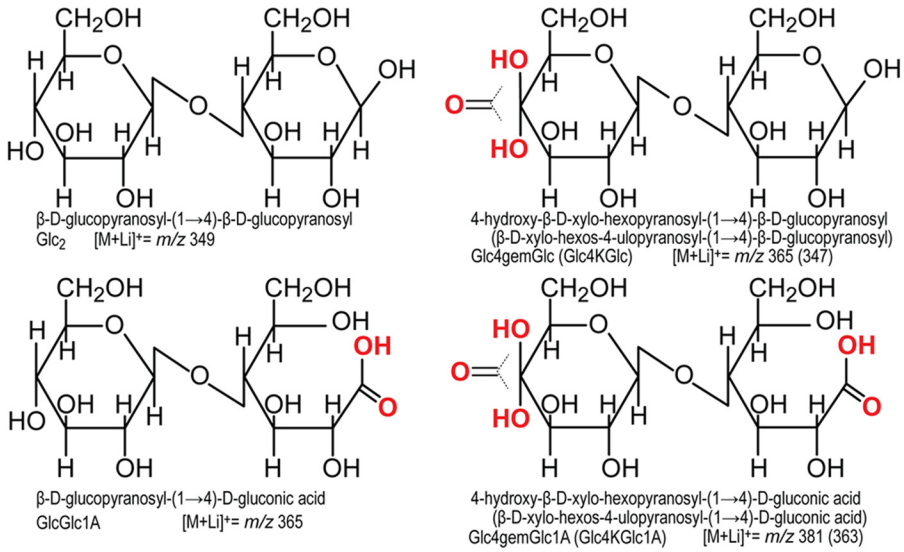

FIGURE 1. **Overview of cellobiose with C1 and/or C4 oxidations.** Oxidations are shown in red. The IUPAC names, the abbreviations, and the *m/z* values for lithium adducts are provided below the structures (please see Table 1 for a complete list of other adducts). In the *right two structures*, the IUPAC name, the abbreviation, and the *m/z* of the lithium adduct for the keto-form are given in *parentheses*.

expressed with LPMOs (8–10). The products of the reaction are oxidized oligosaccharides and native oligosaccharides containing reducing ends originally present in the polymeric substrate (1, 7) (Fig. 1). Although cleavage of polysaccharides by C1 oxidizing LPMOs, which yields aldonic acids, has been thoroughly demonstrated and analyzed (4, 5, 7, 8, 11, 12), oxidation at the nonreducing end is more difficult to analyze. Based on mass spectrometry, both oxidation at C4 and C6 have been suggested (6, 8, 13). For products generated by *Nc*LPMO9D (also known as NCU01050 or *Nc*GH61–4), oxidation at C4 rather than at C6 has been shown indirectly by detection of the C4 epimer of glucose, galactose, upon reduction of reaction products and by the absence of glucuronic acid upon hypoiodite oxidation of reaction products (14). However, direct evidence for the identity of the nonreducing end oxidized species is lacking.

The filamentous ascomycete *Neurospora crassa* is an efficient degrader of plant cell walls and produces a wide range of LPMOs and hydrolytic enzymes. The genome of *N. crassa* is predicted to contain 14 AA9 family LPMOs, six of which are attached to a CBM1 carbohydrate-binding module (15). These CBM1 modules contain ~40 amino acids, typically bind cellulose, and are almost exclusively found in fungi. The activities of three of these LPMOs on cellulose have been qualitatively characterized by HPLC and MS analyses of released products. *Nc*LPMO9E (NCU08760, *Nc*GH61–5; attached to a CBM1) oxidizes C1, *Nc*LPMO9D exclusively oxidizes the nonreducing end, and *Nc*LPMO9M (NCU07898, *Nc*GH61–13) seems to be capable of oxidizing both C1 and the nonreducing end (8, 16). In another study, using CDH as an electron donor, it was shown that three additional *N. crassa* AA9s, *Nc*LPMO9C (NCU02916, *Nc*GH61–3; attached to a CBM1), *Nc*LPMO9F (NCU03328, *Nc*GH61–6), and *Nc*LPMO9J (NCU01867, *Nc*GH61–10; attached to a CBM1), degrade cellulose, but no attempts were made to unravel details of the reaction products of these enzymes (17).

In this study, we have characterized the activity of *Nc*LPMO9C using NMR, mass spectrometry, HPLC, and a previously described activity assay (17). Interestingly, *Nc*LPMO9C turned out to be active on soluble substrates, which is an activity not previously described for LPMOs. Exploiting this unique property, we used NMR analysis to identify the products generated by *Nc*LPMO9C.

EXPERIMENTAL PROCEDURES

Production and Purification of Enzymes—The AA9 encoding *N. crassa* gene NCU02916 was codon-optimized, cloned with its native signal sequence under control of the methanol inducible AOX1 promotor, and recombinantly produced in *Pichia pastoris* X-33 following a published protocol (10). The protein was purified from 0.4 liter of culture supernatant by three subsequent chromatographic steps following a published method (17). In total, 28 mg of purified *Nc*LPMO9C was obtained, and the homogeneity was verified by SDS-PAGE. Cellobiose dehydrogenase from *Myriococcum thermophilum* carrying a C-terminal CBM1 (*Mt*CDH, Uniprot accession number A9XK88) (18) was recombinantly expressed in *P. pastoris* using methanol for induction (19). The enzyme was purified from 1 liter of culture supernatant by two chromatographic steps according to the procedure described by Harreither *et al.* (20), and 180 mg of homogeneous *Mt*CDH was obtained.

Activity Assays—Standard reaction mixtures (100–300 μl of liquid volume) contained the substrate (0.4 mg/ml cello-oligosaccharide degree of polymerization (DP) 3–6; (Megazyme), 1 mg/ml phosphoric acid swollen cellulose (PASC prepared from Avicel as described by Wood (21), 2.5 mg/ml Avicel PH-101 (Fluka) or (0.15–0.30 mg/ml) steam-exploded spruce), 4.4–8.8 μM *Nc*LPMO9C, 5 mM ammonium acetate buffer, pH 6.0, and 2 mM reducing agent (hydroquinone, ascorbic acid, or catechin; all from Sigma-Aldrich). In some reactions, the reducing agent was replaced by 1.4 μM *Mt*CDH. Additional substrates tested under the same conditions were: maltodextrin (DP4–14; Glu-

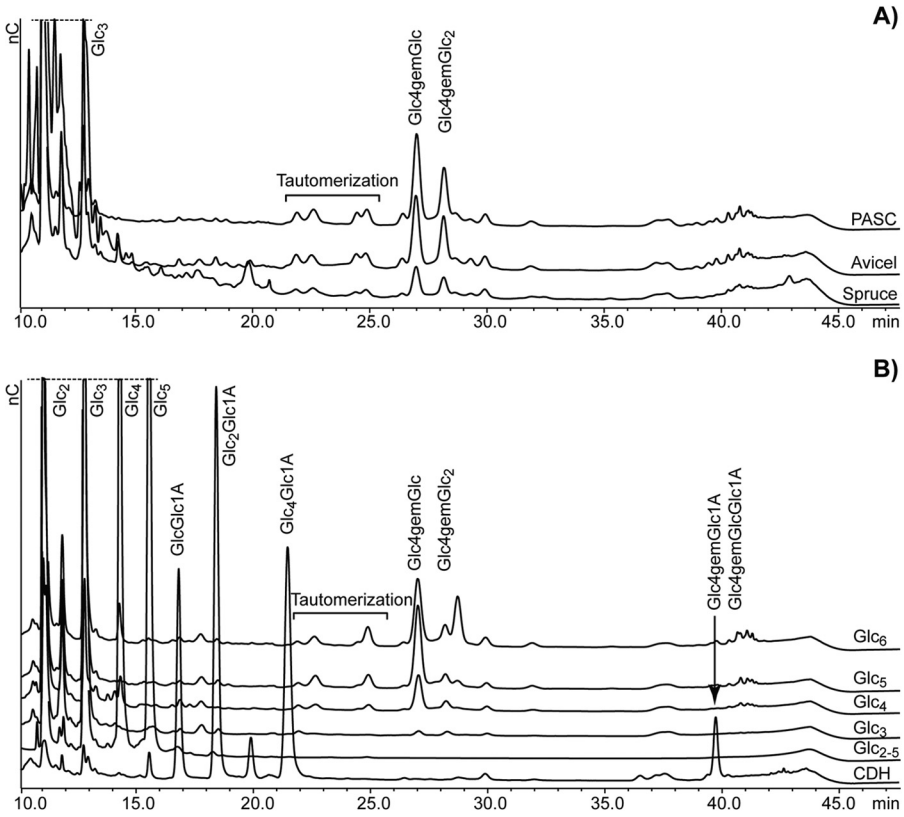

FIGURE 2. **Degradation of cellulose and cello-oligosaccharides by *Nc*LPMO9C.** *A* shows overlaid HPAEC chromatograms of oligosaccharides released upon degradation of 1 mg/ml of PASC, Avicel, and steam-exploded spruce with 8.8 μM *Nc*LPMO9C. *B* shows products released upon incubation of 0.4 mg/ml cello-oligosaccharides with 4.4 μM *Nc*LPMO9C (Glc_n traces), a standard mixture of nonoxidized cello-oligosaccharides (DP2–5, Glc_{2-5} trace), as well as products released upon incubation of cellopentaose with 4.4 μM *Nc*LPMO9C and 1.4 μM *Mt*CDH (CDH trace). All reaction mixtures contained 2 mM ascorbic acid (except the one containing CDH) and 5 mM ammonium acetate buffer, pH 6.0, and were incubated for 24 h at 50 °C. Reactions with insoluble substrates were incubated with horizontal shaking at 850 rpm. For product nomenclature, see Fig. 1. See text for further details.

cidex 9 from Roquette) mannohexaose, xylopentaose, xylohexaose, chitopentaose (all from Megazyme), α-chitin from shrimp shell (HovBio, Tromsø, Norway), and nanofibrillar α-chitin and β-chitin (10–20 nm) (22). All reactions were incubated in 2-ml screw cap microtubes (Sarstedt) at 50 °C in an Eppendorf Thermo mixer with (for insoluble substrates, 850 rpm) or without (for soluble substrates) shaking.

HPLC Analysis—From the standard reactions, samples were taken at different time points, and the reaction was stopped by adding NaOH to a final concentration of 0.05 M. After removing insoluble substrates by centrifugation, the supernatant was centrifuged and analyzed by high performance anion exchange chromatography (HPAEC) using an ICS3000 system (Dionex, Sunnivale, CA) as described previously (12). In brief, a 2-μl sample was injected on a CarboPac PA1 2 × 250 mm analytical column (Dionex) coupled to a CarboPac PA1 2 × 50 mm guard column kept at 30 °C. Cello-oligosaccharides were eluted at 0.25 ml/min using a stepwise linear gradient from 100% eluent A (0.1 M NaOH) toward 10% eluent B (1 M NaOAc in 0.1 M NaOH) 10 min after injection and 30% eluent B 25 min after injection, followed by a 5 min exponential gradient to 100% B. The column was reconditioned between each run by running initial conditions for 9 min.

Analysis by Mass Spectrometry—For time resolved product analysis, electrospray ionization mass spectrometry (ESI-MS) was used with a linear ion trap LTQ Velos Pro (Thermo Scientific, San Jose, CA USA) coupled to an UltiMate 3000 RS UHPLC from Dionex (Sunnyvale, CA USA) which delivered a constant flow and performed injection. No chromatographic separation was employed. The UHPLC delivered a flow of 0.2 ml/min of 30/70 (v/v) H_2O and acetonitrile via the auto-sampler. Standard reaction mixtures were incubated in the thermostatted auto-sampler of the UHPLC at 15 °C during the entire reaction time and samples of 2 μl were injected at given time points. The electrospray was operated in positive mode at 4 kV spray current, with a sheath gas flow of 30 (arbitrary units), an auxiliary gas flow of 5 (arbitrary units) and a capillary temperature of 250 °C. The acquisition time was set to 0.2 min with a data collection time of 10 ms per acquisition. Full scans were performed in the m/z 100–1000 mass range and fragmentation was done using higher energy collisional dissociation (HCD) with N_2 as the collision gas and normalized energy levels of 65, to enable observation of lower mass fragments. During fragmentation, data were collected in the m/z 100–400 mass range. The data were further processed using Xcalibur 2.2 SP1.48 (Thermo Scientific).

NMR Analysis—For NMR analysis, 1.0 mg/ml cellopentaose was dissolved in 99.996% D_2O (Cambridge Isotope Laboratories, Andover, MA) containing 5 mM sodium acetate (pD 6.0) and 0.1 ‰ (v/w) 3-(trimethylsilyl)-propionic-2,2,3,3-d4 acid sodium salt (Aldrich, Milwaukee, WI) used as chemical shift reference for proton and carbon. The NMR tubes contained

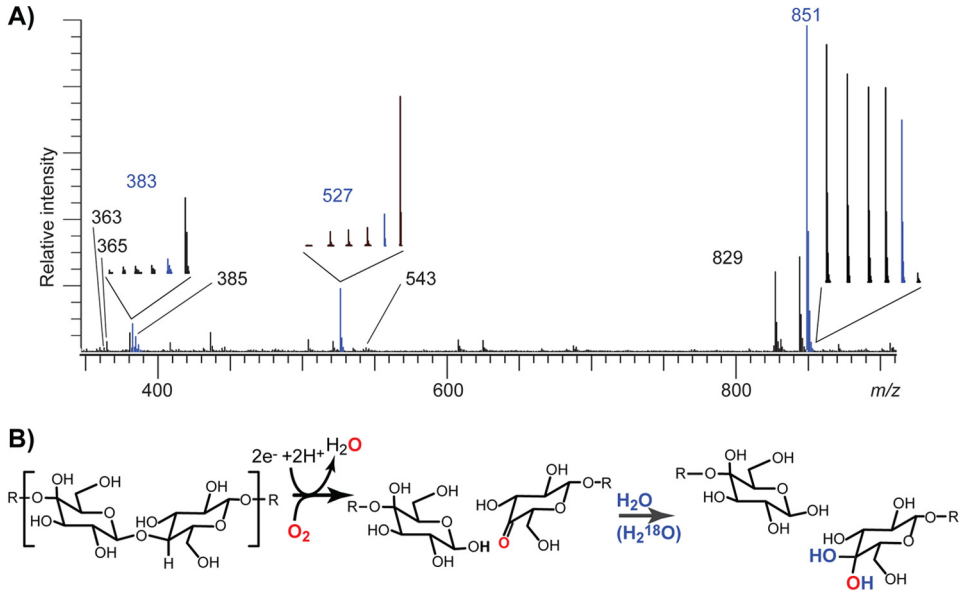

FIGURE 3. *A*, ESI-MS analysis of products generated over time during degradation of Glc_5 (*m/z* 851) in $H_2{}^{18}O$. Reaction products were analyzed after 0, 5, 15, 25, 90, and 1440 min of incubation, and the main picture shows the full spectrum for the 90-min sample. The *insets* show the signal over time for key products (from 0 to 1440 min, from *left* to *right*; the *blue* signal represents the 90-min sample). The *m/z* values 851, 527, and 365 correspond to the sodium adducts of Glc_5, Glc_3, and Glc_2, respectively (*m/z* 829 is $[M+H]^+$ for Glc_5), whereas *m/z* 383 and the minor *m/z* 543 correspond to $[M+Na]^+$ of gemdiol products (Glc4GemGlc and $Glc4GemGlc_2$) carrying one ^{18}O (see Fig. 1; the mass difference between sodium and lithium is 16 Da). The 383 signal is accompanied by a signal at *m/z* 385 because of exchange with the solvent, which gradually leads to incorporation of two ^{18}O atoms. The mass of 363 corresponds to the keto-form of the oxidized dimer ($[M+Na]^+$), and its intensity corresponds to ~10% of the intensity of the gemdiol. *B*, schematic presentation of enzymatic cleavage of the glycosidic bond and introduction of molecular oxygen in the nonreducing end, followed by water incorporation transforming the keto group into a gemdiol. The incorporation of water leads to uptake of ^{18}O in the products.

500 μl of cellopentaose stock solution and 33 μl of either *Mt*CDH (to a final concentration of 0.9 μM) or hydroquinone solution (to a final concentration of 3 or 10 mM). After addition of 17 μl *Nc*LPMO9C (to a final concentration of 2.9 μM), the head space of the NMR tube was flushed with pure oxygen gas (YARA, Trondheim, Norway) for ~10 s before sealing the tube. After incubation of the samples at 25 °C for 24 h, reaction products were analyzed with NMR spectroscopy.

All homo- and heteronuclear NMR experiments were recorded on a Bruker Avance 600 MHz NMR spectrometer (Bruker BioSpin AG, Fällanden, Switzerland) equipped with a 5-mm cryogenic CP-TCI z-gradient probe at 25 °C. For chemical shift assignment, the following spectra were recorded: one-dimensional proton, two-dimensional double quantum filter correlated spectroscopy (DQF-COSY), two-dimensional in-phase correlation spectroscopy (IP-COSY) (23), two-dimensional total correlation spectroscopy (TOCSY) with 70 ms of mixing time, two-dimensional ^{13}C heteronuclear single quantum coherence (HSQC) with multiplicity editing, two-dimensional ^{13}C HSQC-[^1H,^1H]TOCSY with 70 ms of mixing time on protons, and two-dimensional heteronuclear multibond correlation (HMBC) with BRID filter to suppress first order correlation. The NMR data were processed and analyzed with TopSpin 2.1 and TopSpin 3.0 software (Bruker BioSpin).

H_2O_2 Analysis—A fluorimetric assay based on Amplex Red and horseradish peroxidase (17) was used to measure the extent of H_2O_2 generation, which is a futile side reaction catalyzed by the reduced LPMO copper center. The peroxidase catalyzed conversion of Amplex Red to resorufin is proportional to H_2O_2 production (stoichiometry = 1). The increase of fluorescence was measured with an Enspire Multimode plate reader (PerkinElmer Life Sciences) using an excitation wavelength of 569 nm and an emission wavelength of 585 nm. The well plate assay (total volume of 200 μl, 30 °C, 6 min) was performed in 100 mM potassium phosphate buffer, pH 6.0, containing 50 μM Amplex Red, 7.1 units ml^{-1} horseradish peroxidase, 0.87 μM LPMO, and 30 μM ascorbate as reductant in 100 mM sodium phosphate buffer, pH 6.0. Cello-oligosaccharides (DP2–6; Sigma-Aldrich) were added to a final concentration of 5 mM.

Sequence Alignment and Modeling—A structure-guided sequence alignment to compare *Nc*LPMO9C with two structurally characterized AA9-type LPMOs, *Nc*LPMO9D (NCU01050; Protein Data Bank code 4EIR), and *Pc*LPMO9D (*Pc*GH61D; Protein Data Bank code 4B5Q), was constructed using the Expresso mode of the T-Coffee multiple sequence alignment server (24). A homology model of *Nc*LPMO9C was made based on a structure prediction by HHpred (25) and by using Modeler (26) with the crystal structure of *Nc*LPMO9D as a template. All structural comparisons were carried out using PyMOL (PyMOL Molecular Graphics System, version 1.5.0.4; Schrödinger, LLC). The coordinates for cellopentaose were derived from Protein Data Bank entry 2EEX (27).

RESULTS AND DISCUSSION

Enzyme Activity on Cellulose and Cello-oligosaccharides—Initial activity screening of the enzyme was done with a well established HPAEC method for the detection of C1-oxidized cello-oligosaccharide products (12). Incubation of *Nc*LPMO9C with polymeric substrates (PASC, Avicel, steam-exploded spruce; Fig. 2*A*) showed release of soluble cello-oligosaccharides (DP2 and DP3), but C1 oxidized species were not detected. Instead, two dominant later eluting peaks (between

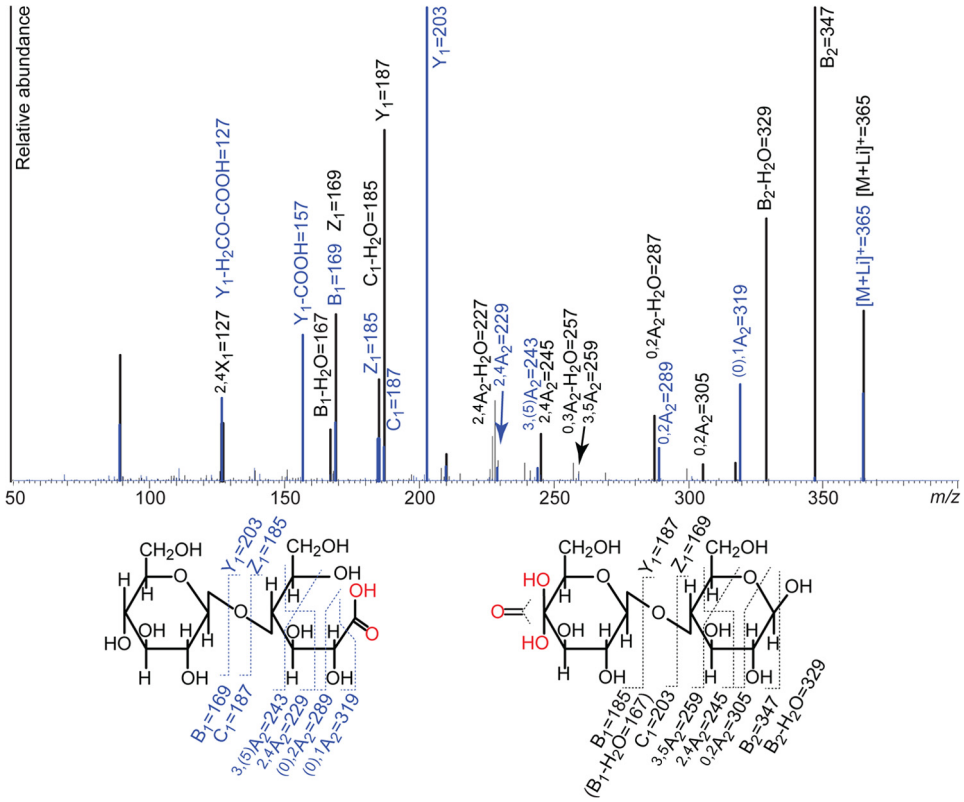

FIGURE 4. **MS/MS analysis of reaction products.** The picture shows overlaid MS/MS spectra obtained for the cellobionic acid generated by *Pc*LPMO9D (*left structure*; *blue trace* and *labeling*) and the oxidized dimer from an *Nc*LPMO9C catalyzed reaction (*right structure*; *black trace* and *labeling*). The analysis was done with lithium doping and all labeled signals reflect lithium adducts.

26 and 29 min) were observed that could represent a different kind of oxidized product. Similar late eluting peaks have also been observed for PASC degradation by *Nc*LPMO9D (8). MS analysis (described in detail below) indicated that the later eluting of these two peaks represented a trimeric product, whereas the earlier peak represented a dimeric product.

Cellulose-active LPMOs characterized so far typically produce oligosaccharides with a DP up to 6 or 7 (5–7, 14). The production of relatively short oligosaccharides by *Nc*LPMO9C could be the result of a glucanase background activity in the enzyme sample or by *Nc*LPMO9C having activity on soluble oligosaccharides. Fig. 2*B* shows that *Nc*LPMO9C indeed is able to oxidatively degrade cello-oligosaccharides. The enzyme readily degraded Glc_5 and Glc_6, while showing lower activity on Glc_4 and no or minute activity on Glc_3. There is abundant evidence for these conversions being caused by *Nc*LPMO9C and not an impurity in the enzyme preparation: 1) the reaction generates oxidized species (see below for detailed characterization), 2) product formation was not observed in the absence of a reducing agent (results not shown), and 3) product formation was not observed in the presence of reducing agent only (results not shown). The oxidized products released were mainly DP2 from Glc_4 and Glc_5 and a mix of DP2 and DP3 from Glc_6 (Fig. 2*B*).

The degradation patterns shown in Fig. 2 (*A* and *B*) were independent of the reductant used (we tested hydroquinone, ascorbic acid, and catechin in concentrations varying from 1.5 to 10 mM). Fig. 2*B* (CDH trace) further shows that use of *Mt*CDH as electron donor for degradation of Glc_5 led to a clear change in the product profile. As expected, native oligosaccharides were no longer observed because they were oxidized to aldonic acids. The oxidized species observed between 26 and 29 min for the reactions with a reducing agent present were not seen, but a new peak appeared at ~40 min that may represent double oxidized species, an interpretation that is supported by mass spectrometry and NMR analysis of the samples (see below). This shift clearly shows that the products eluting between 25 and 30 min cannot be aldonic acids, because the CDH oxidizes reducing ends.

MS analysis confirmed the presence of only two major products upon incubating *Nc*LPMO9C with Glc_5, namely an oxidized dimer and a nonoxidized trimer (see below for details). Nevertheless, the chromatograms (Fig. 2) show additional peaks eluting in between the native and the main oxidized products. We propose that this is due to tautomerization, because keto groups on ring carbons are prone to this process at the elevated pH values used during the HPAEC runs. Migration of the 4-keto group to C3 and maybe C2 will change the elution behavior of the oligosaccharides.

*Nc*LPMO9C is the first LPMO unequivocally shown to be active on soluble cello-oligosaccharides. The fact that the pentamer is degraded faster than the tetramer and the clear preference for releasing an oxidized dimer from Glc_5 indicates the presence of at least five subsites on the enzyme running from −3 to +2 (subsites numbered according to the nomenclature used for glycoside hydrolases) (28). Degradation of Glc_6 yielded both DP2 and DP3 oxidized species, indicating binding to −4 to +2 and −3 to +3. Notably, the location of these subsites and

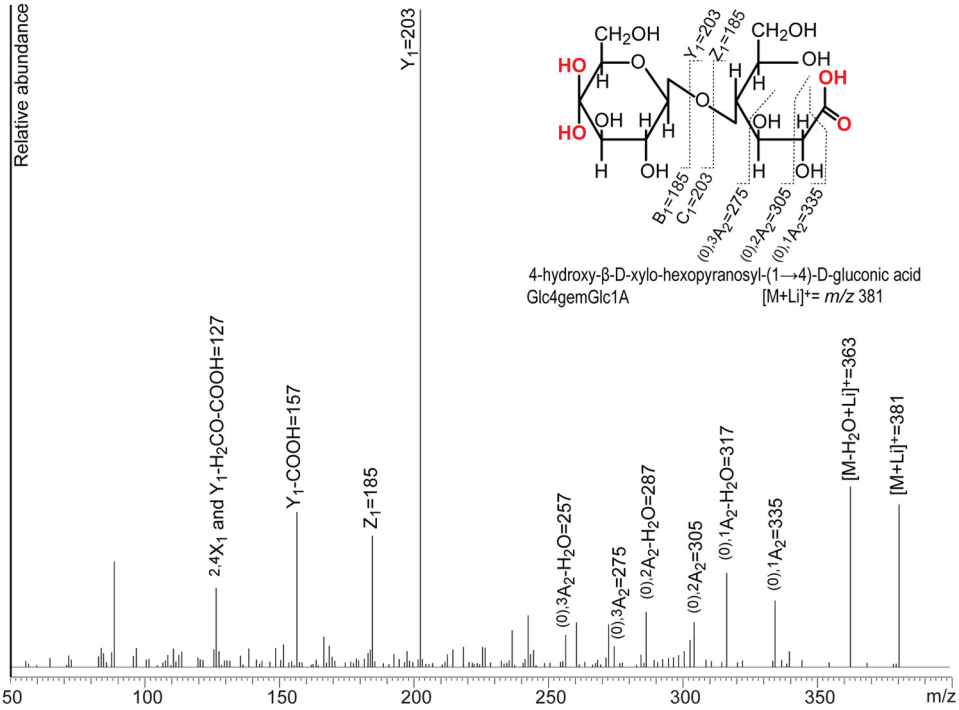

FIGURE 5. **MS/MS analysis of a double oxidized dimer (Glc4gemGlc1A) generated by treating cellopentaose with *Nc*LPMO9C and *Mt*CDH.** The analysis was done with lithium doping, and all labeled signals reflect lithium adducts.

the orientation of the productively bound substrate relative to the catalytic center are currently unknown.

To test whether the cello-oligosaccharide oxidizing activity of *Nc*LPMO9C is specific or an unspecific side reaction, the enzyme was incubated with other oligosaccharides (Man_6, Xyl_5, Xyl_6, chitopentaose, and maltodextrin) and both crystalline α-chitin and nanofibrillar α- and β-chitin. No activity was observed on any of these substrates. Thus, *Nc*LPMO9C seems to be specific for β-1–4-linked glucose units.

Product Identification by MS—MS analysis of product mixtures showed that the main products from degradation of Glc_5 were two species with *m*/*z* values of 527 and 381, corresponding to the sodium adducts of Glc_3 and a cellobiose with a gemdiol (*i.e.*, a hydrated keto group) at the nonreducing end, respectively. Minor amounts of cellobiose and oxidized cellotriose were also detected. To investigate this further, the time course of the enzyme reaction was studied in $H_2{}^{18}O$. This resulted in the same signal of *m*/*z* 527, whereas a new peak at *m*/*z* 383 appeared (Fig. 3*A*). This shows that the oxygen of the glycosidic linkage remains in the Glc_3 product as the OH group at C1, whereas the oxidized product acquires one oxygen atom from water and one from oxygen as schematically shown in Fig. 3*B*. As indicated in Fig. 3*A*, a *m*/*z* 385 peak appears after some time, which is due to the lactone-gemdiol equilibrium leading to exchange with water and, thus, to the eventual incorporation of two ^{18}O atoms.

Because the glycosidic bond in cello-oligosaccharides (and cellulose) links the C1 of one glucose to the C4 of the adjacent glucose, it is logical to suggest that the oxidation carried out by *Nc*LPMO9C takes place at the C4 position of the nonreducing end moiety. The resulting keto sugar will be in equilibrium with the C4 gemdiol in water solution (a feature that is common to keto saccharides (29)). Generally, it is not straightforward to prove the position of LPMO generated oxidations using mass spectrometry because the masses of various possible products are identical and because the mass difference between sodium and potassium adducts equals the mass of an oxygen atom. As shown in Fig. 1, the aldonic acid and gemdiol forms have identical masses, as do the corresponding lactone and keto forms. Thus, MS analysis alone cannot determine the type of LPMO activity. Exploiting the low complexity product mixtures obtainable thanks to the activity of *Nc*LPMO9C on soluble substrates, we have addressed this issue by carrying out MS/MS analyses of products after lithium doping, which facilitates reducing end cross-ring fragmentation (30), as well as by NMR analyses.

Fig. 4 shows the result of MS/MS analysis of the *m*/*z* 365 $[M+Li]^+$ products (oxidized dimers) generated by *Nc*LPMO9C and C1 oxidizing *Pc*LPMO9D (7). The spectra show differences including the occurrence of fragment ions that are diagnostic for C1 or C4 oxidations. Cellobionic acid (fragments indicated in *blue* in Fig. 4) readily loses mass corresponding to one carboxyl group (*m*/*z* 319). The C4 oxidized dimer (fragments indicated in *black* in Fig. 4) readily loses masses corresponding to both one and two water molecules (*m*/*z* 347 and 329) but does not generate ions corresponding to the loss of a carboxyl group. In addition, the most prevalent Y_1 and Z_1 ions (31) from glycosidic bond cleavage will be different for the two oxidized compounds. The Y_1 and Z_1 ions for glycosidic bond cleavage of cellobionic acid have *m*/*z* values of 203 and 185, respectively, species that include the carboxylic acid. The Y_1 and Z_1 ions for the C4 oxidized dimer have *m*/*z* values of 187 and 169, respectively, meaning no oxidation in the reducing end. The corresponding B_1 and C_1 ions (for the

TABLE 1

Fragmentation table for MS/MS analysis of native and oxidized cellobiose and cellotriose compounds

The table shows the *m/z* values of sodium and lithium adducts of the different fragments. n.a., not analyzed. # represents $[M+2Metal-H]^+$. Double ox represents a double oxidized oligosaccharide with a gemdiol in the nonreducing end and an aldonic acid in the downstream end. Fragmentation nomenclature according to Domon and Costello (31).

		Mother ions		Major fragments						
DP	Metal		$[M+metal]^+$	$-H_2O$	$-2H_2O$	B_1/B_2	C_1/C_2	Y_1/Y_2	Z_1/Z_2	–COOH
2	Na	Native	365	347	n.a	185	203	203	185	n.a
		–Glc1A	381	363	n.a	185	203	219	201	335
		–Glc1A #	403	385	n.a	207	225	241	223	357
		Glc4gem–	381	363	345	201	219	203	185	n.a
		Double ox	397	379	361	201	219	219	201	351
2	Li	Native	349	331	n.a	169	187	187	169	n.a
		–Glc1A	365	347	n.a	169	187	203	185	319
		–Glc1A #	371	353	n.a	175	193	209	191	325
		Glc4gem–	365	347	329	185	203	187	169	n.a
		Double ox	381	363	345	185	203	203	185	335
DP	Metal		$[M+metal]^+$	$-H_2O$	$-2H_2O$	B_1/B_2	C_1/C_2	Y_1/Y_2	Z_1/Z_2	–COOH
3	Na	Native	527	509	n.a	185/347	203/365	203/365	185/347	n.a
		–Glc1A	543	525	n.a	185/347	203/365	219/381	201/363	497
		–Glc1A #	565	547	n.a	207/369	225/387	241/403	223/385	519
		Glc4gem–	543	525	507	201/363	219/381	203/365	185/347	n.a
		Double ox	559	541	523	363	381	381	363	513
3	Li	Native	511	493	n.a	169/331	187/349	187/349	169/331	n.a
		–Glc1A	527	509	n.a	169/331	187/349	203/365	185/347	481
		–Glc1A #	533	515	n.a	175/337	193/355	209/371	191/353	487
		Glc4gem–	527	509	491	185/347	203/365	187/349	169/331	n.a
		Double ox	543	525	507	185/347	203/365	203/365	185/347	497

		Mother ions		Cross ring cleavages					
DP	Metal		$[M+metal]^+$	$^{0,2}A_2$	$^{2,4}A_2$	$^{3,5}A_2$	$^{0,2}X_1$	$^{2,4}X_1$	$^{3,5}X_1$
2	Na	Native	365	305	245	259	245	305	291
		–Glc1A	381	305	245	259	261	321	307
		–Glc1A #	403	327	267	281	283	343	329
		Glc4gem–	381	321	261	275	245	305	291
		Double ox	397	321	261	275	261	321	307
2	Li	Native	349	289	229	243	229	289	275
		–Glc1A	365	289	229	243	245	305	291
		–Glc1A #	371	295	235	249	251	311	297
		Glc4gem–	365	305	245	259	229	289	275
		Double ox	381	305	245	259	245	305	291
DP	Metal		$[M+metal]^+$	$^{0,2}A_2$	$^{2,4}A_2$	$^{3,5}A_2$	$^{0,2}X_2$	$^{2,4}X_2$	$^{3,5}X_2$	$^{0,2}A_3$	$^{2,4}A_3$	$^{3,5}A_3$	$^{0,2}X_3$	$^{2,4}X_3$	$^{3,5}X_3$
3	Na	Native	527	305	245	259	245	305	291	467	407	421	407	467	453
		–Glc1A	543	305	245	259	261	321	307	467	407	421	423	483	469
		–Glc1A #	565	327	267	281	283	343	329	489	429	443	445	505	491
		Glc4gem–	543	321	261	275	245	305	291	483	423	437	407	467	453
		Double ox	559	321	261	275	261	321	307	483	423	437	423	483	469
3	Li	Native	511	289	229	243	229	289	275	451	391	405	391	451	437
		–Glc1A	527	289	229	243	245	305	291	451	391	405	407	467	453
		–Glc1A #	533	295	235	249	251	311	297	457	397	411	413	473	459
		Glc4gem–	527	305	245	259	229	289	275	467	407	421	391	451	437
		Double ox	543	305	245	259	245	305	291	467	407	421	407	467	453

C4 oxidized dimer) are *m/z* 185 and 203, respectively, and tend to lose a water molecule, generating *m/z* 167 and 185 species, respectively. Loss of a water molecule from the B_1 and C_1 ions seems a likely event considering the presence of the gemdiol group at C4.

Notably, comparison of the *blue* and *black spectra* in Fig. 4 shows that MS/MS analysis can reveal whether an LPMO oxidizes the reducing or the nonreducing end of the substrate. Distinct *m/z* ions for cellobionic acid are 319, 203, and 157. For Glc4gemGlc, distinct *m/z* ions are 347, 329, and 287. A rela-

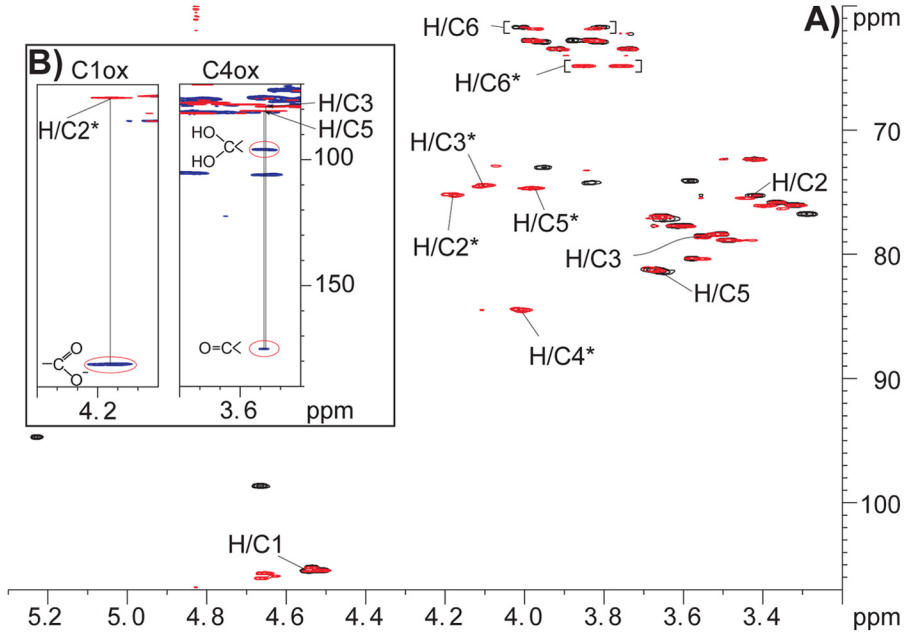

FIGURE 6. **NMR analysis of reaction products.** *A*, overlay of ^{13}C HSQC spectra for reaction products generated by treating 0.9 mg/ml cellopentaose with 2.9 μM *Nc*LPMO9C in the presence of 10 mM hydroquinone (black signals) or 0.9 μM *Mt*CDH (*red signals*). The samples were both in 99.996% D_2O with 5 mM sodium acetate, pH 6.0, and spectra were recorded at 25 °C. Peaks in the proton/carbon signals of the C4 oxidized monosaccharide residue are marked by H/C#, where # refers to the ring carbon number (overlapping *red* and black *signals*). Peaks in the proton/carbon signals of the C1 oxidized monosaccharide residue are marked by H/C#* (only *red signals*). For the sake of simplicity, peaks related to nonoxidized monosaccharide residues are not marked (a full assignment of chemical shifts is provided in Table 2). *B*, two details of an overlay of a ^{13}C HSQC spectrum and a ^{13}C HMBC spectrum recorded for products obtained in a reaction with both *Nc*LPMO9C and CDH. The *left panel* shows a correlation (indicated by a *vertical line*) from the H/C2* peak in HSQC (*red*) to a peak with a carbon chemical shift of 181.1 ppm in HMBC (*blue*), corresponding well to the (expected) presence of a carboxylate group at position C1. The *right panel* shows correlation from the H/C3 and H/C5 peaks in HSQC (*red*) to carbon peaks with a carbon chemical shift at C4 of 95.9 and 175.2 ppm in HMBC (*blue*), corresponding to the presence of a gemdiol and a keto group at C4, respectively.

tively high *m*/*z* 187 signal is also typical for Glc4gemGlc. In the case CDH is used as the electron donor for a C4-oxidizing LPMO, double oxidized products will emerge (Fig. 2*B*, *CDH trace*) with yet another characteristic MS/MS pattern, as illustrated in Fig. 5. Lithium and sodium adducts of MS/MS fragments obtained for the different LPMO products are summarized in Table 1.

Product Identification by NMR—To obtain proof for the identity of the oxidized reaction products, NMR spectroscopy was used to analyze products generated from cellopentaose by *Nc*LPMO9C in the presence of either CDH or hydroquinone (Fig. 6). The individual monosaccharide residues were assigned by starting at the anomeric signal and/or at the primary alcohol group at C6 and then following the proton-proton connectivity using TOCSY, DQF-COSY/IP-COSY, and ^{13}C HSQC-[^{1}H,^{1}H]TOCSY, whereas connectivity between the individual monosaccharide residues was obtained from the HMBC spectrum (see Table 2 for assignment of chemical shifts). These experiments showed the presence of dimeric and trimeric products in both reactions, as expected from the MS analyses described above. An overlay of the ^{13}C HSQC spectra of product mixtures obtained in the presence of hydroquinone or CDH (Fig. 6*A*) readily shows different occurrence of oxidations. C4 oxidation occurs in both reactions, whereas signals reflecting C1 oxidation are only observed in the reaction with CDH (*red signals* in Fig. 6*A*). There were no indications of oxidation of C6 because this part of the ^{13}C HSQC spectrum was essentially identical to that observed for nontreated cellopentaose and because no novel signals possibly reflecting additional oxidations were observed.

^{13}C HMBC spectra provided further insight into the nature of the products. This is illustrated by Fig. 6*B*, showing an overlay of the ^{13}C HSQC and the ^{13}C HMBC spectra of products obtained in the reaction with *Nc*LPMO9C and CDH. The overlay shows correlations from H/C-5 and H/C-3 to peaks with a carbon chemical shift at C4 of 95.9 and 175.2 ppm, corresponding well to the chemical shift of a gemdiol (29) and keto group, respectively. The gemdiol and keto groups account for ~80 and 20% of the signal intensity, respectively. Although documentation on the keto:gemdiol ratio in literature is limited, the gemdiol would be expected to dominate at pH 6.0 because the keto group is easily hydrated (29) in aqueous medium. The overlay also shows a correlation from the H/C-2 to a peak with a carbon chemical shift of 181.1 ppm that corresponds well to the shift expected when a carboxylate group is present at position C1 (12). Even though there are two different Glc1A groups in the product mixture (one in a trimeric and one in a dimeric product), their chemical shifts are very similar, meaning that the peaks appear nearly at the same position in the spectra, looking like one broad peak (see Table 2 for more details). Thus, for the first time, NMR has been used to prove that the products generated by a nonreducing end active LPMO are oxidized in the C4 position and that the products primarily exist as a gemdiol.

Suppression of H_2O_2 Production—In the absence of a cellulosic substrate, activated LPMOs produce H_2O_2 (17). We employed this ability to demonstrate binding of cello-oligo-

TABLE 2

Assignment of chemical shifts

The individual monosaccharide residues were assigned by starting at the anomeric signal and/or at the primary alcohol group at C6 and then following the proton-proton connectivity using TOCSY, DQF-COSY/IP-COSY, and ^{13}C HSQC-TOCSY spectra. ^{13}C HSQC was used for assigning the carbon chemical shifts. The ^{13}C HMBC spectrum provided long range bond correlations, allowing identification of the connectivity between the sugar units and allowing identification of the carboxyl group at C1, as well as the gem-diol group and keto group at C4. Chemical shifts were assigned for the major products resulting from treatment of 0.9 mg/ml cellopentaose with 2.9 μM *Nc*LPMO9C in the presence of 0.9 μM *Mt*CDH or 10 mM hydroquinone in 99.996% D_2O with 5 mM sodium acetate pD 6.0, for spectra recorded at 25 °C. Numbers 1–6 represent ring carbon numbers to which the chemical shift values (^1H, ^{13}C or ^1H, ^1H, ^{13}C for C6) are assigned for a nonreducing end glucose (NR), a second after NR glucose (SNR), the α anomer of glucose (α), the β anomer of glucose (β), the aldonate glucose (C1), and the keto/gemdiol glucose (C4). The chemical shifts reported are with accuracy of 0.01 ppm for ^1H and 0.1 ppm for ^{13}C, based on spectral resolution.

Glucose unit/carbon number	1	2	3	4	5	6
Hydroquinone, Glc_3						
NR	4.51; 105.4	3.32; 76.0	3.49; 78.6	3.51; 78.4	3.42; 72.3	3.93; 3.73; 63.5
SNR	4,54; 105.1	3.37; 75.8	3.67; 77.0	3.62; 81.2	3.61; 77.7	3.97; 3.83; 62.8
α	5.22; 94.7	3.57; 74.1	3.83; 74.1	3.65; 81.5	3.95; 73.0	3.96; 3.83; 62.8
β	4.66; 98.6	3.29; 76.8	3.65; 77.6	3.61; 81.3	3.63; 77.1	3.98; 3.87; 62.8
Hydroquinone, Glc4gemGlc						
C4	4.55; 105.5	3.41; 75.3	3.55; 78.5	175.2/95.9	3.58; 80.4	4.02; 3.81; 61.7
α	5.22; 94.7	3.57; 74.1	3.83; 74.1	3.65; 81.5	3.95; 73.0	3.96; 3.83; 62.8
β	4.66; 98.6	3.29; 76.8	3.65; 77.6	3.61; 81.3	3.63; 77.1	3.98; 3.87; 62.8
***Mt*CDH, Glc_2Glc1A**						
NR	4.51; 105.3	3.32; 76.1	3.42; 78.4	3.50; 78.3	3.49; 72.3	3.91; 3.74; 63.4
SNR	4.65; 105.6	3.41; 76.1	3.66; 76.9	3.67; 81.4	3.60; 77.8	3.97; 3.82; 62.8
C1	181.1	4.18; 75.2	4.10; 74.4	4.01; 84.51	3.99; 74.7	3.85; 3.76; 64.8
***Mt*CDH, Glc4gemGlc1A**						
C4	4.53; 105.3	3.44; 75.5	3.55; 78.6	175.2/95.9	3.65; 80.3	3.82; 3.98; 61.8
C1	181.1	4.18; 75.2	4.10; 74.4	4.01; 84.51	3.99; 74.7	3.85; 3.76; 64.8

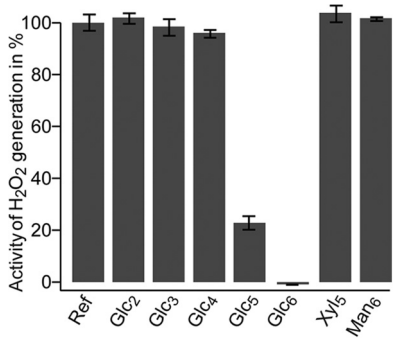

FIGURE 7. **Generation of H_2O_2 by *Nc*LPMO9C.** The enzyme was incubated with reductant and the production of H_2O_2 (here referred to as activity) was measured as described under "Experimental Procedures." The figure shows that the enzyme generates H_2O_2 in the absence of oligosaccharides (*Ref*) and in the presence of Xyl_5, Man_6, or short cello-oligosaccharides, whereas cello-oligosaccharide with DP > 4 diminish H_2O_2 production, indicative of productive binding to the LPMO. The reported results are mean values of five experiments. Control experiments without the reductant or the LPMO showed no H_2O_2 production, regardless of the presence of cello-oligosaccharides.

meric substrates to *Nc*LPMO9C. Fig. 7 shows that production of H_2O_2 was diminished in the presence of cello-oligosaccharides with a DP >4. Although only minor inhibition was observed in the presence of Glc_4, almost complete inhibition of H_2O_2 formation was observed in the presence of Glc_6. Thus, the data in Fig. 7 show that cello-oligomers with a minimal length of five sugars form complexes with the enzyme that are sufficiently strong to suppress the futile H_2O_2 generating side reaction. This corresponds very well with the HPLC data, which showed high activity for Glc_5 and Glc_6 and much lower activity for Glc_4 (Fig. 2*B*). As mentioned above, *Nc*LPMO9C is not active on Xyl_5 or Man_6, and Fig. 7 shows that these substrates do not suppress the H_2O_2 production.

Sequence and Structural Model of NcLPMO9C—For comparison, the sequence of *Nc*LPMO9C was aligned to a C1 oxidizing and a nonreducing end oxidizing LPMO. The AA9 domain of *Nc*LPMO9C shares 36.6% sequence identity with C1-oxidizing *Pc*LPMO9D (7) and 47.5% with nonreducing end oxidizing *Nc*LPMO9D (8, 16) (Fig. 8*A*). The metal coordinating residues (His-1 and His-83 in *Nc*LPMO9C) as well as four residues surrounding the copper site (Asn-26, His-155, Gln-164, and Tyr-166 in *Nc*LPMO9C) are conserved in all three AA9s. Looking at the LPMO structures available, AA9s tend to contain several surface-exposed aromatic residues that seem to be aligned to interact with a cellulose chain that would then traverse the catalytic center (32). Indeed, these aromatic residues have roughly the same spatial orientation as residues that, by experiment, have been shown to interact with chitin in an AA10 LPMO (33). This adds confidence to the notion that these aromatic residues interact with the substrate, possibly analogous to what is been suggested for *Pc*LPMO9D (32).

Based on the crystal structure of *Nc*LPMO9D a model was built for *Nc*LPMO9C. Comparison of the surface-exposed residues potentially involved in substrate binding (Fig. 8*B*) shows that there are more aromatic residues on the surfaces of *Nc*LPMO9D and *Pc*LPMO9D compared with *Nc*LPMO9C. Because of this difference, the length of the substrate-binding surface seems shorter in *Nc*LPMO9C. Even if one considers the contribution of protruding polar residues on the surface, the binding surface of *Nc*LPMO9C seems less extended (Fig. 8*B*) and possibly more adapted to binding shorter substrates compared with *Pc*LPMO9D and *Nc*LPMO9D. Interestingly, the structural model of *Nc*LPMO9C indicates that this enzyme has a cluster of three asparagine residues (Asn-25, Asn-26, and Asn-27) in a location that could potentially be a +2 subsite. One could speculate that these asparagines interact with the reducing end sugar. Indeed, as illustrated in Fig. 8*B*, a cellopentaose would span the putative binding surface of *Nc*LPMO9C, when its reducing end is positioned near these asparagines. Such an orientation would be in accordance with the observation that cellopentaose primarily is cleaved into a cellotriose and a Glc4gemGlc.

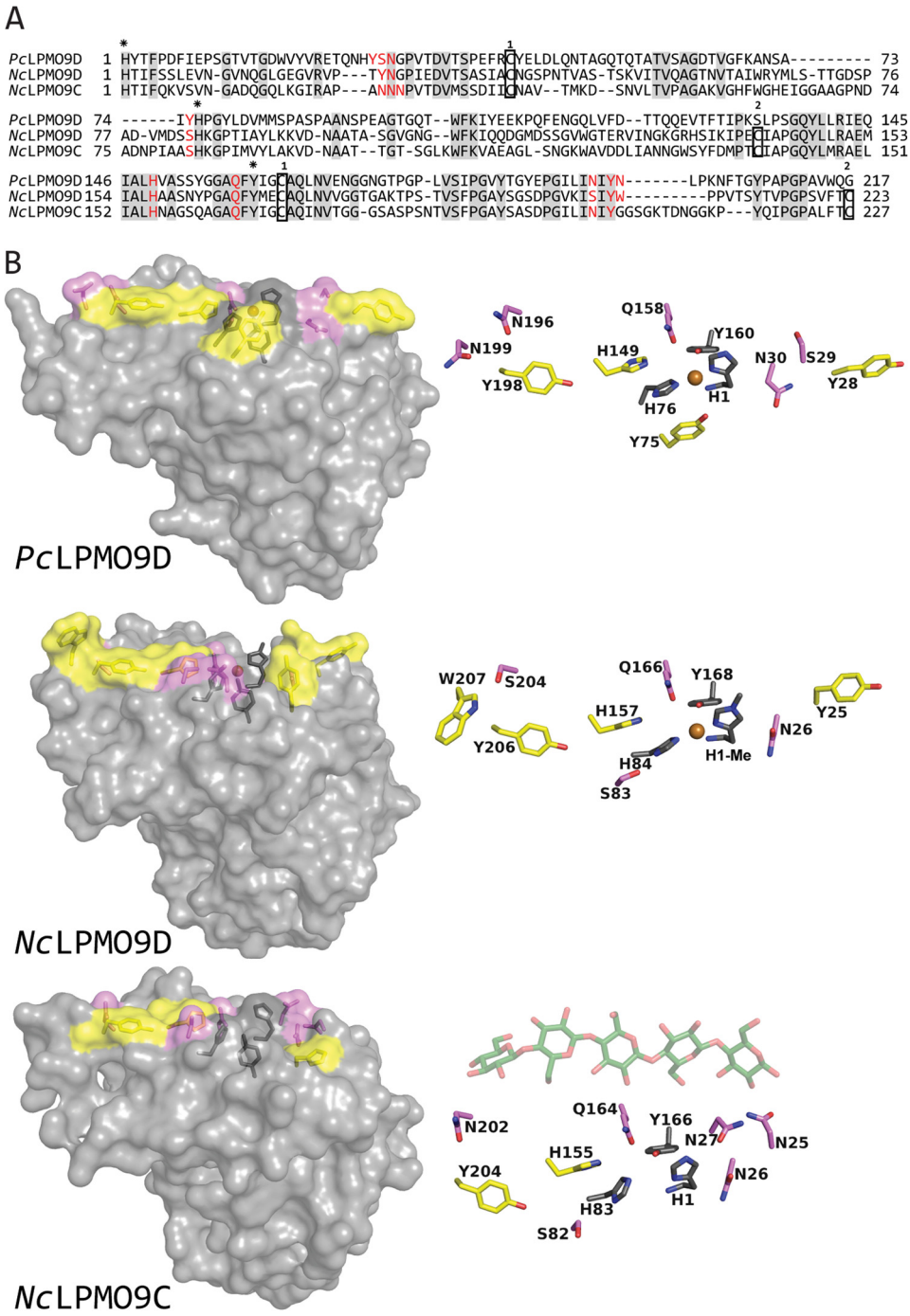

FIGURE 8. **Sequence and structural comparisons of *Pc*LPMO9D, *Nc*LPMO9D and *Nc*LPMO9C.** *A*, structure-guided sequence alignment of *Pc*LPMO9D, *Nc*LPMO9D, and *Nc*LPMO9C with fully conserved residues shaded in *gray*. Residues involved in coordination of the copper are marked with *asterisks* above the sequence (His-1, His-83 and Tyr-166 in *Nc*LPMO9C), whereas aromatic surface residues and protruding polar surface residues potentially involved in substrate binding are colored in red (all highlighted in *B*). Cysteines forming disulfide bonds are indicated in *boxes* with *numbers* above showing which cysteines are connected. *B*, side chains on the substrate binding surface of *Pc*LPMO9D (Protein Data Bank code 4B5Q) and *Nc*LPMO9D (Protein Data Bank code 4EIR) compared with the modeled structure of *Nc*LPMO9C. Copper coordinating residues are colored *dark gray*, aromatic and protruding polar residues putatively involved in substrate binding are colored *yellow* and *violet*, respectively, and the copper is shown as an *orange sphere*. For illustration purposes only, a cellopentaose (coordinates derived from Protein Data Bank code 2EEX; shown in *green* with oxygens in *red*) is placed above the surface of the modeled structure of *Nc*LPMO9C. The view in the *right panels* is rotated 90° relative to the view in the *left panels*, looking down at the flat surface containing the copper binding site.

Conclusions—In this article, we have shown that *Nc*LPMO9C cleaves both crystalline cellulose as well as cello-oligosaccharides yielding products oxidized in the nonreducing end. By applying NMR, we unambiguously showed that the nonreducing end sugar was oxidized at the C4 position and that this sugar primarily exists as a gemdiol in solution. We have also shown that MS/MS fragmentation analysis of oxidized products can be applied to differentiate between C1 and C4 oxidizing LMPOs. A range of different substrates were tested, but *Nc*LPMO9C was only active on β-1,4-linked glucose units, and the enzyme

seemed to require a minimum stretch of at least four glucose units. It has not escaped our attention that such an activity on short cello-oligosaccharides could imply that the enzyme is active on hemicellulose structures containing β-1,4-linked glucose units.

Acknowledgments—We thank Prof. Johannes Vliegenthart and Dr. Derek Horton for suggesting IUPAC nomenclature for the enzyme products. We thank Morten Skaugen for help in managing our HPLC and MS systems.

REFERENCES

1. Horn, S. J., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V. G. (2012) Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5,** 45
2. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active enzymes database (CAZy). An expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238
3. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6,** 41
4. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V. G. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330,** 219–222
5. Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S. J., and Eijsink, V. G. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20,** 1479–1483
6. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., and Walton, P. H. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 15079–15084
7. Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G., Igarashi, K., Samejima, M., Ståhlberg, J., Horn, S. J., and Sandgren, M. (2011) The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One* **6,** e27807
8. Phillips, C. M., Beeson, W. T., Cate, J. H., and Marletta, M. A. (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6,** 1399–1406
9. Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., and Sweeney, M. D. (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* **77,** 7007–7015
10. Sygmund, C., Kracher, D., Scheiblbrandner, S., Zahma, K., Felice, A. K., Harreither, W., Kittl, R., and Ludwig, R. (2012) Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Appl. Environ. Microbiol.* **78,** 6161–6171
11. Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J., and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135,** 6069–6077
12. Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrøm, Y. H., and Eijsink, V. G. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J. Chromatogr. A* **1271,** 144–152
13. Bey, M., Zhou, S., Poidevin, L., Henrissat, B., Coutinho, P. M., Berrin, J.-G., and Sigoillot, J.-C. (2013) Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (Family GH61) from *Podospora anserina*. *Appl. Environ. Microbiol.* **79,** 488–496
14. Beeson, W. T., Phillips, C. M., Cate, J. H., and Marletta, M. A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134,** 890–892
15. Tian, C., Beeson, W. T., Iavarone, A. T., Sun, J., Marletta, M. A., Cate, J. H., and Glass, N. L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 22157–22162
16. Li, X., Beeson, W. T., 4th, Phillips, C. M., Marletta, M. A., and Cate, J. H. (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* **20,** 1051–1061
17. Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., and Ludwig, R. (2012) Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol. Biofuels* **5,** 79
18. Zámocký, M., Schumann, C., Sygmund, C., O'Callaghan, J., Dobson, A. D., Ludwig, R., Haltrich, D., and Peterbauer, C. K. (2008) Cloning, sequence analysis and heterologous expression in *Pichia pastoris* of a gene encoding a thermostable cellobiose dehydrogenase from *Myriococcum thermophilum*. *Protein Expr. Purif.* **59,** 258–265
19. Flitsch, A., Prasetyo, E. N., Sygmund, C., Ludwig, R., Nyanhongo, G. S., and Guebitz, G. M. (2013) Cellulose oxidation and bleaching processes based on recombinant *Myriococcum thermophilum* cellobiose dehydrogenase. *Enzyme Microb. Technol.* **52,** 60–67
20. Harreither, W., Sygmund, C., Augustin, M., Narciso, M., Rabinovich, M. L., Gorton, L., Haltrich, D., and Ludwig, R. (2011) Catalytic properties and classification of cellobiose dehydrogenases from Ascomycetes. *Appl. Environ. Microbiol.* **77,** 1804–1815
21. Wood, T. M. (1988) Preparation of crystalline, amorphous, and dyed cellulase substrates. *Methods Enzymol.* **160,** 19–25
22. Ifuku, S., and Saimoto, H. (2012) Chitin nanofibers. Preparations, modifications, and applications. *Nanoscale* **4,** 3308–3318
23. Xia, Y., Legge, G., Jun, K. Y., Qi, Y., Lee, H., and Gao, X. (2005) IP-COSY, a totally in-phase and sensitive COSY experiment. *Magn. Reson. Chem.* **43,** 372–379
24. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso. Automatic incorporation of structural information in multiple sequence alignments using 3D-coffee. *Nucleic Acids Res.* **34,** W604–W608
25. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33,** W244–W248
26. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815
27. Kitago, Y., Karita, S., Watanabe, N., Kamiya, M., Aizawa, T., Sakka, K., and Tanaka, I. (2007) Crystal structure of Cel44A, a glycoside hydrolase family 44 endoglucanase from *Clostridium thermocellum*. *J. Biol. Chem.* **282,** 35703–35711
28. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321,** 557–559
29. Kopper, S., and Freimund, S. (2003) The composition of keto aldoses in aqueous solution as determined by NMR spectroscopy. *Helv. Chim. Acta* **86,** 827–843
30. Cancilla, M. T., Wong, A. W., Voss, L. R., and Lebrilla, C. B. (1999) Fragmentation reactions in the mass spectrometry analysis of neutral oligosaccharides. *Anal. Chem.* **71,** 3206–3218
31. Domon, B., and Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* **5,** 397–409
32. Wu, M., Beckham, G. T., Larsson, A. M., Ishida, T., Kim, S., Payne, C. M., Himmel, M. E., Crowley, M. F., Horn, S. J., Westereng, B., Igarashi, K., Samejima, M., Ståhlberg, J., Eijsink, V. G., and Sandgren, M. (2013) Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the basidiomycota fungus *Phanerochaete chrysosporium*. *J. Biol. Chem.* **288,** 12828–12839
33. Aachmann, F. L., Sørlie, M., Skjåk-Braek, G., Eijsink, V. G., and Vaaje-Kolstad, G. (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 18779–18784

Paper II

Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation

Jane W. Agger[a], Trine Isaksen[a,1], Anikó Várnai[a,1], Silvia Vidal-Melgosa[b], William G. T. Willats[b], Roland Ludwig[c], Svein J. Horn[a], Vincent G. H. Eijsink[a,2], and Bjørge Westereng[a,b,d]

[a]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), 1432 Aas, Norway; [b]Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen,1871 Frederiksberg, Denmark; [c]Department of Food Sciences and Technology, BOKU–University of Natural Resources and Life Sciences, 1190 Vienna, Austria; and [d]Department of Geosciences and Natural Resource Management, Faculty of Science, University of Copenhagen, 1958 Frederiksberg C, Denmark

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved March 21, 2014 (received for review December 20, 2013)

The recently discovered lytic polysaccharide monooxygenases (LPMOs) are known to carry out oxidative cleavage of glycoside bonds in chitin and cellulose, thus boosting the activity of well-known hydrolytic depolymerizing enzymes. Because biomass-degrading microorganisms tend to produce a plethora of LPMOs, and considering the complexity and copolymeric nature of the plant cell wall, it has been speculated that some LPMOs may act on other substrates, in particular the hemicelluloses that tether to cellulose microfibrils. We demonstrate that an LPMO from *Neurospora crassa*, *Nc*LPMO9C, indeed degrades various hemicelluloses, in particular xyloglucan. This activity was discovered using a glycan microarray-based screening method for detection of substrate specificities of carbohydrate-active enzymes, and further explored using defined oligomeric hemicelluloses, isolated polymeric hemicelluloses and cell walls. Products generated by *Nc*LPMO9C were analyzed using high performance anion exchange chromatography and multidimensional mass spectrometry. We show that *Nc*LPMO9C generates oxidized products from a variety of substrates and that its product profile differs from those of hydrolytic enzymes acting on the same substrates. The enzyme particularly acts on the glucose backbone of xyloglucan, accepting various substitutions (xylose, galactose) in almost all positions. Because the attachment of xyloglucan to cellulose hampers depolymerization of the latter, it is possible that the beneficial effect of the LPMOs that are present in current commercial cellulase mixtures in part is due to hitherto undetected LPMO activities on recalcitrant hemicellulose structures.

biorefinery | metallo enzymes | GH61 | CBM33

Plant cell walls constitute the largest source of biomass on earth and have evolved over some 450 million years to provide support and protection to the plant body. The effective deconstruction of these highly complex and recalcitrant structures is dependent on efforts to discover and understand the batteries of enzymes evolved by microorganisms for cell wall degradation. One important class of enzymes is the copper-dependent lytic polysaccharide monooxygenases (LPMOs) (1–6). Several LPMOs have been characterized, and for all of these activity on either cellulose or chitin has been demonstrated, indicating that LPMOs act on highly crystalline substrates (1, 4, 5, 7, 8). LPMOs cleave the β-(1→4) glycosidic bonds in these substrates, leaving the C1 or the C4 carbon oxidized (1, 9–12). The only exception so far emerged in a recent study showing that the strictly C4-oxidizing *Nc*LPMO9C from *Neurospora crassa* is active on water-soluble, cellulose-derived oligosaccharides (12). This finding suggests that *Nc*LPMO9C, and possibly other LPMOs, might be active on other noncrystalline polysaccharides, for example hemicelluloses, which are important structural components of plant cell walls (13).

Hemicelluloses are among the most abundant and diverse class of plant cell wall polysaccharides and are typically based on a backbone of β-(1→4) glycosidic linkages. In contrast to cellulose, hemicelluloses are usually substituted with a variety of sugar residues and the backbone may be interspersed with other linkage types instead of being strictly β-(1→4) as in cellulose (13, 14). Both these structural features prevent chain-association and crystallization, and enable hemicelluloses to form a highly resistant coextensive load-bearing network with cellulose microfibrils (10, 15). Several studies have shown that xylan and xyloglucan can adsorb to cellulose, forming a partial outer coating to microfibrils, which likely provides protection against cell wall deconstruction by glycoside hydrolases (16, 17).

Our current understanding of microbial degradation of the cellulose/hemicellulose network envisages that a wide variety of glycoside hydrolase enzymes (GHs) degrade hemicelluloses, whereas LPMOs target crystalline cellulose which is then susceptible to further degradation by cellulases. However, the vast sequence variation among LPMOs (10, 11) and the fact that multiple LPMOs are often cosecreted by biomass degrading microorganisms (18–20) are inconsistent with cellulose and chitin being the only LPMO substrates. It seems likely that additional LPMO activities exist, and if that is the case, this would be a significant new insight into the biological roles of LPMOs and

Significance

Plant cell walls are recalcitrant copolymeric structures mainly comprising polysaccharides and lignin. Enzymatic degradation of the polysaccharides is a crucial step in biorefining of biomass. Recently, it was discovered that nature employs copper-dependent redox enzymes called lytic polysaccharide monooxygenases (LPMOs) to promote degradation of the most recalcitrant and crystalline of these polysaccharides, cellulose. By carrying out oxidative cleavage of otherwise inaccessible cellulose chains, LPMOs create access points for classical hydrolytic enzymes such as cellulases. Intriguingly, the genomes of biomass degrading microorganisms encode a plethora of LPMOs (up to over 40). To our knowledge, we demonstrate for the first time that LPMOs act on hemicelluloses. This finding dramatically widens the scope of LPMOs and oxidative processes in plant cell wall degradation and biorefining.

Author contributions: J.W.A., S.V.-M., W.G.T.W., V.G.H.E., and B.W. designed research; J.W.A., T.I., A.V., S.V.-M., and B.W. performed research; J.W.A., S.V.-M., W.G.T.W., R.L., and B.W. contributed new reagents/analytic tools; J.W.A., A.V., W.G.T.W., S.J.H., V.G.H.E., and B.W. analyzed data; and J.W.A., T.I., A.V., S.V.-M., W.G.T.W., R.L., S.J.H., V.G.H.E., and B.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]T.I. and A.V. contributed equally to this manuscript.

[2]To whom correspondence should be addressed. E-mail: vincent.eijsink@nmbu.no.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323629111/-/DCSupplemental.

also extend the enzyme toolbox available for studying plant cell wall structures and valorization.

We report here important findings about the activity of *Nc*LPMO9C on a number of α- and β-linked polysaccharides, obtained using glycan microarray substrate screening combined with chromatography and mass spectrometry. We show unequivocal activity against the biologically important noncellulosic soluble polysaccharides tamarind xyloglucan, mixed linkage (1→3),(1→4)-β-D-glucan and, more weakly, against glucomannan. Furthermore, we demonstrate, to our knowledge for the first time, that *Nc*LPMO9C is active on two naturally occurring and structurally distinct xyloglucan substrates from *Arabidopsis thaliana* and *Solanum lycopersicum* (tomato) and also on (1→3),(1→4)-β-D-glucan from the lichen *Cetraria islandica*. These findings have significant implications both for our understanding of *Nc*LPMO9C in nature and for our utilization of LPMOs for plant biomass degradation.

Results

Activity Screening of *Nc*LPMO9C Using Glycan Microarrays. We used glycan microarray technology (21) to rapidly screen *Nc*LPMO9C toward an extensive collection of putative substrates including the most abundant polysaccharides present in plant and fungal cell walls (Fig. 1 and Fig. S1). This analysis clearly showed that *Nc*LPMO9C was active on xyloglucan, (1→3),(1→4)-β-D-glucan, and, to a lesser extent, also on glucomannan, and that activity was increased in the presence of the reductant. These data indicate that *Nc*LPMO9C requires short stretches of contiguous (1→4)-β–linked glucose units for activity. Indeed, there was no indication of activity on polysaccharides not containing some degree of (1→4)-β-D-glucan, exemplified by (1→4)-β-D-xylan, (1→4)-α-D-polygalacturonan, (1→4)-β-D-mannan (without glucose in the backbone), (1→4)-β-D-galactan, yeast β-D-glucan [(1→3),(1→6)-β-D-glucan], callose [(1→3)-β-D-glucan], and arabinoxylan (Fig. 1 and Fig. S1).

Activity on Xyloglucan. The data obtained from the glycan microarray screening indicate that *Nc*LPMO9C acts on (1→4)-β-D-glucans, independent of substitution of the polysaccharide backbone (as in, for example, xyloglucan). We selected a set of biologically relevant polysaccharides to confirm the findings from the microarrays and to analyze the products. We compared the product profiles generated by *Nc*LPMO9C to those generated by two endo-glucanases, *Ta*Cel5A (22) and *Af*Cel12A (23).

Fig. 2*A* shows that *Nc*LPMO9C and *Af*Cel12A generated different products from tamarind xyloglucan. Products generated by *Nc*LPMO9C include two major oxidized products corresponding to the keto-form of single oxidized species, the most probable product configurations being $G^{ox}XXL$ (*m/z* 1,245.3) and $G^{ox}XLL$ (*m/z* 1,407.3). Note that the relative position of the X and L units can vary [ref. 24; we use the generally accepted nomenclature of xyloglucan according to ref. 24, where G = β-D-Glc; X = α-D-Xyl-(1→6)-β-D-Glc; L = β-D-Gal-(1→2)-α-D-Xyl-(1→6)-β-D-Glc; F = α-L-Fuc-(1→2)-β-D-Gal-(1→2)-α-D-Xyl-(1→6)-β-D-Glc); S = α-L-Ara*f*-(1→5)-α-D-Xyl-(1→6)-β-D-Glc]. These assignments were made based on the known primary structure of tamarind xyloglucan with a nonsubstituted G appearing every fourth residue in the glucan backbone (25). The appearance of oxidized products indicates that a single glucose unit in the (1→4)-β-D-glucan backbone is sufficient for the LPMO to access and cleave xyloglucan. In comparison, the reaction with *Af*Cel12A generated native products only appearing as *m/z* 1,085.1, 1,247.2, and 1,409.2, most probably corresponding to XXXG, XLXG/XXLG and XLLG, respectively (25). These assignments were made based on the notion that glycoside hydrolases degrading xyloglucan tend to release oligosaccharides with an unsubstituted glucose unit at the reducing end (24).

Although Fig. 2*A* shows the generation of oxidized products from xyloglucan by *Nc*LPMO9C, the assignment of oligosaccharide sequences to the observed products is somewhat arbitrary, due to the heterogeneity of the substrate, and leaves questions as to where exactly the LPMO and the endo-glucanase cleave. Therefore, we studied the degradation of $XXXGXXXG^{OH}$ ($XG14^{OH}$), a reasonably pure 14-mer generated by endo-glucanase treatment of tamarind xyloglucan, with known sequence and with a reduced glucose (glucitol) in the reducing end (+2 Da). Reaction products were analyzed using both high-performance anion exchange chromatography (HPAEC) (Fig. S2 and ref. 26) and multidimensional mass spectrometry (MS) (Fig. 3 *A* and *B*), again revealing clear differences between *Nc*LPMO9C and *Af*Cel12A product profiles. The product profile generated by *Nc*LPMO9C shows that the lytic reaction primarily happened at the nonsubstituted internal glucosyl unit and caused an oxidation at the nonreducing end, since oxidations were only observed on fragments

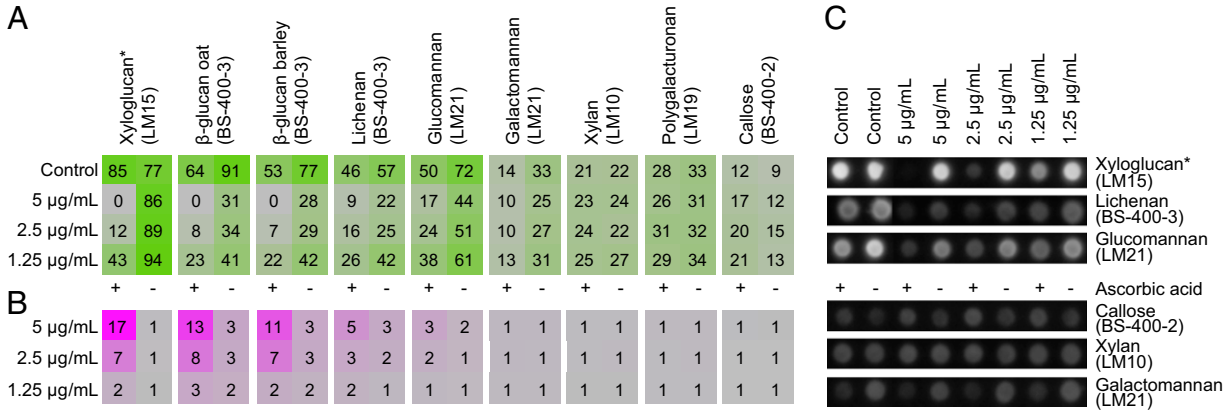

Fig. 1. Glycan microarray screening of *Nc*LPMO9C activity on various polysaccharide substrates. Activities were detected by the loss or reduction of polysaccharide-borne epitopes recognized by substrate-specific monoclonal antibodies (mAb). (*A*) Heatmap showing degradation of the substrates listed above the panel following treatment with *Nc*LPMO9C at different concentrations (listed left), in the presence (+) or absence (–) of ascorbic acid. Control means no enzyme treatment. The mAbs used to recognize the substrates are in parentheses, top row. The heatmap shows relative mean spot signal intensities; the lower the signal when comparing controls and *Nc*LPMO9C treated samples, the greater the degree of degradation. (*B*) Fold-change heatmap of the data shown in *A*. This heatmap shows the degree of change in mAb binding resulting from *Nc*LPMO9C treatment of the substrates. A cutoff of 5 was imposed before the fold changes were calculated. (*C*) Images of the arrays used to produce the heatmaps, showing consistent spot morphologies and depleted signals caused by *Nc*LPMO9C. * denotes that final concentration of xyloglucan was 1 mg/mL, whereas the other substrates were 0.1 mg/mL.

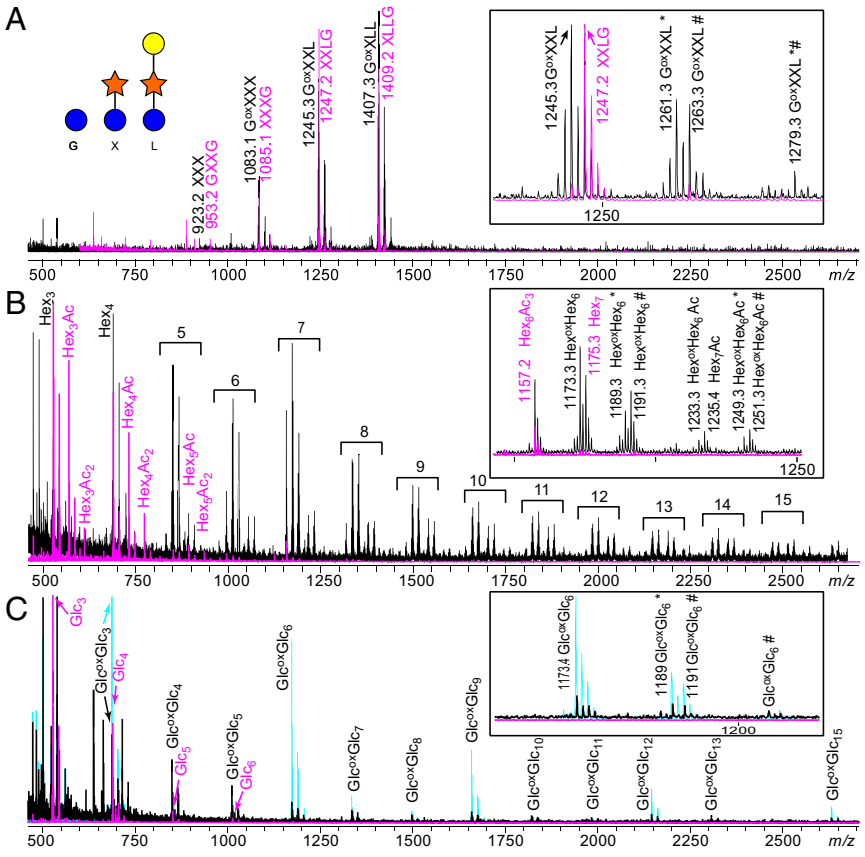

Fig. 2. MALDI-ToF MS analysis of product profiles. The spectra show products generated from tamarind xyloglucan (*A*; see *Inset* regarding nomenclature; blue, glucose; orange, xylose; yellow, galactose), konjac glucomannan (*B*), lichenan and mixed linked glucan (MLG) from barley (*C*) treated with either endo-glucanase (magenta) or LPMO (black, turquoise for MLG). Product profiles upon endo-glucanase treatment of lichenan and MLG were essentially the same and therefore only one of the spectra is shown. Brackets in *B* indicate product clusters of same DP, indicated by the number. In the main spectra, only sodium adducts are labeled, whereas the inserts also show potassium adducts (marked *) and various forms of oxidized species where both the keto-group formed upon C4 oxidation (−2 Da) and its gemdiol form (marked #, i.e., addition of H_2O, +18 Da) appear. Abbreviations: G, X. and L, see *A Inset*; Glc, glucose (+ 162 Da); Hex, hexose (+ 162 Da); Ac, acetyl group (+ 42 Da); ox, oxidized (−2 Da if keto form). Analysis of control reactions, without enzyme addition, showed no signals related to carbohydrates.

carrying the reduced terminal glucose (Fig. 3 *A* and *B*). Accordingly, the only native fragment detected in the reaction with *Nc*LPMO9C corresponds to XXX. The endo-glucanase only hydrolyzed $XG14^{OH}$ at the reducing end of the internal nonsubstituted glucosyl (G) unit, generating two products, XXXG (*m/z* 1,085) and $XXXG^{OH}$ (*m/z* 1,087) (Fig. 3 *A* and *B* and Fig. S2). The experiments with $XG14^{OH}$ confirmed that xyloglucan cleavage by *Nc*LPMO9C occurs at a single unsubstituted glucosyl unit.

Activity on Glucomannan and β-Glucan. The activity of *Nc*LPMO9C on konjac glucomannan, a linear (1→4)-β–linked mannan comprising randomly distributed glucose (~35%; ref. 27), generated clusters of products up to at least DP15, each cluster containing the keto-, gemdiol and native form of hexose oligosaccharides and acetylated species (Fig. 2*B*). The fact that relatively high DP species appeared indicates that larger parts of the substrate could not be cleaved, which is likely due to the occurrence of mannoses, because galactomannan is not cleaved by the enzyme (Fig. 1). The endo-glucanase *Ta*Cel5A seemed rather efficient in degrading konjac glucomannan, generating primarily DP3 and DP4 and the corresponding mono- and diacetylated species, i.e., a totally different product profile compared with *Nc*LPMO9C. The differences may be due to *Ta*Cel5A being less affected by the high mannose content of the backbone or by the enzyme being able to hydrolyze (1→4)-β-D-mannan linkages (28).

MALDI-ToF analysis of products generated from barley β-glucan and lichenan (Fig. 2*C*) confirmed *Nc*LPMO9C activity on these substrates. The MALDI-ToF spectrum of barley β-glucan showed dominance of products increasing by three sugars in size, i.e., DP4, DP7, DP10, DP13, and so forth (Fig. 2*C*), which correlates with the known repetitive structure of this polysaccharide, from barley with every third linkage being a (1→3)-β-D linkage (29). Apparently, (1→3)-β-D linkages affect the substrate's suitability for cleavage by *Nc*LPMO9C. Products generated from lichenan did not show this repetitive pattern, indicating that the organization of its backbone is less repetitive. Activity of *Nc*LPMO9C seemed lower on lichenan compared with barley β-glucan, which is also indicated by the heatmap in Fig. 1*B*.

In-Depth Analysis of Substrate Specificity of *Nc*LPMO9C on Xyloglucan. The assignment of MS signals to certain products with certain sugar compositions is somewhat arbitrary, due to the many overlapping masses and contaminations in $XG14^{OH}$ (Fig. 3*B*). The data discussed so far demonstrate occurrence of oxidation and show that *Nc*LPMO9C can accommodate a considerable degree of substitution of the glycan backbone close to the cleavage point. In fact, the data discussed suggest that every glucose near the cleavage point may be substituted, except, perhaps, the one that becomes oxidized. To gain further insight into these matters, MS2 experiments were conducted. Fig. 3*C* shows MS2 fragmentation of an ion with *m/z* 1,249 generated by *Nc*LPMO9C from $XG14^{OH}$ and assigned in Fig. 3*A* as $G^{ox}XXXG^{OH}$ (*m/z* 1,249 represents the Li-adduct, whereas *m/z* 1,265, in Fig. 3*A*, is a Na-adduct). Many signals in the spectrum in Fig. 3*C* provide little structural information because they result from (dominating) fragmentation reactions leading to removal of substitutions (primarily xyloses, −132/150 Da) from the glucan backbone. However, the presence or absence of several key ions does provide useful information. First, the *m/z* 853/835 species represents a pentahexose carrying both an oxidized and a reduced end; this can only be a backbone glucose pentamer, confirming that *Nc*LPMO9C cuts at the nonreducing end of the intermediate G-unit of $XG14^{OH}$ and produces $G^{ox}XXXG^{OH}$, as indicated in Fig. 3 *A* and *B*. Second, the occurrence of an *m/z* 939 fragment implies the loss of an oxidized X-unit, whereas *m/z* 1,069 implies the loss of a terminal native hexose. The latter ion can only arise if an L-unit,

BIOCHEMISTRY

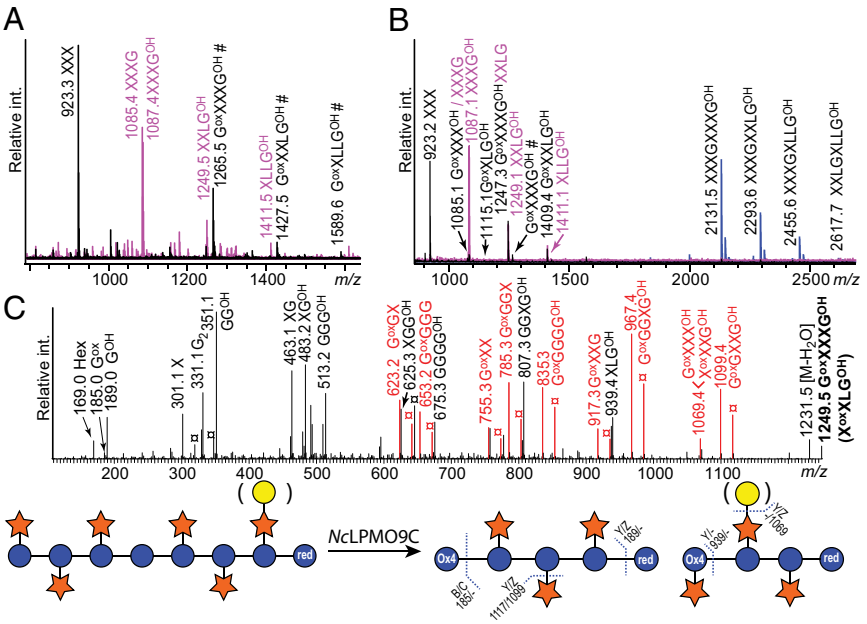

Fig. 3. Analysis of products generated from XG14^{OH}. (*A*) ESI-MS spectra highlighting Na-adducts of products generated by *Nc*LPMO9C (black) or *Af*Cel12A (magenta). ESI-MS spectra show primarily the gemdiol form (marked #) of oxidized products. (*B*) MALDI-ToF-MS spectra highlighting Na-adducts of XG14^{OH} (blue) and of products generated from XG14^{OH} by *Nc*LPMO9C (black) or *Af*Cel12A (magenta). MALDI-ToF-MS spectra show primarily the keto-form of oxidized products, which in the case of a reduced substrate has the same *m/z* as native products (keto group gives *m/z* -2, and reduction gives *m/z* +2). Note that panel B shows that XG14^{OH} (XXXGXXXG^{OH}) is contaminated with other species containing one or more additional hexoses, probably galactoses coupled to one or more of the X units as this is a very common moiety in xyloglucan from tamarind (hence annotation as L in the figure). Some products derived from these contaminations are annotated in the mass spectra. (*C*) MS2 fragmentation of an *m/z* 1,249 species generated upon lithium doping of the product mixture shown in panel A (*m/z* 1,249 corresponds to the Li-adduct of the *m/z* 1,265 species in *A*). Gemdiol-fragments readily lose a water molecule thus occurring in the spectrum as B or C −18 Da (fragmentation nomenclature as in ref. 30); this has been observed previously for gemdiol products (12). The substrate, XG14^{OH}, and possible products of *m/z* 1,249 are shown as cartoons according to the nomenclature of (31): blue circle, glucose; orange star, xylose; yellow circle, galactose. Parenthesis surrounding galactosyl-units denote that the position of these units may vary. "Ox" denotes the position of the oxidation. "Red" denotes the position of reduction. Note that dominating fragmentation reactions lead to removal of substitutions from the glucan backbone, explaining why several oligo-G products are detected.

which has a terminal galactose, is present within the *m/z* 1,249 product pool from XG14^{OH}, which again would mean that the oxidation is carried on an X-unit, as indeed indicated by the presence of the *m/z* 939 species (so the *m/z* 1,249 ion would then be X^{ox}XLG^{OH}, where the position of the L may vary). Fragmentation of the corresponding ion (*m/z* 1,247) generated from a xyloglucan polymer also supports that oxidation of substituted glucose may occur (Fig. S3). All in all, the data show that *Nc*LPMO9C is promiscuous when it comes to accepting substitutions in or close to the catalytic site. The only absolute requirement, indicated by the substrate screening experiments described above, is that the backbone contains (1→4)-β-D-glucan.

Previous studies have shown that *Nc*LPMO9C has low activity on cellotetraose, yielding dimeric products (12), and this may be considered a minimal substrate. We addressed both the effect of substitution and the type of glycosidic bond by testing the enzyme on substrates with four backbone glucosyl units. It was not capable of cleaving between two X-units in XG7 (XXXG), but was able to very slowly convert this substrate to XXX and G^{ox} (Fig. S4). Similarly, the activity on mixed linked cellotetraosyls was also very low and the data clearly showed that 1→3 linkages close to the catalytic site limited enzyme activity (Fig. S4). Of the tetraoses G4G3G4G and G4G4G3G, only the latter was cleaved slowly and only between the first two glucoses (producing G + G^{ox}4G3G).

To evaluate the relative importance of xyloglucan and cellulose as substrates, we compared degradation rates for cellopentaose versus XG14^{OH} and polymeric xyloglucan versus phosphoric acid swollen cellulose (Fig. S5). The oligomeric substrates may be considered as "minimal" substrates as they have only one major cleavage point (Fig. 3*A* and ref. 12). These experiments showed that the enzyme degraded XG14^{OH} approximately twice as fast as cellopentaose (0.06 s^{-1} and 0.03 s^{-1}, respectively, at 40 °C). A similar difference was not observed when using the polymeric substrates; in this case, the rates were very similar (0.11 s^{-1} at 50 °C).

***Nc*LPMO9C Activity on Plant Cell Walls.** To explore the activity of *Nc*LPMO9C on hemicellulose polymers in biologically relevant contexts, we then studied product release from various plant materials. Fig. 4 shows that *Nc*LPMO9C released a variety of oxidized xyloglucan fragments from *Arabidopsis thaliana* and tomato (*Solanum lycopersicum*), including fragments with more complex substitution patterns than those obtained with the model substrates described above, i.e., fragments containing fucosylations, acetylations, and arabinosyl substitutions. Acetylations in *Arabidopsis* xyloglucan are positioned on the galactose units that in addition can be further substituted by a fucosyl unit (32). Acetylations in the Solanaceae may be positioned either directly on the glucan backbone or on arabinosyl units (33). The product mixtures thus further illustrate the ability of *Nc*LPMO9C to accept a variety of substitutions. Although the mass spectra clearly show the occurrence of oxidation in the case of the LPMO only, they do not allow firm conclusions about differences in substrate specificity between the LPMO and the endo-glucanase, due to the many overlapping masses.

The product profiles obtained with Icelandic moss (Fig. S6) showed clear differences between *Nc*LPMO9C and the endo-glucanase (*Ta*Cel5A), similar to what was seen for purified lichenan (Fig. 2).

Discussion

The discovery of LPMO activity was a significant breakthrough in our understanding of how nature degrades recalcitrant biomass and has already had significant implications for industrial enzymatic conversion of lignocellulosic biomass. Until recently, LPMO activity had been demonstrated for chitin and cellulose only. The "flat" substrate-binding surfaces of these enzymes (3, 34, 35) suggest that they are indeed optimized for binding the surfaces of crystalline materials, rather than single polysaccharide chains. However, several observations indicate that LPMOs may have a wider role in biomass degradation. First, LPMOs also act on rather amorphous insoluble substrates such as phosphoric acid swollen cellulose. Second, the genomes of biomass-degrading microbes often contain many LPMO-encoding genes (over 40 in some cases; ref. 10), and expression studies show that the upregulation of these genes varies with the substrates, including

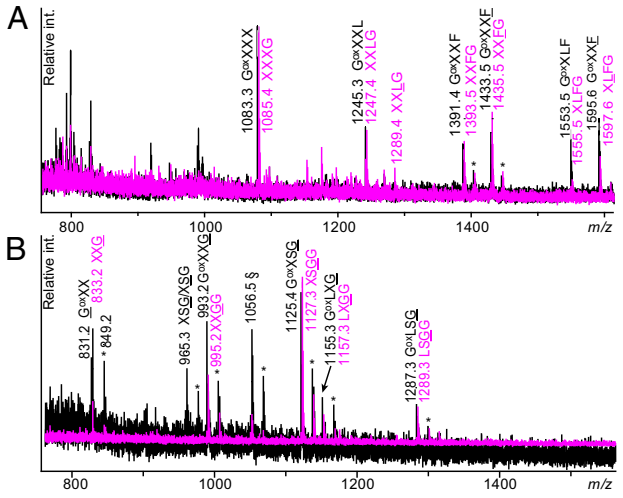

Fig. 4. MALDI-ToF spectra of products generated from natural substrates. *Arabidopsis* (*A*) or tomato stem (*B*) cell wall material was incubated with either *Nc*LPMO9C (black line) or *Af*Cel12A (magenta line) and product mixtures were analyzed. Peaks were assigned by combining the mass information with knowledge on the substrate and the preferences of the enzymes in question; note that these assignments to some extent are arbitrary, because the exact positions of modifications/substitutions cannot always be inferred. Also note that additional sugars and modifications occur in these natural substrates: F, fucosylation on an L-unit (+146 Da); S, arabinosyl substitution on an X-unit (+132 Da). Acetylations (+42 Da) are also observed and marked by underlined G, F, or S units. All assigned masses are sodium adducts, except when indicated otherwise. Potassium adducts are marked by an asterisk. §, background signal. Control reactions showed that ascorbic acid alone did not release any products and did not change the product profile generated by the endo-glucanase.

hemicelluloses (20, 36, 37). Third, apart from a conserved catalytic center, LPMOs show large sequence variation (10, 11). All these observations indicate that LPMO activity on plant cell walls may not be restricted to cellulose but could include a variety of other polysaccharides, including xylan, xyloglucan, mannan, pectin and even starch. In this study, we provide, to our knowledge, the first evidence for LPMO activity on xyloglucans, β-glucans, and glucomannan.

By using high-throughput glycan microarray technology, we were able to rapidly screen putative hemicellulosic substrates, leading to the discovery of hitherto unidentified activities. This initial screening effort was then followed by detailed product analysis using model substrates and, importantly, also natural plant materials. Although quantification of substrate degradation is difficult, our data indicate that the primary hemicellulose targeted by *Nc*LPMO9C is xyloglucan. The data also indicate that the enzyme is able to accommodate a wide variety of modifications on the glucan backbone. The structural basis of the substrate-specificity of *Nc*LPMO9C needs further research. Interestingly, a structural model of *Nc*LPMO9C suggests that the enzyme has a more extended and more polar substrate-binding surface relative to *Pc*LPMO9D, which is not active on xyloglucan. The additional protruding polar side chains in *Nc*LPMO9C may engage in hydrogen bonding with the carbohydrate moieties of branched substrates such as xyloglucan (Fig. S7).

A main biological function of hemicellulose is to tether cellulose microfibrils, thereby strengthening the cell wall (13). Xyloglucan in particular plays a major role in building the load-bearing network in type I primary plant cell walls and interacts very closely with cellulose (14, 16, 38). The biological implications of *Nc*LPMO9C having broad substrate specificity on xyloglucan could be to enable *N. crassa* to "shave off" hemicelluloses and thereby increase accessibility to otherwise recalcitrant regions of cellulose. Perhaps these enzymes attack junction zones between cellulose and hemicellulose, which may not be good substrates for glycoside hydrolases (39), because the activity of these enzymes depends on interactions with accessible single polysaccharide chains. LPMOs may be better suited to attack these zones because they are adapted to act on surfaces.

To our knowledge, the present study provides the first example of LPMO activity on hemicellulose and it is likely that more such activities will be discovered in the future. Considering the present observations and the fact that LPMO-encoding genes are abundant in biomass-degrading microorganisms, it seems likely that LPMO-catalyzed oxidative processes play a major role in natural biomass conversion. Considering the broad specificity on xyloglucan, it is intriguing that *Nc*LPMO9C is not active on xylan because the (1→4)-β-D-xylan structure in many aspects resembles the (1→4)-β-D-glucan structure. Xylans are another prominent and structurally important class of hemicelluloses and it seems plausible that other LPMOs are targeted toward this polysaccharide.

Another important implication of the present findings concerns current worldwide efforts in establishing efficient methods for enzymatic saccharification of lignocellulosic biomass. Recent improvements in the efficiency of commercial cellulase mixtures are in part due to the addition of LPMOs (40). The present findings raise questions as to why these LPMOs actually work in practice. The LPMO effect may of course be due to higher efficiency of the degradation of cellulose itself (2, 41), but might also be due to hitherto nondetected LPMO activities on recalcitrant hemicellulose structures that block cellulase access (38). Notably, whereas cellulose is homogeneous and similar in all types of plant biomass, hemicelluloses vary a lot within the plant kingdom. So, if hemicellulose degradation by LPMOs already is or eventually will be a factor determining the efficiency of commercial enzyme mixtures for biomass saccharification, it is likely that the LPMO content of such mixtures needs to be adapted to the substrate.

Materials and Methods

Enzymes. Endo-glucanases *Ta*Cel5A from *Thermoascus aurantiacus* (UniProt: Q8TG26) and *Af*Cel12A from *Aspergillus fumigatus* (UniProt:Q8TG26) were expressed in *Pichia pastoris* (*ade2*, *prb1*, *pep4*) under the AOX1 promoter using the PichiaPink expression system (Invitrogen, Life Technologies), and purified according to procedures described in *SI Materials and Methods*. Lytic polysaccharide monooxygenase 9C from *Neurospora crassa*, *Nc*LPMO9C, was expressed in *Pichia pastoris* X33 and subsequently purified according to Kittl et al. (42).

Microarray-Based Detection of LPMO Activity. ***Microarray construction.*** *Nc*LPMO9C was screened against nine polysaccharide substrates: xyloglucan (tamarind), β-glucan (oat), β-glucan (barley), lichenan (icelandic moss), glucomannan (konjac), galactomannan (carob), polygalacturonic acid, and callose (1,3-β-D-glucan) (Megazyme International Ireland) and birch wood xylan (X0502, Sigma-Aldrich). Additional substrates tested are described in *SI Materials and Methods*. Solutions were prepared in 100 mM sodium phosphate buffer, pH 8.0, and added into a microtiter plate in the following order (final concentrations in parenthesis): polysaccharide (0.1 mg/mL, except xyloglucan which was 1 mg/mL), ascorbic acid (2 mM) and *Nc*LPMO9C (5 μg/mL, 2.5 μg/mL or 1.25 μg/mL) solutions. Each reaction was performed in triplicate. The plate containing the reaction mixtures was incubated for 2 h at 50 °C following 10 min at 80 °C. Afterward, printing buffer (55.2% glycerol, 44% water, 0.8% Triton X-100) was added in each well as 1/4 of the final volume. Reaction mixtures were then printed using a piezoelectric non-contact microarray robot (Sprint, Arrayjet) onto a nitrocellulose membrane with a pore size of 0.45 μm (Whatman).

Probing and analysis of microarrays. Nitrocellulose microarrays were probed as described (21). Briefly, the printed arrays were probed with a panel of anti-rat and anti-mouse monoclonal antibodies (PlantProbes and Biosupplies). Antibodies were diluted in PBS containing 5% wt/vol low fat milk powder (MPBS) to 1/10 and 1/1,000, respectively. For secondary antibodies, anti-rat and anti-mouse secondary antibodies conjugated to alkaline phosphatase (Sigma) were diluted in MPBS to 1/5,000. A complete list of antibodies may be found in the supplementary information. Developed microarrays were scanned at 2,400 dpi (CanoScan 8800F), converted to TIFFs and signals were

measured using appropriate software (Array-Pro Analyzer 6.3, Media Cybernetics). Data are shown in a heatmap where color intensity is correlated to mean spot signal value. A cut off of 5 arbitrary units was applied. A fold-change heatmap was produced by calculating the following ratio: control signal:enzyme-treated signal.

Substrates for Enzyme Reactions. The following substrates were used for exploring enzymatic activities of the LPMO and endo-glucanases: xyloglucan (XG) from tamarind seed (*Tamarindus indica*); glucomannan from konjak (*Amorphophallus konjac*); mixed linked β-glucan from barley (*Hordeum vulgare*); lichenan from Icelandic moss (*Cetraria islandica*); cellopentaose (Glc_5); xyloglucan-heptamer (XG7) with known composition XXXG; a reduced 14-mer ($XG14^{OH}$) derived from tamarind xyloglucan mainly containing $XXXGXXXG^{OH}$ (the reducing end glucose is reduced to a glucitol); two types of mixed linked cellotetraosyl (G4G4G3G and G4G3G4G). All of these substrates were from Megazyme International Ireland. Finally, enzymes were tested against three natural sources of mixed linked β-glucan and xyloglucan: a suspension of ball-milled Icelandic moss and alcohol insoluble residues from *Arabidopsis thaliana* and from stems of tomato (*Solanum lycopersicum*).

Enzyme Reactions. Reaction mixtures of 200-μL total volume contained 0.4 mg/mL substrate, 0.05 mg/mL enzyme and 1 mM ascorbic acid in 10 mM Na-acetate, pH 6.0. Samples were incubated at 40 °C with shaking at 600 rpm for 18 h. Products were analyzed using MALDI-ToF-MS, direct infusion ESI-MS or HPAEC (*SI Materials and Methods*). Determination of enzyme rates is described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. This work was funded by Norwegian Research Council Grants 216162 and 214613 and by the Seventh Framework Programme of the European Union through the project Waste2Go under Contract 308363 with the European Commission.

1. Vaaje-Kolstad G, et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330(6001):219–222.
2. Forsberg Z, et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20(9): 1479–1483.
3. Aachmann FL, Sørlie M, Skjåk-Bræk G, Eijsink VG, Vaaje-Kolstad G (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc Natl Acad Sci USA* 109(46): 18779–18784.
4. Phillips CM, Beeson WT, Cate JH, Marletta MA (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol* 6(12):1399–1406.
5. Quinlan RJ, et al. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci USA* 108(37):15079–15084.
6. Wu M, et al. (2013) Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem* 288(18):12828–12839.
7. Westereng B, et al. (2011) The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE* 6(11):e27807.
8. Bey M, et al. (2013) Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospora anserina*. *Appl Environ Microbiol* 79(2):488–496.
9. Beeson WT, Phillips CM, Cate JHD, Marletta MA (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc* 134(2): 890–892.
10. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH (2012) Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels* 5(1):45.
11. Li X, Beeson WT, 4th, Phillips CM, Marletta MA, Cate JH (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* 20(6):1051–1061.
12. Isaksen T, et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem* 289(5):2632–2642.
13. Scheller HV, Ulvskov P (2010) Hemicelluloses. *Annu Rev Plant Biol* 61(61):263–289.
14. Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat Chem Biol* 6(10):724–732.
15. Somerville C, et al. (2004) Toward a systems approach to understanding plant cell walls. *Science* 306(5705):2206–2211.
16. Vincken JP, de Keizer A, Beldman G, Voragen AGJ (1995) Fractionation of xyloglucan fragments and their interaction with cellulose. *Plant Physiol* 108(4):1579–1585.
17. Hayashi T, Marsden MPF, Delmer DP (1987) Pea xyloglucan and cellulose: VI. Xyloglucan-cellulose interactions in vitro and in vivo. *Plant Physiol* 83(2):384–389.
18. Galagan JE, et al. (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422(6934):859–868.
19. Znameroski EA, et al. (2012) Induction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. *Proc Natl Acad Sci USA* 109(16):6012–6017.
20. Floudas D, et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.
21. Pedersen HL, et al. (2012) Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research. *J Biol Chem* 287(47):39429–39438.
22. Parry NJ, et al. (2002) Biochemical characterization and mode of action of a thermostable endoglucanase purified from *Thermoascus aurantiacus*. *Arch Biochem Biophys* 404(2):243–253.
23. Vlasenko E, Schülein M, Cherry J, Xu F (2010) Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. *Bioresour Technol* 101(7):2405–2411.
24. Fry SC, et al. (1993) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiol Plant* 89(1):1–3.
25. York WS, van Halbeek H, Darvill AG, Albersheim P (1990) Structural analysis of xyloglucan oligosaccharides by 1H-n.m.r. spectroscopy and fast-atom-bombardment mass spectrometry. *Carbohydr Res* 200:9–31.
26. Westereng B, et al. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J Chromatogr A* 1271(1):144–152.
27. Kato K, Matsuda K (1969) Studies on chemical structure of konjac mannan. I. Isolation and characterization of oligosaccharides from partial acid hydrolyzate of mannan. *Agric Biol Chem* 33(10):1446.
28. Mikkelson A, Maaheimo H, Hakala TK (2013) Hydrolysis of konjac glucomannan by *Trichoderma reesei* mannanase and endoglucanases Cel7B and Cel5A for the production of glucomannooligosaccharides. *Carbohydr Res* 372:60–68.
29. Ajithkumar A, Andersson R, Siika-aho M, Tenkanen M, Aman P (2006) Isolation of cellotriosyl blocks from barley beta-glucan with endo-1,4-beta-glucanase from *Trichoderma reesei*. *Carbohydr Polym* 64(2):233–238.
30. Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in Fab-Ms Ms spectra of glycoconjugates. *Glycoconj J* 5(4):397–409.
31. Varki A, et al. (2009) Symbol nomenclature for glycan representation. *Proteomics* 9(24):5398–5399.
32. Gille S, et al. (2011) O-acetylation of *Arabidopsis* hemicellulose xyloglucan requires AXY4 or AXY4L, proteins with a TBL and DUF231 domain. *Plant Cell* 23(11): 4041–4053.
33. Jia ZH, Cash M, Darvill AG, York WS (2005) NMR characterization of endogenously O-acetylated oligosaccharides isolated from tomato (*Lycopersicon esculentum*) xyloglucan. *Carbohydr Res* 340(11):1818–1825.
34. Vaaje-Kolstad G, Horn SJ, van Aalten DMF, Synstad B, Eijsink VGH (2005) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem* 280(31):28492–28497.
35. Karkehabadi S, et al. (2008) The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 A resolution. *J Mol Biol* 383(1): 144–154.
36. Ray A, et al. (2012) *Phanerochaete chrysosporium* produces a diverse array of extracellular enzymes when grown on sorghum. *Appl Microbiol Biotechnol* 93(5): 2075–2089.
37. Hori C, Igarashi K, Katayama A, Samejima M (2011) Effects of xylan and starch on secretome of the basidiomycete *Phanerochaete chrysosporium* grown on cellulose. *FEMS Microbiol Lett* 321(1):14–23.
38. Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *Plant J* 3(1):1–30.
39. Vincken JP, Beldman G, Voragen AGJ (1994) The effect of xyloglucans on the degradation of cell-wall-embedded cellulose by the combined action of cellobiohydrolase and endoglucanases from *Trichoderma viride*. *Plant Physiol* 104(1):99–107.
40. Cannella D, Hsieh CW, Felby C, Jørgensen H (2012) Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol Biofuels* 5(1):26.
41. Harris PV, et al. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry* 49(15):3305–3316.
42. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R (2012) Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in Pichia pastoris monitored by a fluorimetric assay. *Biotechnol Biofuels* 5(1):79.

Supporting Information

Agger et al. 10.1073/pnas.1323629111

SI Materials and Methods

Purification of Endo-Glucanases. Endo-glucanase containing culture broths were concentrated with Vivaflow 50 tangential crossflow concentrator cassettes (molecular weight cutoff 10 kDa, Sartorius Stedim Biotech), and thoroughly washed with ultrapure water to remove residual salts remaining from the culture media. Toward the end of this ultrafiltration procedure, water was exchanged with the equilibration buffer of the subsequent purification step. Endoglucanase *Ta*Cel5A from *Thermoascus aurantiacus* was loaded onto a 5 mL DEAE-Sepharose column (GE Healthcare Bio-Sciences AB) equilibrated with 20 mM Na-acetate buffer, pH 5.0, and eluted by increasing the NaCl concentration from 0 to 0.15 M linearly within 6 column volumes; protein fractions eluting at 0.12–0.15 M NaCl were collected. Endoglucanase *Af*Cel12A from *Aspergillus fumigatus* was purified on a 5-mL SP-Sepharose-FF column (GE Healthcare Bio-Sciences AB) equilibrated with 20 mM Na-citrate buffer, pH 3.0; the protein was eluted with a stepwise NaCl gradient and the fractions eluting between 0.10 and 0.13 M NaCl were collected and pooled. Pooled enzyme fractions were concentrated by ultrafiltration with 20 mL Vivaspin tubes, 10 kDa (Sartorius Stedim Biotech).

Additional Glycan Microarray Array Studies. Lytic polysaccharide monooxygenase 9C from *Neurospora crassa*, *Nc*LPMO9C, was also screened against β-Glucan (yeast), arabinoxylan (wheat), pectic galactan (lupin), glucomannan (konjac), and galactomannan (carob) (Megazyme International Ireland), using the protocol described in the main text with the following modifications, namely 100 mM sodium acetate buffer pH 6.5, *Nc*LPMO9C concentration of 50 μg/mL and overnight incubation at 40 °C.

Antibodies Used in the Arrays. Rat monoclonal antibodies used in the experiments were (the epitopes recognized are in parenthesis): LM15, (Xyloglucan, XXXG motif) (1); LM21, [(1→4)-β-D-mannan/galactomannan/glucomannan] (2); LM10, [(1→4)-β-D-xylan], and LM11, [(1→4)-β-D-xylan/arabinoxylan] (3); LM19, (partially methylesterified and nonesterified homogalacturonan) (4) and LM5, [(1→4)-β-D-galactan] (5). Mouse monoclonal antibodies used were: BS-400–2, [(1→3)-β-D-glucan] (6) and BS-400–3, [(1→3),(1→4)-β-D-glucan] (7).

MALDI-ToF Mass Spectrometry. MALDI-ToF analysis was done on an Ultraflex MALDI-TOF/TOF instrument (Bruker Daltonics) equipped with nitrogen 337 nm laser beam as described (8). In short, the instrument was operated in positive mode and two μL 9 mg/mL mixture of 2,5-dihydroxybenzoic acid (DHB) in 30% acetonitrile was applied to an MTP 384 ground steel target plate TF (Bruker Daltonics). A 2-μL sample was mixed with the DHB droplet followed by drying under a stream of warm air. All spectra were obtained using the reflectron mode with an acceleration voltage of 25 kV, a reflector voltage of 26 and pulsed ion extraction of 40 ns in the positive ion mode. The data were collected from averaging 400 laser shots, with the lowest laser energy necessary to obtain sufficient signal to noise ratios. Peak lists were generated from the MS spectra using Bruker flexAnalysis software (version 3.3).

Direct Infusion ESI-Mass Spectrometry. Mass spectrometry analysis by ESI mass spectrometry was performed on a linear ion trap LTQ Velos Pro (Thermo Scientific) coupled to an Ultimate 3000RS UHPLC from Dionex used to deliver a constant injection flow. The flow was set to 0.2 mL/min (30% 1 mM formic acid/70% acetonitrile vol/vol). The analyses were done by direct infusion to the MS without chromatographic separation and 2-μL sample was injected. The electrospray was operated in positive mode at 4 kV spray current and sheath gas and auxiliary gas flows of 30 and 5 (arbitrary units), respectively. The temperature of the ion transfer tube was set to 250 °C. Acquisition time was set to 0.4 min with normal scan rate, and spectra were averaged from approx. 0.1–0.2 min corresponding to approximately 30 spectra. Full scans were performed in the *m/z* 250–2,000 mass range and MS2 fragmentation was performed with HCD (high collision dissociation), using He as the colliding gas. MS2 fragmentation was performed on Li-adducts. Lithium doping was used to promote improved fragmentation efficiency (9). Li-adduct formation was controlled by addition of 0.5 mM LiCl in 50% ACN via a syringe pump delivering 30 μL/min directly before the ESI probe. Data from ESI-MS was processed in Xcalibur 2.2 SP1.48 (Thermo Scientific).

HPAEC Analysis. High-performance anion exchange chromatography (HPAEC) analysis was performed on a Dionex ICS3000 system equipped with pulsed amperometric detection (PAD) and a CarboPac PA1 column (2 × 250 mm) and CarboPac PA1 guard (2 × 50 mm). For most samples the elution was done at 0.25 mL/min using a 75 min gradient mixing eluent A (0.1 M NaOH) with eluent B (1 M NaOAc in 0.1 M NaOH), as follows: 100% A to 10% B, in 35 min (linear gradient), then to 30% B at 60 mins using Dionex curve 6, then to 100% B at 65 min using Dionex curve 6, and then back to 100% A over 1 min, followed by re-equilibration for an additional 9 min. Reaction rate experiments with Glc_5 were analyzed using a shortened version (29 min) of this gradient: starting with 100% A and reaching 10% B after 10 min, 14% B at 15 min, 30% B at 16 min, 100% B at 18 min, using the same Dionex curves as above, and back to starting conditions in 0.1 min followed by equilibration for 9 min. Quantification for rate comparison experiments were done for cellopentaose by determining the decrease in substrate concentration compared with a standard curve from authenticated cellopentaose (Megazyme International Ireland). For $XG14^{OH}$ the quantification was done based on an external standard curve generated from the substrate and quantified as decrease in substrate concentration (calculated from the major eluting peak, see Fig. S2, green chromatogram).

Determination of Enzyme Rates. For oligomeric substrates reactions were carried out in 100 μL total volume containing 0.2 mM cellopentaose or 0.2 mM $XG14^{OH}$ in 25 mM NH_4OAc, pH 8, with 0.05 mg/mL (1.47 μM) *Nc*LPMO9C, with or without 1 mM ascorbic acid. Reactions were incubated at 40 °C and 600 rpm. Reactions were quenched by adding methanol to a final concentration of 66% (vol/vol). Reaction products were analyzed using HPAEC and reaction rates determined based on substrate conversion.

Assays for polymeric substrates [tamarind xyloglycan or phosphoric acid swollen cellulose (PASC)] were run at 5 mg/mL in 5 mM sodium phosphate buffer (pH 8) at 50 °C and 1000rpm with 0.05 mg/mL (1.47 μM) *Nc*LPMO9C and 1 mM ascorbic acid in triplicates. Product formation was monitored by determining the increase in reducing ends with the dinitrosalisylic acid (DNS) assay (10). Reactions were quenched by adding DNS-reagent at appropriate time points. A 100-μL enzyme reaction volume was added to 150 μL DNS solution (0.04 M 3,5-dinitro salicylic acid in 1.0 M K-Na-tartrate, 0.4 M NaOH). The samples were mixed properly and incubated for 5 min in boiling water. Hereafter the samples were centrifuged and 200 μL of the reaction mixture was transferred to

96-well plates for measuring the absorbance at 540 nm. Calibration was done with cellobiose, for PASC samples or XG7, for xyloglucan samples. Control reactions containing substrate and ascorbic acid, substrate and enzyme and substrate alone were included and results were corrected for background absorbance.

1. Marcus SE, et al. (2008) Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biol* 8:60.
2. Marcus SE, et al. (2010) Restricted access of proteins to mannan polysaccharides in intact plant cell walls. *Plant J* 64(2):191–203.
3. McCartney L, Marcus SE, Knox JP (2005) Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *J Histochem Cytochem* 53(4):543–546.
4. Verhertbruggen Y, Marcus SE, Haeger A, Ordaz-Ortiz JJ, Knox JP (2009) An extended set of monoclonal antibodies to pectic homogalacturonan. *Carbohydr Res* 344(14): 1858–1862.
5. Jones L, Seymour GB, Knox JP (1997) Localization of pectic galactan in tomato cell walls using a monoclonal antibody specific to (1[->]4)-[beta]-D-galactan. *Plant Physiol* 113(4):1405–1412.
6. Meikle PJ, Bonig I, Hoogenraad NJ, Clarke AE, Stone BA (1991) The location of (1→3)-β-glucans in the walls of pollen tubes of *Nicotiana alata* using a (1→3)-β-glucan-specific monoclonal antibody. *Planta* 185(1):1–8.
7. Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA (1994) A (1—>3,1—>4)-beta-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1—>3,1—>4)-beta-glucans. *Plant J* 5(1):1–9.
8. Vaaje-Kolstad G, et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330(6001):219–222.
9. Cancilla MT, Wong AW, Voss LR, Lebrilla CB (1999) Fragmentation reactions in the mass spectrometry analysis of neutral oligosaccharides. *Anal Chem* 71(15):3206–3218.
10. Sumner JB (1924) The estimation of sugar in diabetic urine, using dinitrosalicylic acid. *J Biol Chem* 62(2):287–290.

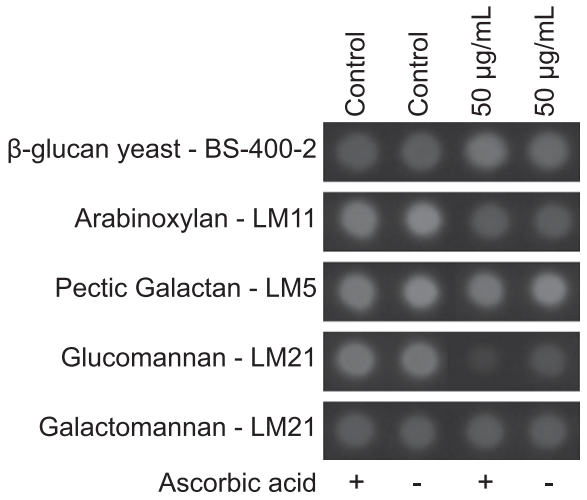

Fig. S1. Glycan microarrays for testing *Nc*LPMO9C activity on additional substrates. The substrates were treated with *Nc*LPMO9C at 50 μg/mL in the presence (+) or absence (–) of ascorbic acid. In the control reactions, no enzyme was added. The antibodies used for detection are provided next to the substrate names. Glucomannan was included as a positive control.

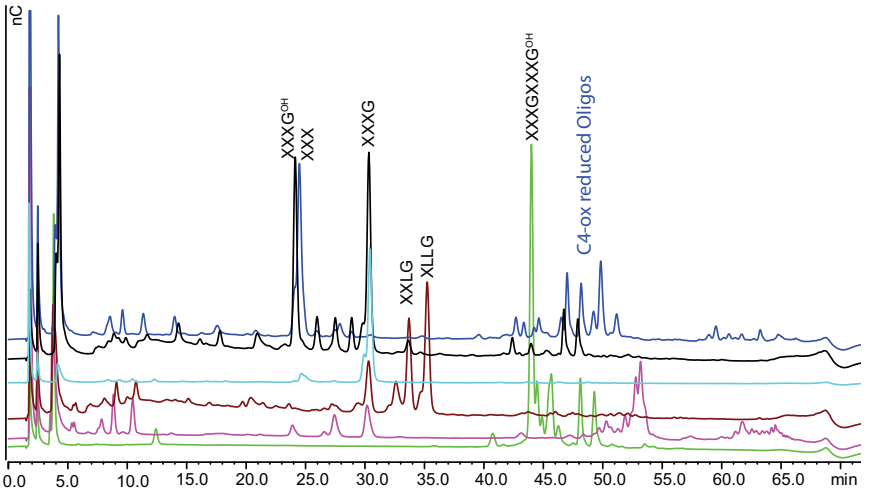

Fig. S2. ICS chromatograms showing products generated from xyloglucan and XG14^{OH} (XXXGXXXG^{OH}). The peak annotations are putative in part and based on comparison with product profiles obtained by MALDI-ToF MS and previously described order of elution of xyloglucan oligosaccharides (1). Magenta: products generated upon incubation of polymeric xyloglucan with *Nc*LPMO9C; the late eluting peak cluster (50–54 min) most probably contains C4-oxidized species. Green: nontreated XG14^{OH}; the major peak at 44 min is XG14^{OH} and the following peaks correspond to the impurities (see Fig. 3*B*). Blue: products generated upon incubation of XG14^{OH} with *Nc*LPMO9C; the peak at 24.5 min most likely corresponds to XXX and C4-oxidized species elute late (47–50 min). The modification of the reducing end in these products is probably causing the early elution compared with xyloglucan derived oxidized products (magenta). Cyan: XXXG (XG7) without enzyme treatment, elutes at 30 min. Black: products generated upon incubation of XG14^{OH} with *Af*Cel12A; the peak at 24 min most likely corresponds to the reduced version of XXXG because the peak at 30.5 min corresponds to XXXG as verified by elution of XG7 (in cyan). Brown: products generated upon incubation of polymeric xyloglucan with *Af*Cel12A; the peak at 33.7 min corresponds to XXLG, and the peak at 35.2 min corresponds to XLLG. These latter peaks are annotated according to the order of elution described in ref. 1. Minor peaks in the proximity of the annotated peaks may be different configurations of products (e.g., variation in the position of L).

1. Hilz H, Bakx EJ, Schols HA, Voragen AGJ (2005) Cell wall polysaccharides in black currants and bilberries–characterisation in berries, juice, and press cake. *Carbohydr Polym* 59(4):477–488.

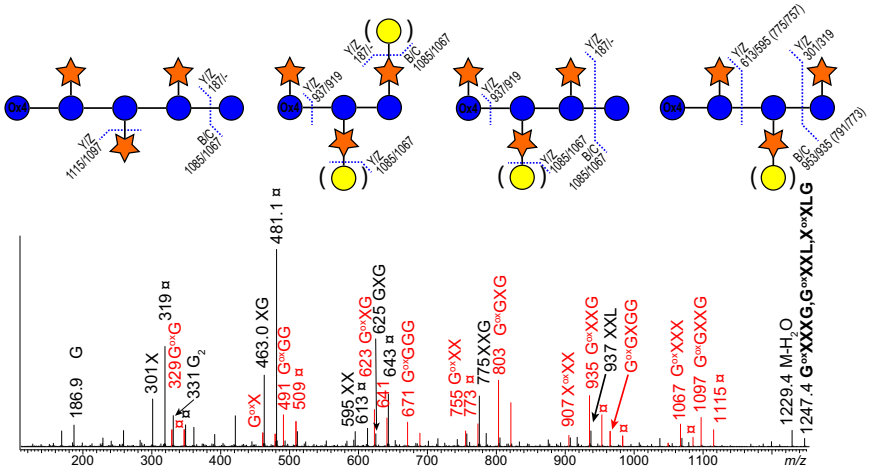

Fig. S3. MS2 fragmentation of *m/z* 1,247 (Li-adduct), produced upon incubation of polymeric xyloglucan (from tamarind) with *Nc*LPMO9C. The occurrence of an *m/z* 937 fragment implies the loss of an oxidized X-unit, whereas *m/z* 1,067 implies the loss of a terminal native hexose. The latter ion can only arise if an L-unit, which has a terminal galactose, or a G-unit in the reducing end is present within the *m/z* 1,247 product pool from XG, which again would mean that the oxidation could be carried on an X-unit, as indeed indicated by the presence of the *m/z* 937 species (so the *m/z* 1,247 ion would then be X^{ox}XLG, where the position of the L may vary). Possible product configurations giving *m/z* 1,247 are shown as cartoons including *m/z* values for key fragmentation ions. The symbols are as suggested in ref. 1: blue circle, glucose; orange star, xylose; yellow circle, galactose. Parenthesis surrounding galactosyl-units denote that the position of these units may vary. Black annotated masses are fragments originating from the reducing end (dominated by Y/Z-ions, Y ions marked ¤). Red annotated masses are fragments originating from the nonreducing end hence carrying the oxidation, dominated by B and C-ions. Gemdiol-fragments readily lose a water molecule thus occurring in the spectrum as B or C −18 Da (C ions −18 Da marked ¤). Such loss of water has been observed previously for gemdiol products (2).

1. Varki A, et al. (2009) Symbol nomenclature for glycan representation. *Proteomics* 9(24):5398–5399.
2. Isaksen T, et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem* 289(5):2632–2642.

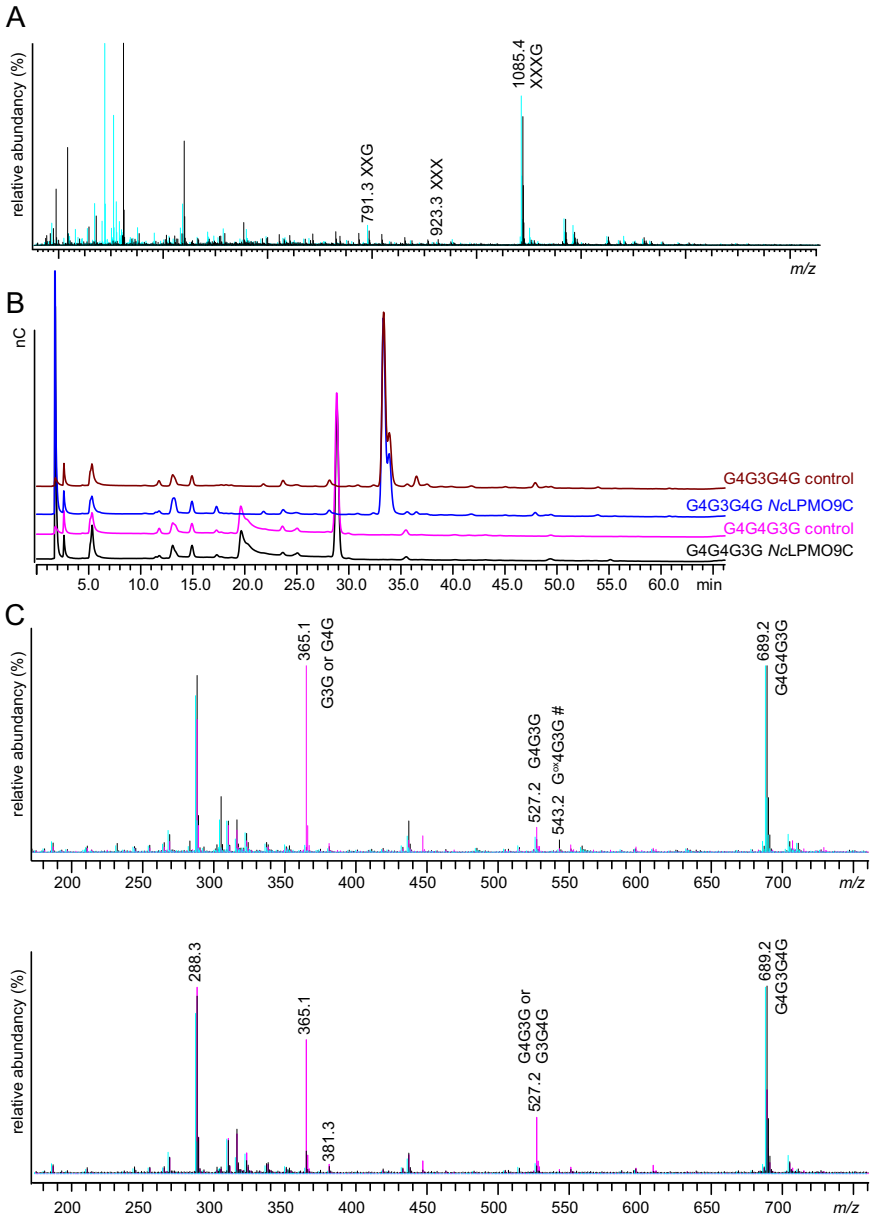

Fig. S4. (*A*) ESI-MS spectra showing products generated by *Nc*LPMO9C upon incubation with XG7 (XXXG) and control without enzyme. Note that the spectra seem to suggest that conversion of XXXG to XXX occurs at very low rate. However, minor amounts of XXX were also present in the controls. For clarity, the control spectrum (turquoise) is presented with an offset of 1 amu to the left compared with the *Nc*LPMO9C treated sample (black). (*B*) HPAEC chromatograms showing products generated upon incubation of *Nc*LPMO9C with the mixed-linkage cellotetraosyl G4G3G4G (blue) and G4G4G3G (black) with corresponding controls without enzyme (brown and magenta, respectively). The chromatograms show that there is no activity on G4G3G4G and limited activity on G4G4G3G (minor peaks eluting at 50 and 55 min may represent C4-oxidized products). (*C*) ESI-MS spectra showing activity of *Nc*LPMO9C or *Ta*Cel5A toward G4G4G3G (*Upper*) and G4G3G4G (*Lower*) with the corresponding controls. Controls (turquoise) are presented with an offset of 1 amu to the left compared with the *Nc*LPMO9C treated samples (black) and the *Ta*Cel5A treated samples (magenta). The data show that *Ta*Cel5A is able to hydrolyze both tetramers, producing nonoxidized glucose dimers (*m/z* 365). *Nc*LPMO9C is only active on G4G4G3G, which is converted very slowly to an oxidized hexose trimer (*m/z* 543, gemdiol #). Most likely, the substrate is converted to G + G^{ox}4G3G. What appears to be oxidized dimer (*m/z* 381) in the lower spectrum is most likely not a valid signal because it is also present in the controls at similar intensity.

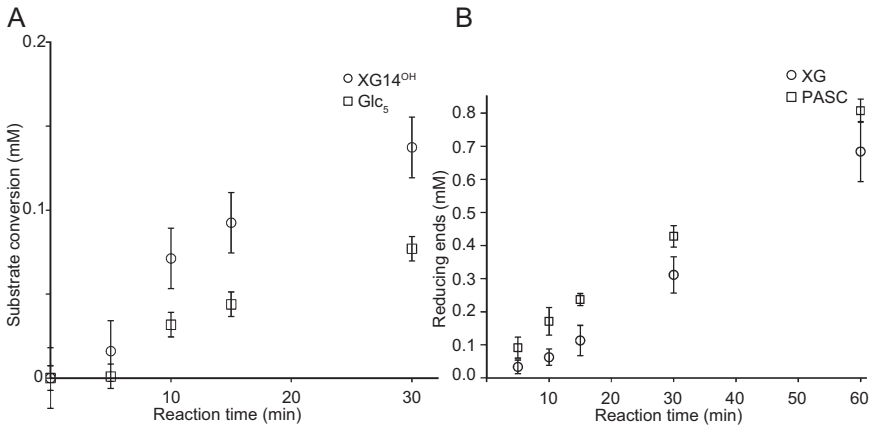

Fig. S5. Comparison of degradation rates. (*A*) Degradation of 0.2 mM $XG14^{OH}$ and 0.2 mM cellopentaose (Glc_5) by 1.47 μM *Nc*LPMO9C, monitored by recording substrate disappearance. Data points reflect conversion of substrate (mM) with SDs. These curves yield approximate rates of 0.06 s^{-1} and 0.03 s^{-1}, respectively at 40 °C. (*B*) Degradation of 5 mg/mL tamarind xyloglucan and 5 mg/mL PASC by 1.47 μM *Nc*LPMO9C at 50 °C, monitored by recording formation of reducing ends. Curves for reactions performed at 50 °C yield approximate rates of 0.11 s^{-1}. Note that control reactions with enzyme but without reductant did not yield products. The only other rate ever described for an LPMO, namely for chitin degradation by CBP21, an LPMO10 from *Serratia marcescens*, was in the order of 0.02 s^{-1} (1).

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

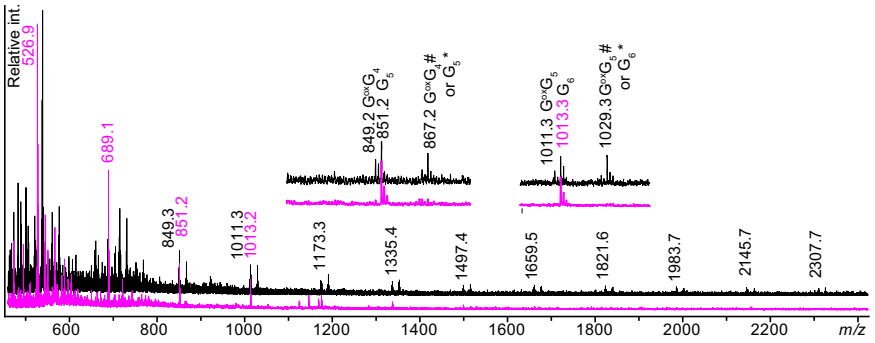

Fig. S6. MALDI-ToF spectra of products generated from natural Icelandic moss incubated with either *Nc*LPMO9C (black line) or endo-glucanase *Ta*Cel5A (magenta line). The product profiles showed clear differences between *Nc*LPMO9C and the endo-glucanase, similar to what was seen for purified lichenan (Fig. 2). Whereas the endo-glucanase generates short, nonoxidized products, longer, oxidized products accumulate in the reaction with the LPMO. # denotes the gemdiol-form of a keto-group in the C4-position. All assigned masses are sodium adducts, except when indicated otherwise. Potassium adducts are marked by an asterisk. Control reactions showed that ascorbic acid alone did not release any products and did not change the product profile generated by the endo-glucanases.

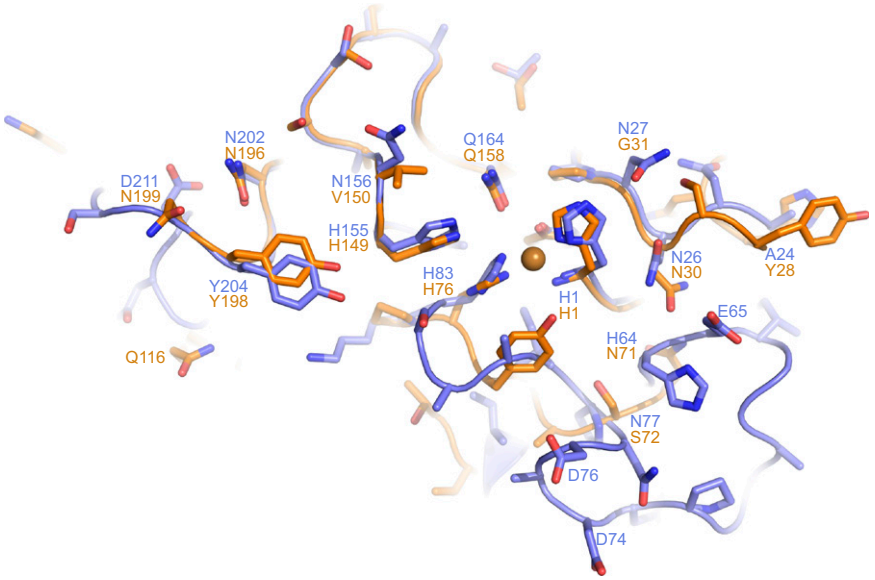

Fig. S7. Comparison of *Nc*LPMO9C and *Pc*LPMO9D highlighting protruding surface residues. The figure shows a 3D model of *Nc*LPMO9C (light blue carbon atoms) superposed on the crystal structure of *Pc*LPMO9D (orange carbon atoms; ref. 1). The model of *Nc*LPMO9C was made based on a structure prediction by HHpred (2) and built using Modeler (3) with the crystal structure of *Nc*LPMO9D (4) as a template (47.5% sequence identity). The orange sphere represents the copper ion in the structure of *Pc*LPMO9D. The side chains of the most protruding surface residues are labeled (coloring as for the carbon atoms). The superposition suggests that the substrate-binding surface of *Nc*LPMO9C is more extended and has a more polar character compared with *Pc*LPMO9C. Notably, the activity of *Pc*LPMO9C on $XG14^{OH}$ was tested using the same conditions as for *Nc*LPMO9C (Fig. 3), and these experiments showed that *Pc*LPMO9C is not active on this xyloglucan fragment. The figures were made with PyMOL (PyMOL Molecular Graphics System, version 1.5.0.4; Schrödinger).

1. Wu M, et al. (2013) Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem* 288(18):12828–12839.
2. Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33(Web Server issue):W244-8.
3. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779–815.
4. Li X, Beeson WT, 4th, Phillips CM, Marletta MA, Cate JH (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* 20(6): 1051–1061.

Paper III

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 38, pp. 22955–22969, September 18, 2015

Structural and Functional Characterization of a Lytic Polysaccharide Monooxygenase with Broad Substrate Specificity*

Received for publication, April 21, 2015, and in revised form, July 2, 2015 Published, JBC Papers in Press, July 15, 2015, DOI 10.1074/jbc.M115.660183

Anna S. Borisova[‡1], Trine Isaksen[§1], Maria Dimarogona[‡], Abhishek A. Kognole[¶], Geir Mathiesen[§], Anikó Várnai[§], Åsmund K. Røhr[‖], Christina M. Payne[‡¶2], Morten Sørlie[§], Mats Sandgren[‡3], and Vincent G. H. Eijsink[§4]

From the [‡]Department of Chemistry and Biotechnology, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden, the [§]Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway, the [‖]Department of Biosciences, University of Oslo, N-0316 Oslo, Norway, and the [¶]Department of Chemical and Materials Engineering, University of Kentucky, Lexington, Kentucky 40506

Background: The recently discovered lytic polysaccharide monooxygenases (LPMOs) are important in enzymatic conversion of lignocellulosic biomass.
Results: We describe structural and functional studies of *Nc*LPMO9C, which cleaves both cellulose and certain hemicelluloses.
Conclusion: *Nc*LPMO9C has structural and functional features that correlate with the enzyme's catalytic capabilities.
Significance: This study shows how LPMO active sites are tailored to varying functionalities and adds to a growing LPMO knowledge base.

The recently discovered lytic polysaccharide monooxygenases (LPMOs) carry out oxidative cleavage of polysaccharides and are of major importance for efficient processing of biomass. *Nc*LPMO9C from *Neurospora crassa* acts both on cellulose and on non-cellulose β-glucans, including cellodextrins and xyloglucan. The crystal structure of the catalytic domain of *Nc*LPMO9C revealed an extended, highly polar substrate-binding surface well suited to interact with a variety of sugar substrates. The ability of *Nc*LPMO9C to act on soluble substrates was exploited to study enzyme-substrate interactions. EPR studies demonstrated that the Cu^{2+} center environment is altered upon substrate binding, whereas isothermal titration calorimetry studies revealed binding affinities in the low micromolar range for polymeric substrates that are due in part to the presence of a carbohydrate-binding module (CBM1). Importantly, the novel structure of *Nc*LPMO9C enabled a comparative study, revealing that the oxidative regioselectivity of LPMO9s (C1, C4, or both) correlates with distinct structural features of the copper coordination sphere. In strictly C1-oxidizing LPMO9s, access to the solvent-facing axial coordination position is restricted by a conserved tyrosine residue, whereas access to this same position seems unrestricted in C4-oxidizing LPMO9s. LPMO9s known to produce a mixture of C1- and C4-oxidized products show an intermediate situation.

Naturally abundant polysaccharides such as cellulose and chitin are enticing supplements to fossil resources in the future production of fuels, chemicals, and other materials, but the economical integration of these alternative carbon sources into fuel production pathways is hampered by their recalcitrance to hydrolytic deconstruction. Until recently, the degradation of polysaccharides in plant biomass, such as cellulose and various hemicelluloses, was thought to be entirely driven by hydrolytic enzymes. However, it has been shown recently that oxidative degradation catalyzed by lytic polysaccharide monooxygenases (LPMOs)[5] (1) plays a significant role (2–6). LPMOs are copper-dependent biocatalysts that use molecular oxygen and an electron donor to break glycosidic bonds (5–9). Electrons may be provided by a reducing agent such as ascorbic acid, gallate, or reduced glutathione (5, 6) or by a co-secreted enzyme such as cellobiose dehydrogenase (10). The catalytic copper-binding site of LPMOs has unique features that are shared with only one other type of enzyme, methane monooxygenase (5, 11, 12).

Because of their potential use in commercial biomass decomposing enzyme mixtures (3), LPMOs have attracted considerable attention since their discovery in 2010 (6). Initially, LPMO activity was demonstrated for cellulose and chitin. However, it was noted that the abundance, the large sequence variation, and the varying domain composition of LPMOs encoded in the genomes of biomass-degrading microorganisms suggest that some of these enzymes could have other substrates (1). Indeed,

* This work was supported by the Faculty for Natural Resources and Agriculture, Swedish University of Agricultural Sciences through the research program MicroDrivE (Microbially Derived Energy), and by the Norwegian Research Council Grants 209335, 214613, and 216162. The authors declare that they have no conflicts of interest with the contents of this article.
The atomic coordinates and structure factors (codes 4D7U and 4D7V) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Both authors contributed equally to this work.
[2] Supported by the August T. Larsson Guest Researcher Programme at the Swedish University of Agricultural Sciences.
[3] To whom correspondence may be addressed: Dept. of Chemistry and Biotechnology, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden. Tel.: 46-18673179; E-mail: mats.sandgren@slu.se.
[4] To whom correspondence may be addressed: Dept. of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway. Tel.: 47-67232463; E-mail: vincent.eijsink@nmbu.no.

[5] The abbreviations used are: LPMO, lytic polysaccharide monooxygenase; ITC, isothermal titration calorimetry; PDB, Protein Data Bank; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; TMP, *N,N,N′,N′*-tetramethyl-1,4-phenylenediamine; AA, auxiliary activity; PASC, phosphoric acid-swollen cellulose; YPD, yeast extract/peptone/dextrose.

three studies published in 2014 have broadened the LPMO paradigm by demonstrating cleavage of soluble cellodextrins (13), hemicelluloses with β(1→4)-glucan backbones (14), and even starch (15). Based on sequence characteristics, LPMOs are currently categorized in auxiliary activity (AA) families 9–11 and 13 of the CAZy database (16). AA families 9, 11, and 13 almost exclusively contain fungal enzymes, whereas bacterial LPMOs occur in family 10. So far, chitin activity has been described for families 10 and 11, cellulose activity for families 9 and 10, hemicellulose activity for family 9, and starch activity for family 13. It has been shown that the action of family 9 and 10 LPMOs on β(1→4)-glycosidic linkages may lead to oxidation of either the C1 or the C4 carbon of the glycan (8, 17). Based on sequence analysis and the known variation in oxidative outcome of the reaction, it has been suggested that AA9 LPMOs may be divided into C1-oxidizing (LPMO1 group), C4-oxidizing (LPMO2 group), and both C1- and C4-oxidizing (LPMO3 group) enzymes (18).

We have recently discovered that *Nc*LPMO9C, a C4-oxidizing AA9 LPMO (hereafter LPMO9) from *Neurospora crassa* also known as NCU02916 or *Nc*GH61–3, shows activity on soluble cellodextrins and hemicelluloses with β(1→4)-glucan backbones, including xyloglucan, glucomannan, and β-glucan (13, 14). *Nc*LPMO9C is a two-domain protein, containing an N-terminal catalytic AA9 domain and a family 1 carbohydrate-binding module (CBM1), that is connected through a serine- and threonine-rich linker comprising ~50 amino acid residues. CBM1s are primarily found in multidomain proteins from fungi and are known to bind both crystalline and amorphous cellulose (19, 20). Approximately 20% of all AA9s contain at least one CBM1 domain. Until now, the effect of such binding domains on LPMO activity and substrate specificity has not been thoroughly investigated.

In this study, we set out to explore the structural basis of the unique functional properties of *Nc*LPMO9C. We report the cloning, purification, and structural determination of *Nc*LPMO9C-N, which is a C-terminally truncated variant of *Nc*LPMO9C comprising only the AA9 domain. Further insights into *Nc*LPMO9C properties were obtained from EPR studies in the presence or absence of soluble substrate, studies of metal and substrate binding affinities using ITC, determination of the redox potential, and activity assays using several substrates and *Nc*LPMO9C variants with and without the CBM1. Notably, the unique functional properties of *Nc*LPMO9C allowed for enzyme-substrate interaction studies that have not been possible to date for LPMOs, and the novel structure of the enzyme enabled a structural comparison revealing structural determinants of the oxidative regioselectivity of these enzymes.

Experimental Procedures

Cloning, Expression, and Purification—Full-length LPMO9C from *N. crassa* (NCU02916; *Nc*LPMO9C) was produced in *Pichia pastoris* using a codon-optimized gene inserted into the pPICZα-A plasmid (21) and purified as described previously (22). A gene fragment encoding the AA9 domain of *Nc*LPMO9C (nucleotides 1–729, encoding 243 residues) was PCR-amplified from the pPICZαA_*Nc*LPMO9C vector. The amplification primers containing EcoRI and Acc65I restriction sites for cloning (underlined) were F1, TTTCGAAACGGAATTCGAAACGATGAAGACTGGTTCCATCTTGGC, and R2, ATGGCCGGCCGGTACCTCAACAAGTGAACAAAGCTGGACCT. The PCR product was ligated into the vector fragment emerging after cleaving the pPINK_GAP_AfCel12A vector (23) with restriction enzymes EcoRI and Acc65I, using the In-Fusion HD cloning kit (Clontech). This construct leads to constitutive expression of the protein driven by the GAP promoter, and secretion of the protein is driven by its native signal peptide. After the DNA sequence was confirmed by GATC Biotech AG (Germany), the resulting pGAP_*Nc*LPMO9C-N plasmid was linearized with the restriction enzyme AflII and was used to transform competent PichiaPink™ strain 4 (Invitrogen), according to the supplier's instructions. Positive transformants were selected on *Pichia* adenine dropout plates (Teknova, Hollister, CA) and were further cultivated in yeast extract/peptone/dextrose (YPD) media for detection of LPMO expression and activity.

For the production of *Nc*LPMO9C-N, 50 ml of YPD medium, containing 2% (w/v) dextrose, was inoculated with a single colony of a positive *P. pastoris* transformant. After growth at 30 °C and 150 rpm for 16 h, these 50 ml were used as preculture to inoculate 1.5 liters of YPD medium to a starting optical density (OD) of 0.15. Cultivation was carried out for 5 days at 30 °C and 150 rpm in shake flasks with standard baffles until the final OD of 19. Starting YPD media contained 1% (w/v) yeast extract, 2% (w/v) peptone, and 1% (w/v) glucose in 0.1 M potassium phosphate buffer, pH 6.0. Every 24 h, 1% (w/v) of glucose was added to the culture medium, and the *Nc*LPMO9C-N expression level was monitored by SDS-PAGE.

As the first purification step, culture supernatant was recovered by centrifugation at 18,000 rpm for 15 min followed by sequential filtering through 0.45- and 0.2-μm polyethersulfone (PES) filters (Millipore, Billerica, MA). A Vivaflow 200 tangential cross-flow concentrator (molecular mass cutoff of 10,000 Da, Sartorius Stedim Biotech GmbH, Germany) was then used to concentrate the supernatant 10-fold and to exchange buffer to 25 mM Tris-HCl, pH 8.0. The concentrated sample was loaded onto a 150-ml DEAE-Sepharose FF CL-6B column (GE Healthcare) equilibrated in 25 mM Tris, pH 8.0, and bound proteins were eluted by applying a linear gradient up to 0.25 M NaCl in the same buffer. Selected fractions were pooled, concentrated, and applied to a Superdex 75 gel filtration column (GE Healthcare) in 50 mM MES buffer, pH 6.5, supplemented with 150 mM NaCl. Protein purity was assessed by SDS-PAGE, and protein concentrations were determined by measuring absorbance at 280 nm, using theoretical extinction coefficients calculated using the ExPASy server (24) (*Nc*LPMO9C, 46,910 M^{-1} cm^{-1}; *Nc*LPMO9C-N, 32,680 M^{-1} cm^{-1}).

Preparation of Crystals and Structural Determination—The search for *Nc*LPMO9C-N crystallization conditions was performed in 96-well sitting drop trays, using a Mosquito crystallization robot (TTP Labtech, UK) and commercially available screens. Diffracting crystals grew at room temperature in a variety of conditions containing PEG 3350 and at a protein concentration of 1.4 mg/ml (see "Results" for more details). Various efforts were made to obtain crystals of *Nc*LPMO9C-N in complex with substrate analogs, such as thio-linked cello- and xylo-

TABLE 1

***Nc*LPMO9C-N ligand docking cases considered in this study**

The "receptor" column describes the metal ions included in the active site. The "torsion angles" column describes whether torsion angles along the glycosidic linkages were held fixed or not. Search parameters for each case are given in the remaining columns. More details are provided under "Experimental Procedures."

Receptor	Ligand	Torsion angles	Grid box	Runs	Population	Minimum survivors
No metals	Glc_4	Rigid	60*120*120	200	200	3
1 Cu^+, 2 zinc	Glc_4	Rigid	60*120*120	100	300	3
2 Cu^+, 1 zinc	Glc_4	Rigid	60*120*120	200	300	3
1 Cu^+, 2 zinc	Glc_4	Flexible	100*100*100	300	300	1
2 Cu^+, 1 zinc	Glc_4	Flexible	100*100*100	300	300	1
2 Cu^+, 1 zinc	Glc_5	Rigid	60*120*120	200	200	3
2 Cu^+, 1 zinc	Glc_6	Rigid	60*120*120	200	200	3
No metals	*XXGX*	Flexible	60*120*120	200	200	3
1 Cu^+, 2 zinc	*XXGX*	Flexible	60*120*120	200	200	3
2 Cu^+, 1 zinc	*XXGX*	Flexible	60*120*120	200	200	3

oligosaccharides. In attempts to obtain ligand complexes, *Nc*LPMO9C-N was deactivated using EDTA or potassium cyanide prior to co-crystallization or soaking experiments with cellulosic or xyloglucan oligomers. *Nc*LPMO9C-N crystals were soaked in mother liquor supplemented with 20% glycerol as cryoprotectant and then flash-frozen in liquid N_2 prior to transportation and x-ray data collection on beamline I911-3 at MAX-lab (Lund, Sweden). The data were integrated with XDS (25) and scaled using the program Aimless in the CCP4 suite (26). The structure was solved by molecular replacement using PHASER (27) with the coordinates of another *N. crassa* LPMO, *Nc*LPMO9D, also known as *Nc*GH61–4 or NCU01050 (PDB code 4EIR (18)), as a search model. *Nc*LPMO9D is a single domain C4-oxidizing LPMO sharing 47% sequence identity with *Nc*LPMO9C-N.

REFMAC5 (28) was used for structure model refinements, and manual model rebuilding was performed with Coot (29, 30), with maximum likelihood σA-weighted $2F_{obs} - F_{calc}$ electron density maps (29). For cross-validation and R and R_{free} calculations, 5% of the data were excluded from the structure refinement (31). Solvent molecules were automatically added using the automatic water picking function in the ARP/wARP package (32). Picked water molecules were selected or discarded manually by visual inspection of the $2F_{obs} - F_{calc}$ electron density maps. The copper or the zinc ions bound in the active site of the enzyme in the two different *Nc*LPMO9C-N structures were introduced at a final stage of the structure refinement. The coordinates for the two final structure *Nc*LPMO9C models, and the structure factors, have been deposited in the Protein Data Bank (PDB) with accession codes 4D7U and 4D7V. PyMOL (version 1.5) and UCSF Chimera (version 1.10.1) were used for analysis of the structures and figure preparations.

Substrate Docking Experiments—The docking study evaluated 10 different scenarios (Table 1) in an attempt to identify putative soluble substrate-binding sites for *Nc*LPMO9C-N. Autodock version 4.2 was used to compute and cluster the grid-based free energies of the enzyme·ligand complexes (33). Four different ligands, corresponding to those evaluated in the biochemical assessment, were included in the docking studies as follows: xyloglucan (XXGX), cellotetraose (Glc_4), cellopentaose (Glc_5), and cellohexaose (Glc_6). The XXGX xyloglucan pattern was selected to encompass the glycosidic linkage thought to be the cleavage site in tamarind xyloglucan. The ligand conformational state was either fixed with the pyranose ring "backbone" in a near-linear conformation or allowed to be flexible (Table 1). Oligomers allowed to adopt different conformations throughout the docking calculation are termed "flexible" in this study. This simply means the torsion angles of the molecules were free of constraints. A rigid version of cellulosic oligomers was considered because the backbone of cello-oligomers does not change significantly in solution. Here, "rigid" refers to a molecule in which the torsion angles associated with the glycosidic linkages were restrained.

The participation of *Nc*LPMO9C-N-bound metal ions in substrate binding was considered through variation of both type and number. Three different protein scenarios were considered as follows: 1) *Nc*LPMO9C without any metal ions; 2) *Nc*LPMO9C with one Cu^+ in the histidine brace at the active site and two zinc ions in the structurally resolved metal-binding locations; and 3) *Nc*LPMO9C with Cu^+ in the histidine brace, one Cu^+ in the metal-binding site closest to the active center, and a zinc ion in the distal loop region. The latter scenario was considered based on an early hypothesis that two copper ions were participating in catalysis. Ultimately, as described under "Results," the EPR studies clearly indicate only a single copper atom is required for *Nc*LPMO9C to catalyze cleavage of the β-1,4-linkage.

Autogrid4 was used to set up the grid maps centered on the *Nc*LPMO9C-N active site metal atoms and the surrounding protein surface. Grid dimensions for each case are given in Table 1; a grid spacing of 0.375 Å was used for each case. Using Lamarckian Genetic Algorithm, the search for the most favorably docked ligand structures was carried out over 200 independent docking runs. The search parameters for the calculations were set as follows: mutation rate of 0.02, crossover rate of 0.80, and a maximum number of energy evaluations per run of 25 million. Initial population and maximum number of top individual survivors are given in Table 1.

Copper is a nonstandard atom type in Autodock 4.2. Thus, parameters for non-hydrogen-bonding copper were incorporated in the parameter file (34). The sum of the van der Waals radii of two copper atoms was set to 3.40 Å. The van der Waals well depth was 0.050 kcal/mol. The atomic solvation volume was set to 20.58 $Å^3$, and the atomic solvation parameter was −0.00110.

Electron Spin Resonance Studies—To prepare apo *Nc*LPMO9C for electron spin resonance (EPR) experiments, purified enzyme was treated with 100 mM EDTA, followed by desalting using a NAP-5 column (GE Healthcare) equilibrated

with Chelex-treated 20 mM BisTris/HCl buffer, pH 6.0. To reconstitute the copper-active site, 10 molar eq of Cu^{2+} were added from an acidic $CuCl_2$ solution. After 10 min of incubation at room temperature, excess Cu^{2+} was removed by repeating the desalting procedure described above. EPR samples containing 225 μM apo *Nc*LPMO9C, 160 μM Cu^{2+}-loaded *Nc*LPMO9C, 160 μM Cu^{2+}-loaded *Nc*LPMO9C and 20 mg/ml cellohexaose (Megazyme International, Ireland), or 160 μM Cu^{2+}-loaded *Nc*LPMO9C and 15 mg/ml xyloglucan isolated from tamarind seeds (Megazyme International) were frozen and stored in liquid nitrogen.

EPR spectra were recorded using a BRUKER EleXsys 560 SuperX instrument equipped with an ER 4122 SHQE SuperX high sensitivity cavity and a cold finger. The spectra were recorded using 0.05–1 milliwatt microwave power and 2 or 10 G modulation amplitude at a temperature of 77 K. Spin quantification was carried out by comparing double integrated spectra of samples to a standard of 500 μM $CuCl_2$ in 1 M $HClO_4$. The EasySpin toolbox developed for MATLAB (MathWorks) was used to simulate and fit EPR spectra (35). Only the g_z and $|A_z|$ values were determined accurately, whereas g_x, g_y, $|A_x|$, and $|A_y|$ were estimated using the EasySpin automatic fitting procedure.

Isothermal Titration Calorimetry Studies—ITC experiments were performed with a VP-ITC system from Microcal, Inc. (Northampton, MA) (36). Solutions were thoroughly degassed prior to experiments to avoid air bubbles in the calorimeter. Chelex-treated 20 mM MES buffer, pH 5.5, was used for all binding studies. For binding of Cu^{2+} to apo *Nc*LPMO9C-N (prepared as described above for *Nc*LPMO9C, eluted with 20 mM Chelex-treated MES buffer, pH 5.5), typically 5 μM enzyme was placed in the reaction cell with a volume of 1.42 ml, and 150 μM $CuSO_4$ was placed in the ITC syringe. Aliquots of 4 μl were injected into the reaction cell at 180-s intervals with a stirring speed of 260 rpm and a temperature of 10 °C. The titrations were complete after 40 injections. For binding of Glc_6 to *Nc*LPMO9C, 30 μM copper-saturated enzyme (prepared as described above) was placed in the reaction cell and 11 mM ligand in the syringe. Aliquots of 8 μl were injected at 180-s intervals with a stirring speed of 260 rpm and temperature of 25 °C. The titrations were complete after 40 injections. For binding of *Nc*LPMO9C or *Nc*LPMO9C-N to the xyloglucan, 0.9 μM xyloglucan (225 kDa, from tamarind seeds, Megazyme) was placed in the reaction cell, and 500 μM enzyme in the ITC syringe. Aliquots of 6 μl were injected into the reaction cell at 180-s intervals with a stirring speed of 260 rpm and temperature of 10 °C. The titrations were complete after 50 injections. For binding of *Nc*LPMO9C or *Nc*LPMO9C-N to phosphoric acid-swollen cellulose (PASC; prepared from Avicel as described by T. M. Wood (37)), 0.146 mg/ml substrate was placed in the reaction cell. The concentration of PASC was set to be 4.5 μM based on an estimation that the average chain length of Avicel-derived PASC is 200 glucose units (38). 500 μM enzyme was placed in the ITC syringe. Aliquots of 6 μl were injected into the reaction cell at 180-s intervals with a stirring speed of 260 rpm and temperature of 25 °C. The titrations were complete after 50 injections. ITC data were collected automatically using the Microcal Origin version 7.0 software accompanying the VP-ITC system (36). Prior to further analysis, all data were corrected for heat of dilution by subtracting the heat produced by continuing injections of ligand into the reaction cell after completion of the binding reaction. These heats had the same magnitudes as the heats of titrating ligand into buffer alone. The data were fitted using a nonlinear least squares algorithm using a single-site binding model employed by the Origin software that accompanies the VP-ITC system, yielding the stoichiometry (n), the equilibrium binding association constant (K_a), and the enthalpy change (ΔH_r^0) of the reaction. Errors in ΔH_r^0, K_d, and ΔG_r^0 were obtained as standard deviations of at least three experiments. Errors in ΔS_r^0 and $-T\Delta S_r^0$ were obtained through propagation of error.

Determination of Redox Potential—The cell potential (E_0) for the *Nc*LPMO9C-Cu^{2+}/*Nc*LPMO9C-Cu^{+} redox couple was determined as described previously (7, 39). In short, 50 μl of 70 μM Cu^{2+}-saturated *Nc*LPMO9C was incubated at room temperature with 50 μl of an oxygen-free solution of 200 μM reduced *N,N,N′,N′*-tetramethyl-1,4-phenylenediamine (TMP_{red}) in Chelex-treated 20 mM MES, pH 5.5. Solutions were made oxygen-free by bubbling of N_2 (g) through the buffer for 1 h prior to addition to TMP_{red} and concentrated LPMO solutions. The reactions took place in UVettes (Eppendorf), and the absorbance at 610 nm was monitored using a Hitachi U-1900 spectrophotometer until the signal became stable. From this, the concentration of TMP_{ox} was calculated based on its extinction coefficient of 14 mM^{-1} cm^{-1} (39) followed by determination of the equilibrium constant for the electron transfer reaction. Finally, from the known cell potential of 273 mV for the TMP_{ox}/TMP_{red} redox couple (40), the cell potential of *Nc*LPMO9C-Cu^{2+}/*Nc*LPMO9C-Cu^{1+} was determined as outlined previously (7, 17).

Substrate Degradation Studies—Prior to activity measurements, pure *Nc*LPMO9C and *Nc*LPMO9C-N were copper-saturated as described above for *Nc*LPMO9C. To analyze activity on polymeric substrates, we employed the fact that the C4-oxidizing *Nc*LPMO9C generates novel reducing ends (14). 5 mg/ml xyloglucan from tamarind seed (Megazyme International) or PASC was incubated with 4 μM copper-saturated *Nc*LPMO9C or *Nc*LPMO9C-N in 40 mM sodium phosphate, pH 6.5, and in the presence or absence of 2 mM ascorbic acid. Reactions were incubated at 50 °C and shaking at 1000 rpm. At appropriate time points, 40-μl samples were taken, and the enzyme reaction was quenched by addition of 60 μl of DNS solution (0.04 M 3,5-dinitrosalicylic acid in 1.0 M potassium-sodium tartrate, 0.4 M NaOH). After mixing, the samples were incubated for 10 min in boiling water. The samples were centrifuged, and 50 μl of the supernatant was transferred to a 384-well plate, and the absorbance at 540 nm was measured using a Multiscan FC microplate photometer (Thermo Scientific, Waltham, MA). Calibration was done with cellobiose (Glc_2; Megazyme) for PASC samples and a xyloglucan heptasaccharide called XG7 (a glucose tetramer where all glucose units, except the one at the reducing end, are substituted with xylose, *XXXG*; Megazyme) for xyloglucan samples. All results shown have been corrected for background absorbance from reactions containing only the substrate or substrate and ascorbic acid.

Reactions for degradation of soluble cello-oligosaccharides contained 0.2 mM cellopentaose (Glc_5) and 1.5 μM copper-sat-

TABLE 2
Diffraction data and refinement statistics for the two *Nc*LPMO9C-N structures

	1-metal structure	3-metal structure
Space group details		
Cell dimensions (a, b, c) (Å)	$a = 44.3$, $b = 66.8$, $c = 66.08$	$a = 44.3$, $b = 67.5$, $c = 65.6$
α, β, γ (°)	90, 101.3, 90	90, 100.9, 90
Space group	P21	P21
No. of molecules per asymmetric unit	2	2
Data collection and processing statistics		
Station (synchrotron)	I911-3 Lund max lab	I911-3, Lund max lab
Wavelength (Å)	1.000	1.000
No. of observations	166,092	198,749
No. of unique reflections	49,150	29,831
Maximum resolution (A)	1.56	1.9
Completeness (outermost shell) (%)	91.3 (29.2)	99.3 (98.4)
*R*merge (%)[a]	4.0	31.1
Mean((*I*)/sd(*I*)) (outermost shell)	15.5 (3.1)	6.8 (2)
CC1/2	99.9 (94.8)	98.9 (87.2)
Multiplicity	3.3 (1.8)	6.7 (6.6)
Outermost shell (Å)	1.59–1.56	1.94–1.9
Refinement statistics		
Resolution range (Å)	64.81–1.56	64.41–1.9
R_{work}/R_{free} (%)	15.7/18.6 (25/28)	21.3/26.8 (31/31)
R.m.s.d., bond lengths (Å)	0.005	0.01
R.m.s.d., bond angles (°)	1.1	1.44
No. of reflections	46,509	28,300
No. of protein atoms	3376	3362
No. of solvent molecules	561	427
No. of glycerol molecules	5 (2 in chain A, 3 in chain B)	2 (chain A)
No. of acetate ions		8 (3 in chain A and 5 in chain B)
No. of metal atoms	2 (copper)	6 (zinc)
Average *B* factor (Å^2) for protein residues		
Overall	12.6	9.7
Main chain atoms	12.6	9.8
Side chain atoms	12.6	9.7
Average *B* factor (Å^2) for heteroatoms		
Water molecules	24.14	18.6
Metal atoms	8.04	16.4
Glycerol molecules	16.67	23.2
Acetate ions		16.3
Ramachandran plot (%)		
Favored region	97.3	97.5
Allowed region	2.7	2.5
PDB entry	4D7U	4D7V

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$ (48), calculated using a strict boundary Ramachandran plot (48).

urated *Nc*LPMO9C or *Nc*LPMO9C-N in 20 mM MES, pH 6.0, with or without 1 mM ascorbic acid. Reactions were incubated at 50 °C and shaken at 1000 rpm. At appropriate time points, 50-μl samples were taken and mixed with 50 μl of 0.2 M NaOH to quench the enzyme reaction. Products were analyzed by high performance anion exchange chromatography using an ICS3000 system (Dionex, Sunnyvale, CA) as described previously (41). A 2.5-μl sample was injected onto a CarboPac PA1 2×250-mm analytical column (Dionex) coupled to a CarboPac PA1 2×50-mm guard column kept at 30 °C. Cello-oligosaccharides were eluted in 0.1 M NaOH at 0.25 ml/min by increasing the concentration of sodium acetate linearly from 0 to 0.1 M in 10 min and then from 0.1 to 0.18 M in 10 min, followed by a 2-min exponential gradient to 1 M sodium acetate. Between each run, the column was reconditioned by running initial conditions for 14 min. Cello-oligosaccharides with a degree of polymerization of 2–5 were used as standards.

Results

Overall Structural Analysis and Metal Binding—*Nc*LPMO9C-N was crystallized in space group $P2_1$ with approximate unit cell dimensions of $a = 44$ Å, $b = 67$ Å, $c = 66$ Å, $\beta = 101°$ and with two molecules in the asymmetric unit. We present two structures of *Nc*LPMO9C-N. The first one (PDB code 4D7U), obtained at 1.56 Å resolution from a crystal grown in the presence 0.2 M ammonium citrate and 20% PEG3350, has one copper ion bound in the catalytic center, similar to other known AA9 structures. The second structure (PDB code 4D7V), determined at 1.9 Å resolution from a crystal grown in the presence of 0.2 M zinc acetate and 20% PEG3350, contains three zinc ions per protein molecule. Data collection and refinement statistics for both structures are summarized in Table 2; Fig. 1 shows structural features that are discussed further below. The two structures are highly similar, with a root-mean-square deviation of 0.55 Å for all Cα atoms. The biggest variations occur in the loop that includes residues 64–78, which binds one of the zinc ions (Fig. 1, *c* and *d*). Similar to other LPMO9s of known structure, *Nc*LPMO9C-N folds into a β-sandwich of two β-sheets consisting of eight β-strands in total. The first β-sheet is formed of β-strands β1 (residues 3–10), β3 (residues 56–63), and β6 (residues 130–136), and the second one is composed of strands β5 (residues 106–114), β4 (residues 86–94), β7 (residues 143–153), and β8 (residues 164–175). In both structures, residues 180–185 display increased temperature factors, and the corresponding electron density was weak or even absent, especially for Ser-183 and Asn-184 in some monomers. Weak electron density is also

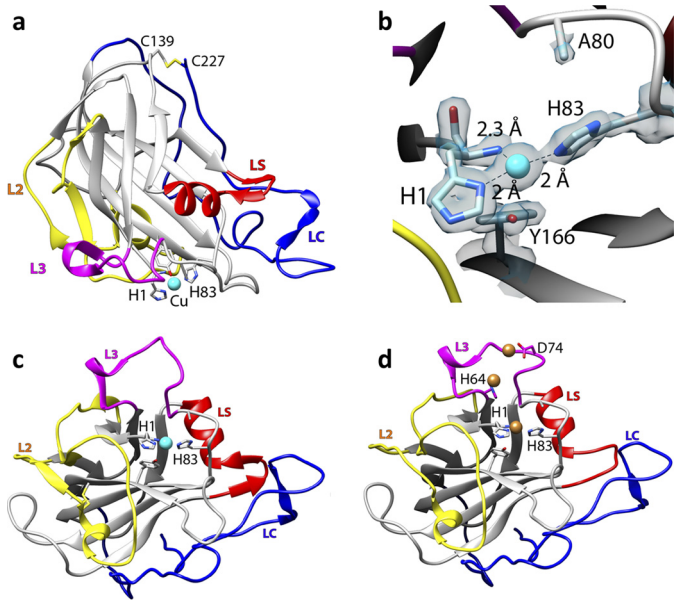

FIGURE 1. **Structural representations of NcLPMO9C-N.** *a*, cartoon representation of the copper-bound structure; copper is depicted as a *cyan sphere*; *b*, close up of the copper-binding site with the electron density map around the active site in *gray mesh* (contoured at 1σ); *c*, overall structure of the copper-loaded protein rotated by 90° along the horizontal axis compared with the view in *a*; *d*, structure of the zinc-loaded protein with the three bound zinc ions depicted as *brown spheres*; the orientation is similar to that in *c*. Note the structural variation in the loop coordinating the third zinc ion (residues 70–76 in the L3 loop, colored in *pink*).

observed in the region spanning residues 68–74 of monomer B of the three-metal structure. Neither of the two structures showed signs of glycosylation.

There are two disulfide bridges in the *Nc*LPMO9C-N structure. The first one, formed between Cys-39 and Cys-169 (partially cleaved due to x-ray radiation damage), is common among all known LPMO9 structures and connects loop L2 to strand β8 (Fig. 2). The second disulfide bond, formed between Cys-139 and the C-terminal amino acid, Cys-227, is observed in about half of LPMO9s with known structures (Fig. 2).

The copper ion lies in the center of a flat surface that interacts with the substrate (Fig. 1, *a* and *b*) (7). The copper site is very similar to that in other LPMOs, resembling an octahedral coordination environment with tetragonal distortion due to the Jahn-Teller effect and with four of six coordination positions occupied by protein ligands (5, 42). The equatorial plane includes the protein's N-terminal amino group (at 2.3 Å), the Nδ of His-1 (2.0 Å), and the Nϵ of His-83 (2.0 Å), forming what has been termed a "histidine-brace" (5). Although His-1 tends to be methylated in LPMO9s produced in fungi (5), His-1 in *Pichia*-produced *Nc*LPMO9C was not methylated, as observed previously for C1-oxidizing *Pc*LPMO9D (or *Pc*GH61D), which was also expressed in *P. pastoris* (42). There is no density in the fourth equatorial coordination position that is facing the solvent. Access to the protein-facing axial position is limited by the phenolic oxygen of a buried tyrosine, Tyr-166 (3.0 Å). In the solvent-facing axial position, there is extra density both in the σA-weighted $F_o - F_c$ map contoured at 3σ and the σA-weighted $2F_o - F_c$ map contoured at 1σ. This density could perhaps reflect a citrate molecule with partial occupancy, because there was 0.2 M citrate in the crystallization solution, but the data did not allow reliable modeling of a citrate molecule or any other compound in the electron density map.

In the structure of the zinc-soaked protein, the histidine brace is occupied by zinc instead of copper. Like the copper ion, the zinc ion is coordinated by the N-terminal amino group and the Nδ of His-1 (with distances of 2.2 and 2.1 Å, respectively) and the Nϵ of His-83 (2.1 Å). An overview over all distances related to metal binding is provided in Table 3. The fourth equatorial position and the solvent-facing axial position are both occupied by acetate ions with distances of 2.3 and 3.3 Å from the copper to the oxygens of the carboxyl group, respectively. The axially interacting acetate is replaced by a water molecule in some of the *Nc*LPMO9C-N molecules and was thus modeled with partial occupancy. The oxygen atom of the Tyr-166 side chain is located at a distance of 3.8 Å from the bound zinc atom. Superposition of the copper and zinc structures indicated that zinc is displaced by ~0.8 Å relative to the copper, explaining the longer distance to Tyr-166 (3.8 and 3.0 Å, respectively). Relative to the copper structure, the zinc-coordinating histidines show an 8° shift in their ψ dihedral angle ($\Delta\psi = -8°$).

The zinc structure reveals two possible additional metal-binding sites that are also located in the flat planar substrate-binding surface. The second zinc ion is coordinated by His-64 and is located at a distance of ~7.7 Å from the active site ion, whereas the third zinc ion is coordinated by Asp-74 and is 7.0 Å away from the second zinc atom (Fig. 1*d*). Both metals are coordinated by amino acids belonging to a sequence insertion that seems unique for strictly C4-oxidizing LPMO9s, also referred to as the LPMO2 subfamily (Figs. 1*d* and 2) (43). Notably, the EPR and ITC experiments described below strongly indicate that *Nc*LPMO9C binds only one copper ion with high affinity and that this is sufficient for activity.

Because *Nc*LPMO9C has been found to act on soluble substrates, including cellodextrins and xyloglucan fragments (13, 14), attempts were made to obtain crystallographic data for an enzyme·substrate complex. Despite the use of a variety of substrates (thio-linked cello-oligosaccharides with a degree of polymerization of 2–5, thio-linked xylopentaose, cello-oligosaccharides with a degree of polymerization of 5 and 7, xyloglucooligosaccharides, low molecular lignin, and β-glucan) and a massive number of crystallization trials using both soaking and co-crystallization, such data were not obtained. The protein molecules in the *Nc*LPMO9C-N crystals are very tightly packed, especially close to the catalytic center of the enzyme, with a distance of ~5 Å between two protein molecules. This distance is relatively small, and there is thus very little room left for a ligand molecule to diffuse into the crystal and bind in the active center of the enzyme.

Comparison with Other LPMO Structures—Using the Dali server (44), the closest structural homolog of *Nc*LPMO9C was found to be another C4-oxidizing LPMO, *Nc*LPMO9D (PDB code 4EIR) (4). The second closest structural homolog is *Nc*LPMO9M (NCU07898, PDB code 4EIS, 39% sequence identity; C1/C4-oxidizing (43)), followed by a *Thermoascus aurantiacus* LPMO (*Ta*GH61A, PDB code 3ZUD, 39% sequence identity; C1/C4-oxidizing (5)), GH61E from *Thielavia terrestris* (*Tt*GH61E, PDB code 3EJA, 44% sequence identity; predicted to be C1-oxidizing (18)), and Cel61B from *Hypocrea jecorina*

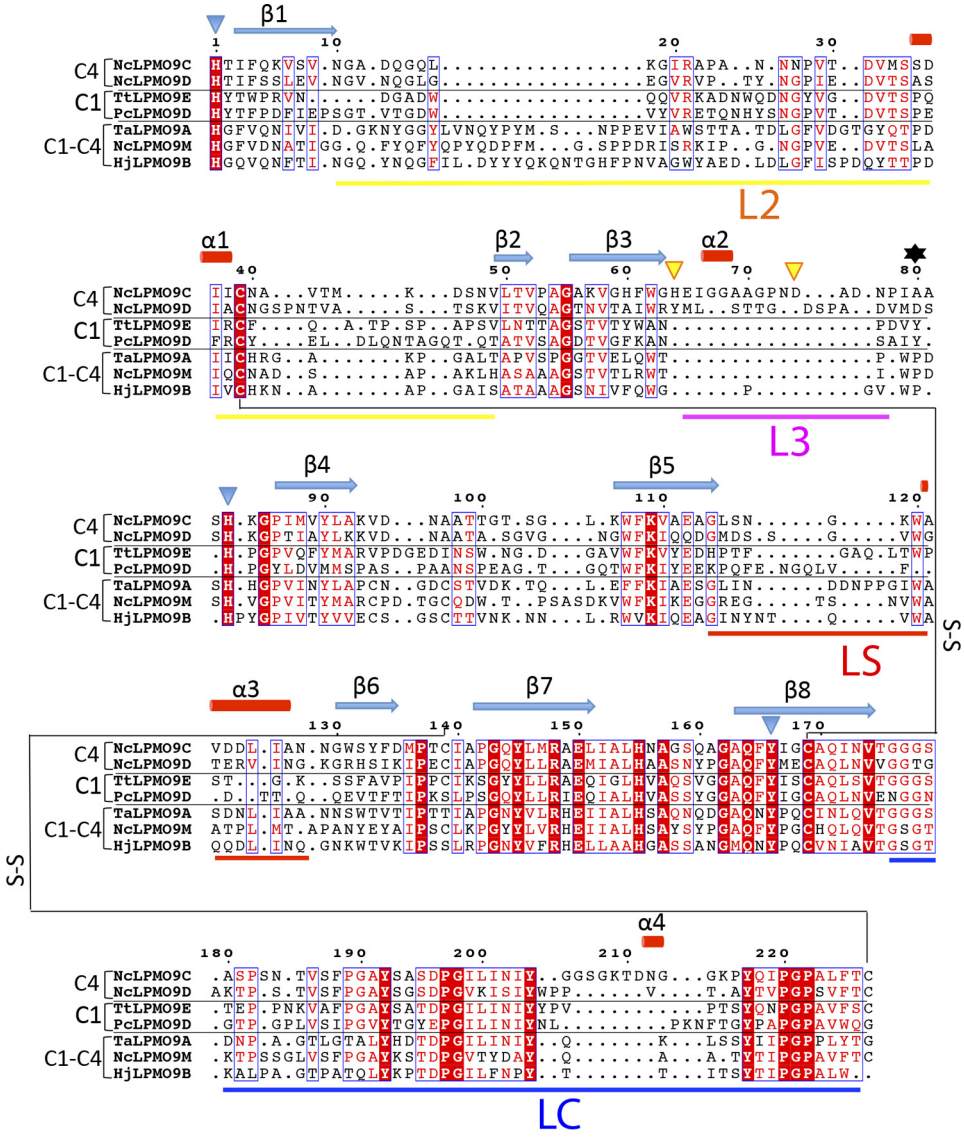

FIGURE 2. **Structure-based sequence alignment of LPMO9s with known structures.** The proteins included are as follows: *Nc*LPMO9C-N (PDB code 4D7U), *Nc*LPMO9D (PDB code 4EIR), *Tt*LPMO9E (PDB code 3EII), *Pc*LPMO9D (PDB code 4B5Q), *Ta*LPMO9A (PDB code 3ZUD), *Nc*LPMO9M (PDB code 4EIS), and *Hj*LPMO9B (PDB code 2VTC). Fully conserved residues are shown in *white* on a *red* background. *Blue frames* indicate that more than 70% of the residues in the corresponding columns exhibit similar physico-chemical properties (indicated as *red* residues on a *white* background). *Blue triangles* indicate residues coordinating the active site metal, and *yellow triangles* indicate residues involved in binding of two additional zinc ions. The secondary structure assignment (β-strands indicated as *blue arrows* and α-helices as *red cylinders*) refers to *Nc*LPMO9C-N and was determined with the program DSSP (56). The oxidative regio-specificity of the LPMOs, indicated on the *left*, was assigned based on experimental evidence (4, 5, 13, 57) or, for *Hj*LPMO9B and *Tt*LPMO9E, by inference from the sequence-based categorization (43, 45). The residue numbered *80*, which affects the accessibility of the solvent-facing axial copper coordination site, as shown in Fig. 4, *c* and *d*, and discussed in detail in the text, is indicated by a *black asterisk*. The loop regions that contribute to shaping the substrate-binding surface, named L2, LC, L3, and LS (see text), are marked by *horizontal bars below* the sequence, with color coding as in Fig. 2. The figure was prepared with ESPript.

TABLE 3

Distances (in Å) between the metal ion bound in the primary metal-binding site (*i.e.* the copper site) and the closest residues in the active site of *Nc*LPMO9C-N in the copper and the zinc structure

NcLPMO9C-N structure	PDB 4D7U (copper)		PDB 4D7V (zinc)	
	Chain A	Chain B	Chain A	Chain B
Me-N(His-1)	2.3	2.3	2.2	2.3
Me-Nδ(His-1)	2.0	2.0	2.1	2.2
Me-Nϵ(His-83)	2.0	2.0	2.1	2.1
Me-O (Tyr-166)	3.0	3.0	3.8	3.5

(*Hj*LPMO9B, PDB code 2VTC, 31% sequence identity; predicted to be C1/C4-oxidizing (18)).

Importantly, although the histidine brace and the tyrosine in the protein-facing axial position are structurally highly conserved, comparison of the seven available LPMO9 structures reveals differences in the accessibility of the solvent-facing axial position that correlate with the regioselectivity of the oxidation reaction. The two C1 oxidizers have a tyrosine in this position, whose hydroxyl group blocks access to the axial position to essentially the same extent as what is achieved by a conserved alanine in C1-oxidizing LPMO10s (11, 17). The mixed C1 and C4 oxidizers have a proline instead of a tyrosine providing better access to the axial position (Fig. 3, *c* and *d*). In strict C4 oxidizers, the axial position appears fully accessible due to the tyrosine being replaced by an alanine in *Nc*LPMO9C and an aspartate pointing away from the copper in *Nc*LPMO9D. Another strictly C4-oxidizing *N. crassa* LPMO9, named

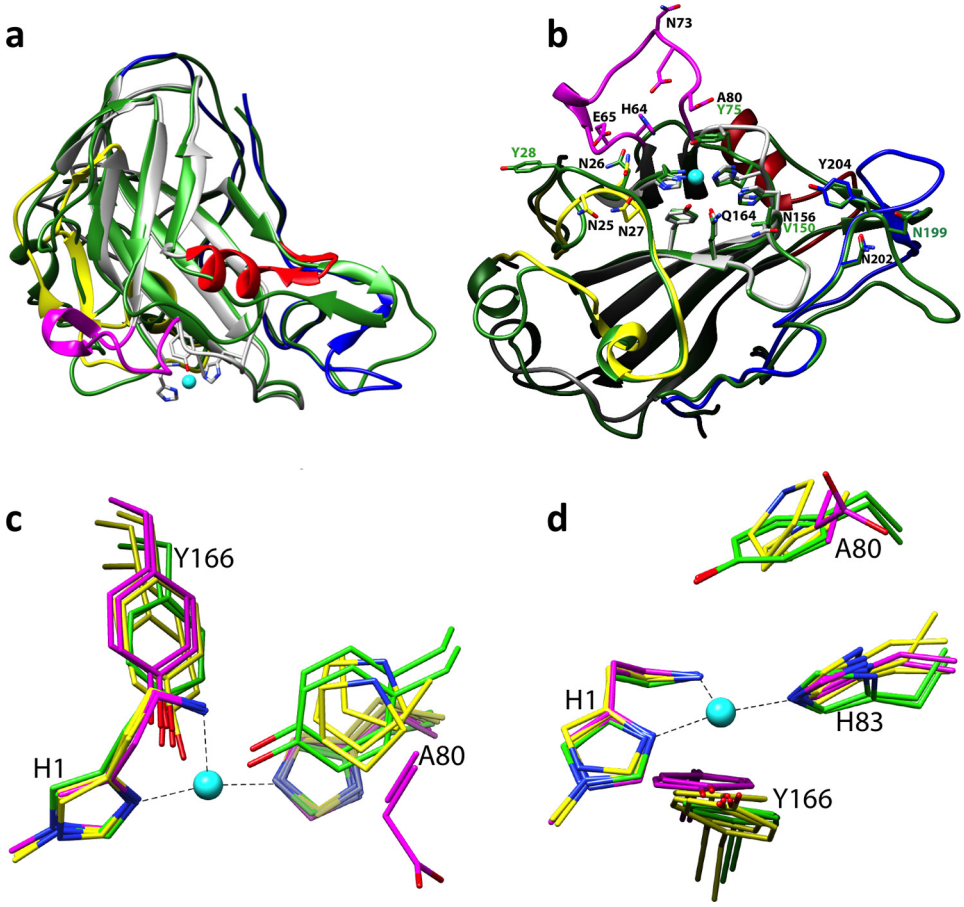

FIGURE 3. **Structural comparison of LPMO9s.** *a,* superposition of PcGH61D (*green*; PDB code 4B5Q) with *Nc*LPMO9C-N (colored as in Fig. 2). *b,* comparison of protruding surface residues in *Nc*LPMO9C (colored as in *a*) with residues at equivalent positions in C1-oxidizing *Pc*LPMO9D (*green*). *b* is rotated 90° along the *horizontal axis* compared with *a*. *c* and *d,* superposition of the copper sites of seven LPMO9s with known structure; C1, C1/C4, and C4 oxidizers are colored *green, yellow,* and *magenta*, respectively. The orientation shown in *c* is similar to that in *a*, whereas the orientation in *d* resembles that of *b*. Residue numbers refer to *Nc*LPMO9C, with the exception of *green labels* in *b*, which refer to *Pc*LPMO9D. The copper in *Nc*LPMO9C is shown as a *cyan sphere* in all panels.

NCU02240 (43), whose structure has not yet been determined, also has an aspartate in this position.

As observed in earlier studies (18, 42, 43) and as clearly visible in Figs. 1 and 3, LPMO9s show large variation in their putative substrate-binding surfaces. Three of these regions (Fig. 1) are known as the L2 (residues 10–49), LS (residues 114–128), and LC (residues 176–226) loops (42, 45). In addition, strictly C4-oxidizing AA9 LPMOs contain an insertion, referred to here as the L3 loop (resides 64–78), that potentially also affects the substrate-binding surface (Fig. 2). The variability in these loops is likely to provide a tunable scaffold for substrate specificity. Fig. 3*b* shows that the variation in these loops indeed results in quite different binding surfaces, when comparing *Pc*LPMO9D (active on cellulose but not on xyloglucan (14)) with *Nc*LPMO9C. Compared with *Pc*LPMO9D, the substrate-binding surface of *Nc*LPMO9C is more extended and crowded with polar residues. It should be noted that nothing is known about possible activities on xyloglucan for the other five LPMOs with known structures.

Further Insight into the Catalytic Center from EPR Spectroscopy, ITC Measurements, and Determination of Redox Potentials—EPR spectroscopy was used to assess whether the unique properties of *Nc*LPMO9C are reflected in the electronic structure of the bound copper ion. The spin Hamiltonian values $g_z = 2.267$ and $|A_z| = 152 \times 10^{-4}$ cm^{-1} obtained for the Cu^{2+}-loaded full-length enzyme are very similar to those obtained previously for cellulose-oxidizing LPMOs in the AA9 and AA10 families analyzed at pH values varying from 5.0 to 6.5 (5, 9, 17, 46) and indicate a type 2 copper active site. Spin quantification of Cu^{2+} in Cu^{2+}-loaded *Nc*LPMO9C was consistent with one copper-binding site/monomer. Interestingly, the addition of the soluble substrates cellohexaose or tamarind seed xyloglucan to Cu^{2+}-loaded *Nc*LPMO9C in the absence of reductant led to clear and similar changes in the EPR spectra (Fig. 4), indicative of structural changes in the Cu^{2+} active site and providing the first ever experimental evidence for substrate-binding effects on the copper site. The simulated spin Hamiltonian values changed to $g_z = 2.226$ and $|A_z| = 175 \times 10^{-4}$ cm^{-1} for both substrates upon binding (Fig. 4*a*) and superhyperfine splittings, reflecting the interaction between the unpaired electron and nitrogen nuclei adjacent to the Cu^{2+}, were greatly enhanced (Fig. 4*b*). The spectral envelopes for *Nc*LPMO9C interacting with D-cellohexaose or xyloglucan were nearly identical, showing that the two substrates affect the catalytic center of *Nc*LPMO9C in a similar manner.

ITC measurements of the binding of Cu^{2+} to *Nc*LPMO9C at pH 5.5 yielded data that fitted well to a single-site binding model, yielding the stoichiometry (n), the equilibrium binding

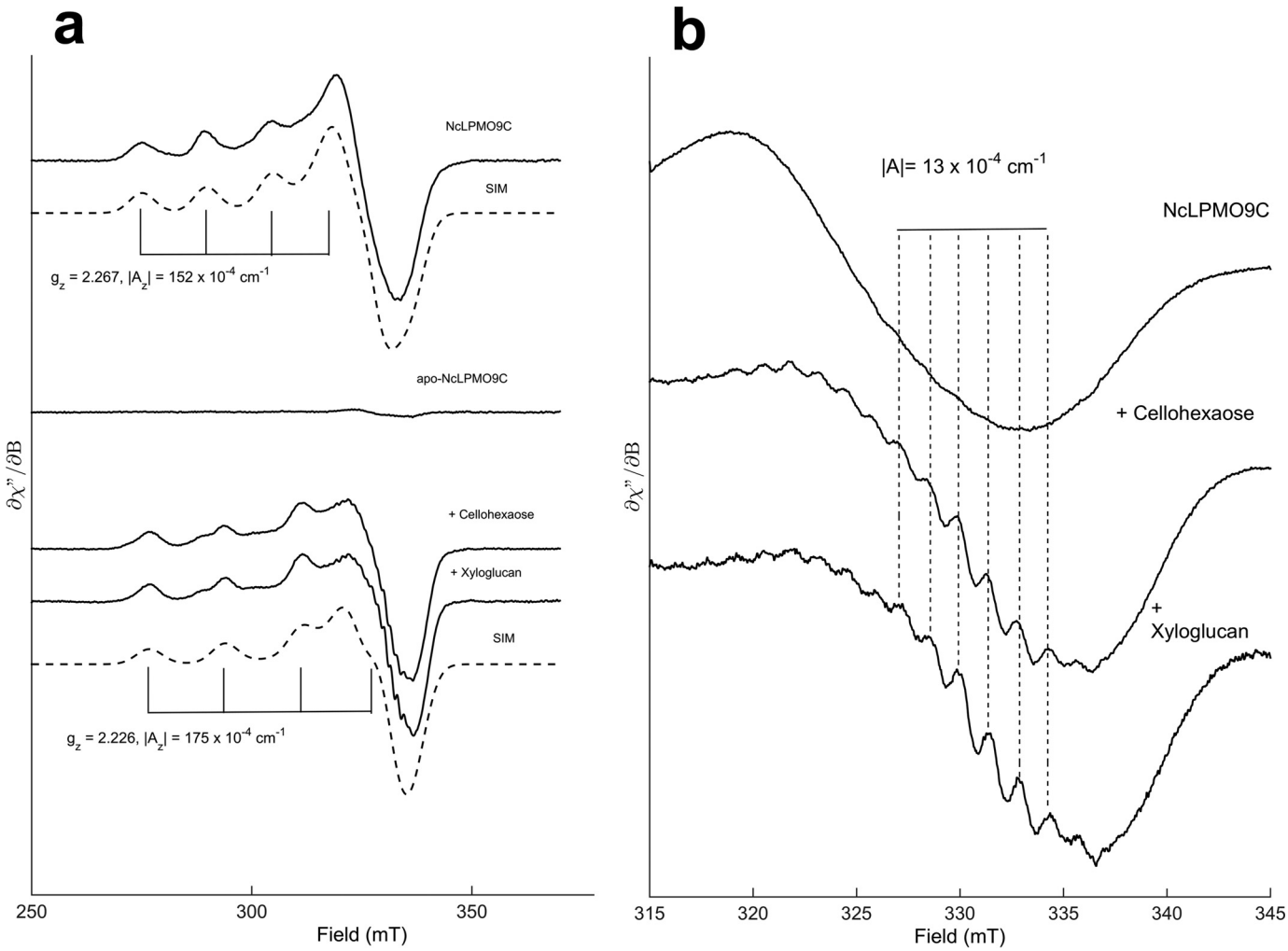

FIGURE 4. **NcLPMO9C EPR spectra.** *a,* EPR spectra of 160 μM Cu^{2+}-loaded *Nc*LPMO9C in the absence (*upper panel*) or presence (*lower panel*) of the soluble substrates cellohexaose (20 mg/ml) or xyloglucan (15 mg/ml). Spectra were recorded at 77 K, 1 milliwatt of microwave power, and 10 gauss modulation amplitude. The simulated spectra (*SIM*) for each species are shown below the corresponding experimental spectra (simulation parameters for the cellohexaose and xyloglucan spectra were identical). *b,* effect of substrate on super hyperfine splitting in the high field region. The splitting constants ($|A| \sim 13 \times 10^{-4}$ cm) are consistent with coupling to nearby nitrogen nuclei. Spectra were recorded at 77 K, 1 milliwatt of microwave power, and 2 gauss modulation amplitude.

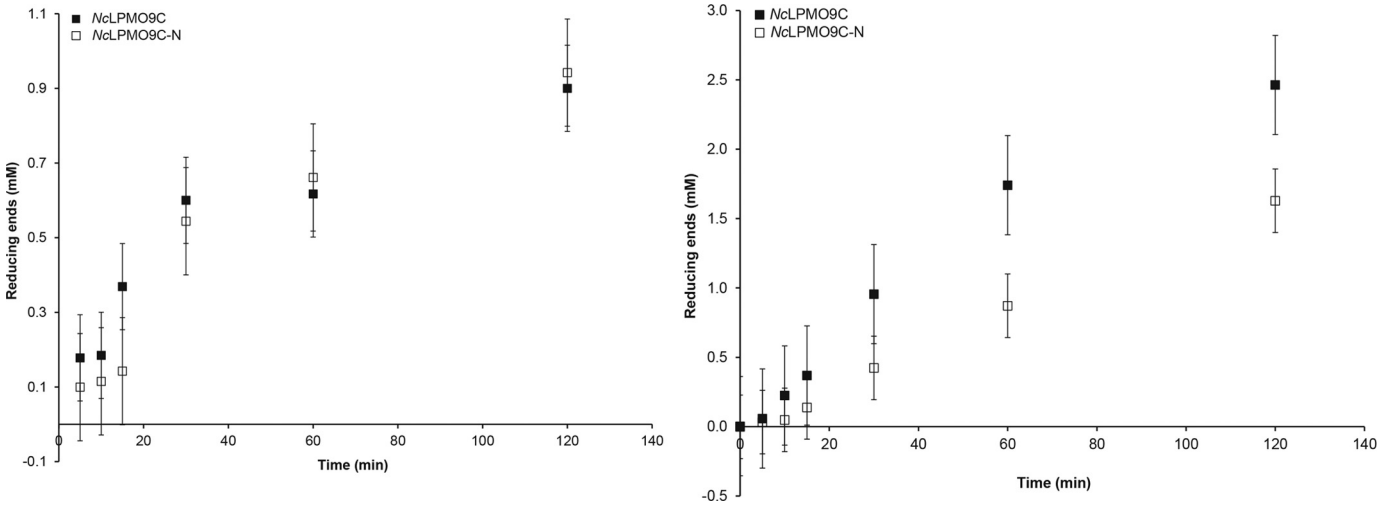

FIGURE 5. **Comparison of substrate degradation rates.** Degradation of 5 mg/ml PASC (*left*) or tamarind xyloglucan (*right*) by 4 μM *Nc*LPMO9C with or without (−*N*) the CBM domain at 50 °C was monitored by measuring the formation of reducing ends. In the absence of a reductant, the enzyme reactions did not yield reducing ends (not shown).

association constant (K_a), and the enthalpy change (ΔH_r^0) of the reaction (data not shown). The value of n was found to be between 0.9 and 1.1 per enzyme molecule. The dissociation constant was determined to be 33 ± 10 nM with $\Delta G_r^0 = -9.7 \pm 0.2$ kcal/mol, $\Delta H_r^0 = -5.4 \pm 0.2$ kcal/mol, and $-T\Delta S_r^0 = -4.3 \pm 0.3$ kcal/mol.

The redox potential for the *Nc*LPMO9C-Cu^{2+}/*Nc*LPMO9C-Cu^{+} redox couple was determined as described previously (7, 39), yielding a value of 224 ± 3 mV. Combining the redox potential and dissociation constants for Cu^{2+} in three thermodynamic relationships (7) allowed estimation of the dissociation constant for reduced copper (Cu^{+}), resulting in a value of 6.1 nM. These values for copper binding and the redox potential are similar to those previously found for LPMO10s at pH 5.5–6.0 (7, 9, 17).

Enzyme Activity and Substrate Specificity—Previous studies have shown that *Nc*LPMO9C is active on both xyloglucan and cellulosic substrates, including soluble β-glucans and cellodextrins (13, 14). To compare the activities on various substrates and to determine the influence of the CBM1 domain, reaction conditions were optimized so as to obtain reasonably linear initial progress curves allowing rate comparisons. This entailed incubating Cu^{2+}-saturated enzyme, with or without the binding module, with tamarind xyloglucan, PASC, or cellopentaose in the presence of ascorbic acid as reducing agent. Samples taken at appropriate time points were analyzed using the DNS assay (for polymeric substrates) or HPLC (for cellopentaose). The results show no effect of the removal of the CBM1 domain on the degradation rate of PASC (Fig. 5*a*) or cellopentaose, but for xyloglucan, removal of the binding module had a negative effect on the degradation rate (Fig. 5*b*). The degradation rate for the full-length enzyme on polymeric xyloglucan was ~0.1 s^{-1}, whereas all other reactions gave rates of ~0.04 s^{-1}.

The binding reaction between Cu^{2+}-saturated *Nc*LPMO9C and different substrates in the absence of reductant was assessed using ITC. The first substrate investigated was polymeric xyloglucan from tamarind seeds with a molecular mass of 225 kDa and estimated degree of polymerization of the main chain of 594 (calculated from the sugar composition). The binding dissociation constant at pH 5.5 and 10 °C was determined to be 2.3 ± 0.5 μM with $\Delta G_r^0 = -7.3 \pm 0.2$ kcal/mol, $\Delta H_r^0 = -0.4 \pm 0.1$ kcal/mol, $-T\Delta S_r^0 = -6.9 \pm 0.2$ kcal/mol, and a binding stoichiometry of 30 ± 8 (Fig. 6). The shape of the ITC binding curve is described by the so-called Wiseman c value (36), which can be expressed as follows: $c = nK_a[M]_t$ (1), where n is the stoichiometry of the reaction; K_a is the equilibrium binding association constant; and $[M]_t$ is the protein concentration. It is well established that c values within the range of $10 < c < 1000$ are a prerequisite for meaningful calculations of K_a (36). Under the given conditions, the Wiseman c value is 14. When the truncated version (*Nc*LPMO9C-N) was titrated against the same substrate, the binding isotherm changed from being sigmoidal to hyperbolic, indicating weaker binding and a c value below 10 (Fig. 6) (36, 47). It has been shown that binding thermodynamics can be obtained even if c is in the range of $0.01 < c < 10$ if a sufficient portion of the binding isotherm is used for analysis (47). This is achieved by ensuring a high molar ratio of ligand *versus* protein at the end of the titration, accurate knowledge of the concentrations of both ligand and receptor, an adequate level of signal-to-noise in the data, and known stoichiometry. Fitting of the theoretical data to the experimental data suggested a 10-fold lower binding affinity (~24 μM). Because the stoichiometry of the reaction is not known, this value is uncertain. Still, it is clear that the truncated form binds weaker to the xyloglucan than the full-length form.

Next, both full-length and truncated forms were titrated with PASC with an estimated average degree of polymerization of 200, at pH 5.5 and t = 25 °C. Interestingly, the full-length form binds in two separate modes (Fig. 6) where one is clearly stronger ($K_d = 0.013 \pm 0.004$ μM with $\Delta G_r^0 = -10.8 \pm 0.2$ kcal/mol, $\Delta H_r^0 = -1.1 \pm 0.1$ kcal/mol, $-T\Delta S_r^0 = -9.7 \pm 0.2$ kcal/mol, and $n = 2.6 \pm 0.6$) than the other ($K_d = 0.64 \pm 0.04$ μM, $\Delta G_r^0 = -8.4 \pm 0.1$ kcal/mol, $\Delta H_r^0 = -1.1 \pm 0.1$ kcal/mol, and $n = 11.4 \pm 0.2$). The truncated form only displayed a single binding mode ($K_d = 0.54 \pm 0.16$ μM, $\Delta G_r^0 = -8.5 \pm 0.2$ kcal/mol, $\Delta H_r^0 = -1.1 \pm 0.1$ kcal/mol, $-T\Delta S_r^0 = -7.4 \pm 0.2$ kcal/mol, and $n = 5.2 \pm 0.4$). All three fits yielded acceptable c values ranging from 60 to 900. The data suggest that upon removal of the CBM1, the high affinity binding mode is lost.

Binding of the soluble substrate Glc_6 to *Nc*LPMO9C (pH 5.5, t = 25 °C) was found to be much weaker, having a K_d of 0.81 ± 0.08 mM, with $\Delta G_r^0 = -4.3 \pm 0.2$ kcal/mol, $\Delta H_r^0 = -2.5 \pm 0.5$ kcal/mol, $-T\Delta S_r^0\Delta = -1.8 \pm 0.5$ kcal/mol (Fig. 6*b*). These values were obtained assuming a 1:1 binding stoichiometry between *Nc*LPMO9C and Glc_6. Considering the binding surface of *Nc*LPMO9C, it is conceivable that two Glc_6 molecules can bind simultaneously. Assuming a 2:1 binding stoichiometry, the fit yielded a virtually identical K_d value within experimental errors (0.85 ± 0.09 mM), albeit with a lower enthalpy change ($\Delta H_r^0 = -1.0 \pm 0.4$ kcal/mol) and a higher entropy change ($-T\Delta S_r^0 = -2.3 \pm 0.5$ kcal/mol).

Substrate Docking Experiments—An extensive computational docking study was carried out to identify potential substrate-binding sites on *Nc*LPMO9C-N. Using Autodock version 4.2 to compute and cluster grid-based free energies of enzyme·ligand complexes (33), we evaluated 10 different scenarios (Table 1) involving four known substrates and three different bound metal ion configurations (Fig. 7). For the substrates, both flexible and rigid conformational states were considered to examine the limitations of fixed torsion angles. Clustering analysis of the docked ligand position conformations (with a 0.5 Å root-mean-square deviation) and calculated binding free energies revealed no clearly preferred conformation for any of the considered ligand binding conditions. As in the case of structural studies attempting to capture *Nc*LPMO9C-N·substrate complexes, we were unable to conclusively identify a favorable binding orientation for any of the four ligands under any metal binding conditions. Increasing degrees of freedom through reduced dihedral restrictions were also unsuccessful in enabling the solution of a minimum energy-bound conformation.

Discussion

The primary aim of this study was to analyze the structural and biochemical features of *Nc*LPMO9C with its unprece-

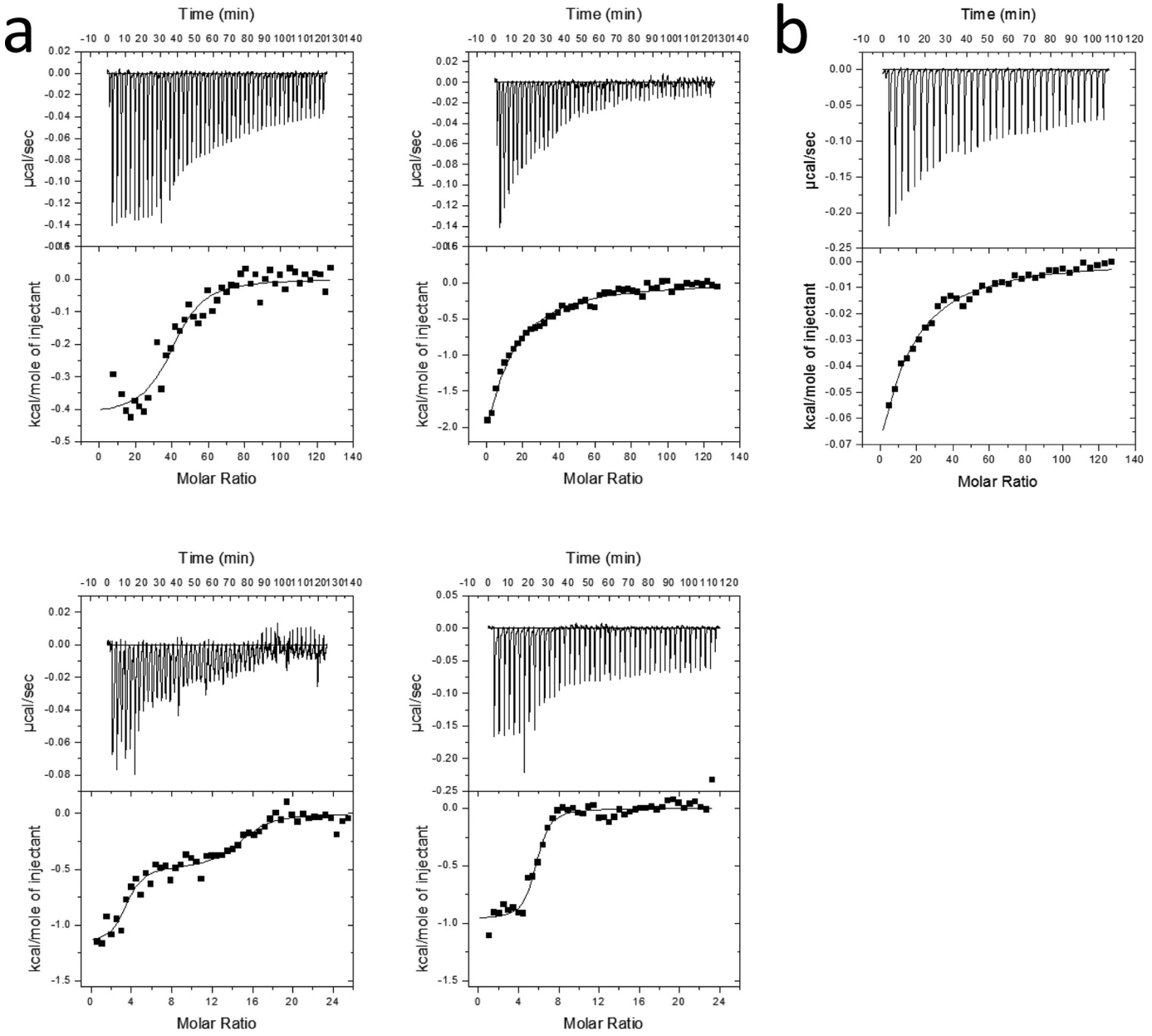

FIGURE 6. **Thermograms.** *a, upper panels*, binding isotherms with theoretical fits (*lower panels*) obtained for the titration of *Nc*LPMO9C (*left*) and *Nc*LPMO9C-N (*right*) into 0.9 μM xyloglucan (*top*) and 0.146 mg/ml PASC (*bottom*). The concentration of PASC was set to be 4.5 μM based on an estimated degree of polymerization of 200 (38). All experiments were carried out at pH 5.5. The temperature was 25 °C for PASC and 10 °C for xyloglucan; *b, upper panel,* binding isotherms with theoretical fits (*lower panel*) obtained for the binding of 11 mM Glc_6 to 30 μM *Nc*LPMO9C-Cu^{2+} at $t = 25$ °C.

dented substrate specificity toward soluble cello-oligosaccharides and hemicellulosic substrates. To this end, the crystal structure of the catalytic domain of the enzyme was determined, and several functional analyses were conducted. Considering the ability of *Nc*LPMO9C to bind productively to soluble substrates, one major goal of our work was to produce the first ever crystallographic insight into LPMO-substrate interactions. However, despite multiple efforts using a wide variety of experimental conditions, such insight was not obtained; no enzyme·substrate complexes could be captured through the structural studies, and docking experiments were inconclusive. In contrast, our EPR data clearly reveal effects of substrate binding on the configuration of the copper site, and the ITC data provide the first glimpse of substrate binding affinities involved in LPMO action.

After some initial uncertainty regarding the nature of the active site metal (3, 6), it is now well established that LPMOs are copper-dependent monooxygenases (5, 7, 9). Although binding of multiple bivalent metal ions has been observed in the crystal structures for several LPMOs (5, 17), available experimental and theoretical data consistently indicate that the active site copper centers are mononuclear (11, 12, 17, 49). The additional metal ions observed in the zinc structure of *Nc*LPMO9C are coordinated by amino acids belonging to a sequence insertion unique for strictly C4-oxidizing LPMO9s, but Fig. 2 shows that the residues interacting with zinc ions in *Nc*LPMO9C are not

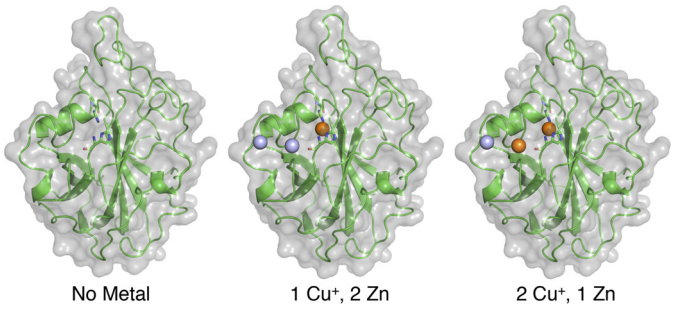

FIGURE 7. **Illustration of the three different bound metal scenarios considered in the computational docking study of *Nc*LPMO9C-N.**

conserved in C4-oxidizing *Nc*LPMO9D. The metal sites do not look like high affinity sites (only one ligand) and are too far away from the copper site to affect the catalytic center. Still, because the additional sites are located in the flat planar substrate-binding surface, it cannot be excluded that they play a role in enzyme-substrate interactions, thus perhaps co-determining the unique specificity of *Nc*LPMO9C.

Detailed functional characterization revealed that *Nc*LPMO9C resembles other well studied LPMOs in many ways. Enzyme rates are in the order of 0.04–0.1/s, *i.e.* similar to rates observed for other LPMOs (6, 50). Likewise, both the copper affinity and the redox potential of *Nc*LPMO9C resemble those of other LPMOs (7, 9, 17), although higher affinity for copper (<1 nM) has been reported for *Ta*LPMO9A (5). The dissociation constant for Cu^{2+} determined here closely resembles dissociation constants that were recently determined for a C1-oxidizing (31 nM) and a C1/C4-oxidizing (12 nM) cellulose-active LPMO10 at the same pH (5.5) (17). Notably, these LPMO10 enzymes are also very similar to *Nc*LPMO9C in terms of their spin Hamiltonian values.

The fact that *Nc*LPMO9C is active on soluble substrates allowed us to obtain a first glimpse of substrate binding affinities in LPMOs. We obtained reliable data for binding of full-length *Nc*LPMO9C to (soluble) polymeric xyloglucan, showing a dissociation constant of 2.3 ± 0.5 μM. Binding to cellohexaose was about 350 times weaker. Encouraged by these results, we also obtained ITC data for binding of (insoluble) PASC, revealing two binding modes, one of with a considerably lower K_d (0.013 ± 0.004 μM), relative to binding of xyloglucan. Interestingly, the data suggest that only the lower affinity binding mode for PASC is maintained in the truncated protein. The data for the truncated protein show that the catalytic domain of *Nc*LPMO9C has higher affinity for cellulose (K_d = 0.54 ± 0.16 μM) than for xyloglucan (K_d roughly estimated at 24 μM; Fig. 6*A*). The CBM1 improves the affinity for both substrates, and the effect seems larger for cellulose, as one could expect considering the known substrate specificities of CBM1s. We note that the affinities for cellulose determined in this study are high compared with literature data (51, 52) and that further work on these affinities is of major interest. Future studies could address the cooperativity between binding of the CBM1 and the catalytic domain and the impact of having different binding modes in the catalytic domain (*e.g.*, LPMOs *versus* processive glycoside hydrolases).

Low affinities for short soluble substrates offer a convenient explanation as to the failed attempts at obtaining crystals of enzyme·substrate complexes and may also explain why extensive computational docking studies did not reveal *Nc*LPMO9C-specific binding of cellotetraose or xyloglucan fragments at any point along the protein surface. Furthermore, low affinities for soluble substrates may be taken to indicate that the natural substrate of the LPMO is polymeric. The latter intuitively makes sense, because cellulolytic microbes have many other enzymes capable of cleaving soluble dextrins, whereas cleavage of polymeric glucans and their co-polymeric complexes is more challenging.

Measurement of LPMO kinetics is notoriously difficult, and the literature is almost devoid of kinetic data. Under the conditions used here to obtain reasonably linear progress curves, the differences in substrate affinity were hardly reflected in the obtained rates. Apparently, under these substrate concentrations, substrate binding is not the main rate-limiting factor. More detailed kinetic analyses, as well as more extended binding studies, preferably under "natural conditions," are needed to obtain further insight into the biological function of *Nc*LPMO9C, and other LPMOs, in biomass conversion. The activity measurements depicted in Fig. 5 indicate that the CBM1 domain of *Nc*LPMO9C is of importance for the activity on xyloglucan, possibly because the affinity of the truncated protein is so low that substrate binding becomes a rate-limiting factor.

In the quantum mechanical calculations described by Kjaergaard *et al.* (12), the closest modeled Cu^{2+} ligands in the equatorial plane include the three nitrogen atoms of the histidine brace and hydroxide. Kjaergaard *et al.* (12) also confirmed the presence of this hydroxide/water using extended x-ray absorption fine structure spectroscopy. Although density corresponding to a hydroxide/water molecule in the fourth equatorial position has been observed in some LPMO structures, such density was not observed in the structure of *Nc*LPMO9C, possibly because of photo-reduction of the copper by the x-ray beam (53). The EPR envelope though confirms a $d_{(x2-y2)}$ ground state ($g_z > g_x, g_y$) typical for an elongated octahedron. Thus, the singly occupied molecular orbital is centered at the Cu^{2+} and encompasses the N/O ligands in the equatorial plane.

The weak axial interaction between the copper ion and the hydroxyl group of Tyr-166 observed in the *Nc*LPMO9C crystal structure is in agreement with a tetragonal distortion of the Cu^{2+} site due to the Jahn-Teller effect. The spectral change observed when binding substrate to *Nc*LPMO9C reveals substrate-induced conformational changes. Such conformational changes may include a general distortion of the amino acid Cu^{2+} ligands or displacement/reordering of coordinating water molecules. The increased intensity of the superhyperfine splittings indicates that the structural change upon substrate binding enhances the interaction between the coordinating ^{14}N atoms of the histidine brace and the unpaired electron. Binding of cellohexaose and xyloglucan had similar effects on super hyperfine splitting in the g_x-g_y area (Fig. 4*b*). This suggests that the copper site only interacts with the (common) β-glucan part of the xyloglucan and that the chemical outcome of the reaction should be the same for both substrates, as is indeed observed (cleavage of a β(1→4) glucose-glucose bond with oxidation of C4 (14)).

All in all, the presented data indicate that the unique functionality of *Nc*LPMO9C is not due to properties of the enzyme's

catalytic site as such but rather to features of its substrate-binding surface and, at least to some extent, the presence of a CBM1. The most noticeable feature of the substrate-binding surface is that it is quite extended and highly polar (Fig. 3*b*), properties that align well with the ability to bind complex and branched polysaccharides such as xyloglucan. The structural determinants of the substrate specificity of LPMOs remain mysterious, and it should be noted that differences in these specificities might not only rely on binding as such but also on binding geometries being sufficiently optimized for binding to become productive.

Most interestingly, accumulating structural data for LPMO9s reveal structural differences in the copper coordination sphere that correlate with the oxidative outcome of the reaction (Fig. 3, *c* and *d*). In the strictly C1-oxidizing LPMO9s, access to the solvent-facing axial position is restricted by the hydroxyl group of a conserved tyrosine, whereas access to this same position seems unrestricted in strictly C4-oxidizing LPMO9s. In terms of access to the solvent-facing axial position, LPMO9s that oxidize C1 and C4 show an intermediate situation (Fig. 3). A similar difference has been observed when comparing a strictly C1-oxidizing LPMO10 with an LPMO10 that can oxidize both C1 and C4 (17). It has been shown that end-on binding of dioxygen is energetically favorable (12, 49); however, although Kim *et al.* (49) model O_2 binding at the exposed axial position, Kjaergaard *et al.* (12) propose a model where the equatorial OH ligand is replaced by O_2. This discrepancy is probably a result of different initial structural models and basis sets used in the density functional theory (DFT) calculations. The preferred binding mode (or modes) of O_2 has not been determined experimentally. It is tempting to speculate that O_2 binding in LPMOs with unrestricted access to the axial site leads to C4 oxidation and that C1 oxidation perhaps requires binding of dioxygen in another position. Alternatively, structural variation at the exposed axial site could affect the geometry of the enzyme·substrate complex, which could cause a different oxidative outcome, even if dioxygen binding to the copper does not vary. From an evolutionary point of view, development of both C1- and C4-oxidizing LPMOs seems logical, because different cellulases, such as cellobiohydrolases moving into opposite directions, will respond differently to the different oxidations (54, 55).

Author Contributions—A. S. B. and M. D. planned and performed all structural experiments, interpreted data, and helped write the manuscript. T. I. planned and performed the activity studies, performed the ITC experiments, interpreted data, and helped write the manuscript. Mo. S. designed, performed, and analyzed the ITC experiments and helped write the manuscript. A. A. K. and C. M. P. designed and carried out ligand docking experiments, interpreted the docking data, and helped write the manuscript. A. S. B., G. M., and A. V. carried out all molecular biology, protein expression and purification, and interpreted data; Å. K. R. planned and performed all EPR experiments, interpreted EPR data, and helped write the manuscript; Ma. S. and V. G. E. designed the study, planned the experiments, interpreted data, and wrote the manuscript. All authors have given approval to the final version of the manuscript.

Acknowledgments—C. M. P. and A. A. K. thank the August T. Larsson Guest Researcher Programme at the Swedish University of Agricultural Sciences for the opportunity to work alongside the Sandgren and Ståhlberg research groups. Computational time for this research was provided in part by the National Science Foundation through Extreme Science and Engineering Discovery Environment, which is supported by National Science Foundation Grant number ACI-1053575 under allocation number TG-MCB090159.

References

1. Horn, S. J., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V. G. (2012) Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5,** 45
2. Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S. J., and Eijsink, V. G. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20,** 1479–1483
3. Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J. C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., and Lo Leggio, L. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* **49,** 3305–3316
4. Phillips, C. M., Beeson, W. T., Cate, J. H., and Marletta, M. A. (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6,** 1399–1406
5. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., and Walton, P. H. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 15079–15084
6. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V. G. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330,** 219–222
7. Aachmann, F. L., Sørlie, M., Skjåk-Bræk, G., Eijsink, V. G., and Vaaje-Kolstad, G. (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 18779–18784
8. Beeson, W. T., Phillips, C. M., Cate, J. H., and Marletta, M. A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134,** 890–892
9. Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J., and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135,** 6069–6077
10. Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., and Sweeney, M. D. (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* **77,** 7007–7015
11. Hemsworth, G. R., Davies, G. J., and Walton, P. H. (2013) Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr. Opin. Struct. Biol.* **23,** 660–668
12. Kjaergaard, C. H., Qayyum, M. F., Wong, S. D., Xu, F., Hemsworth, G. R., Walton, D. J., Young, N. A., Davies, G. J., Walton, P. H., Johansen, K. S., Hodgson, K. O., Hedman, B., and Solomon, E. I. (2014) Spectroscopic and computational insight into the activation of O2 by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 8797–8802
13. Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., Ludwig, R., Haltrich, D., Eijsink, V. G., and Horn, S. J. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J. Biol. Chem.* **289,** 2632–2642
14. Agger, J. W., Isaksen, T., Várnai, A., Vidal-Melgosa, S., Willats, W. G., Ludwig, R., Horn, S. J., Eijsink, V. G., and Westereng, B. (2014) Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc. Natl. Acad. Sci. U.S.A.* **111,**

6287–6292
15. Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R., and Marletta, M. A. (2014) A family of starch-active polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 13822–13827
16. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6,** 41
17. Forsberg, Z., Mackenzie, A. K., Sørlie, M., Røhr, Å. K., Helland, R., Arvai, A. S., Vaaje-Kolstad, G., and Eijsink, V. G. (2014) Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 8446–8451
18. Li, X., Beeson, W. T., 4th., Phillips, C. M., Marletta, M. A., and Cate, J. H. (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* **20,** 1051–1061
19. Mattinen, M. L., Linder, M., Drakenberg, T., and Annila, A. (1998) Solution structure of the cellulose-binding domain of endoglucanase I from *Trichoderma reesei* and its interaction with cello-oligosaccharides. *Eur. J. Biochem.* **256,** 279–286
20. Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, J. K., and Teeri, T. T. (1992) Investigation of the function of mutated cellulose domains of *Trichoderma reesei* cellobiohydrolase I. *Proteins* **14,** 475–482
21. Sygmund, C., Kracher, D., Scheiblbrandner, S., Zahma, K., Felice, A. K., Harreither, W., Kittl, R., and Ludwig, R. (2012) Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Appl. Environ. Microbiol.* **78,** 6161–6171
22. Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., and Ludwig, R. (2012) Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol. Biofuels* **5,** 79
23. Várnai, A., Tang, C., Bengtsson, O., Atterton, A., Mathiesen, G., and Eijsink, V. G. (2014) Expression of endoglucanases in *Pichia pastoris* under control of the GAP promoter. *Microb. Cell Fact.* **13,** 57
24. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) in *The Proteomics Protocols Handbook* (Walker, J. M., ed) pp. 571–601, Humana Press Inc., Totowa, NJ
25. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
26. Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 282–292
27. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
28. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 355–367
29. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
31. Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
32. Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3,** 1171–1179
33. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30,** 2785–2791
34. Op't Holt, B. T., and Merz, K. M., Jr. (2007) Insights into Cu(I) exchange in HAH1 using quantum mechanical and molecular simulations. *Biochemistry* **46,** 8816–8826
35. Stoll, S., and Schweiger, A. (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178,** 42–55
36. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179,** 131–137
37. Wood, T. M. (1988) Preparation of crystalline, amorphous, and dyed cellulase substrates. *Methods Enzymol.* **160,** 19–25
38. Zhang, Y. H., and Lynd, L. R. (2005) Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. *Biomacromolecules* **6,** 1510–1515
39. Sørlie, M., Seefeldt, L. C., and Parker, V. D. (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. *Anal. Biochem.* **287,** 118–125
40. Liu, Y., Seefeldt, L. C., and Parker, V. D. (1997) Entropies of redox reactions between proteins and mediators: the temperature dependence of reversible electrode potentials in aqueous buffers. *Anal. Biochem.* **250,** 196–202
41. Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrøm, Y. H., and Eijsink, V. G. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J. Chromatogr. A* **1271,** 144–152
42. Wu, M., Beckham, G. T., Larsson, A. M., Ishida, T., Kim, S., Payne, C. M., Himmel, M. E., Crowley, M. F., Horn, S. J., Westereng, B., Igarashi, K., Samejima, M., Ståhlberg, J., Eijsink, V. G., and Sandgren, M. (2013) Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the *Basidiomycota fungus Phanerochaete chrysosporium*. *J. Biol. Chem.* **288,** 12828–12839
43. Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H., and Marletta, M. A. (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J. Am. Chem. Soc.* **136,** 562–565
44. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549
45. Morgenstern, I., Powlowski, J., and Tsang, A. (2014) Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monooxygenase family. *Brief. Funct. Genomics* **13,** 471–481
46. Forsberg, Z., Røhr, A. K., Mekasha, S., Andersson, K. K., Eijsink, V. G., Vaaje-Kolstad, G., and Sørlie, M. (2014) Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry* **53,** 1647–1656
47. Turnbull, W. B., and Daranas, A. H. (2003) On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125,** 14859–14866
48. Kleywegt, G. J., and Jones, T. A. (1996) ϕ/ψ-chology: Ramachandran revisited. *Structure* **4,** 1395–1400
49. Kim, S., Ståhlberg, J., Sandgren, M., Paton, R. S., and Beckham, G. T. (2014) Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 149–154
50. Loose, J. S., Forsberg, Z., Fraaije, M. W., Eijsink, V. G., and Vaaje-Kolstad, G. (2014) A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett.* **588,** 3435–3440
51. Guo, J., and Catchmark, J. M. (2013) Binding specificity and thermodynamics of cellulose-binding modules from *Trichoderma reesei* Cel7A and Cel6A. *Biomacromolecules* **14,** 1268–1277
52. Colussi, F., Sørensen, T. H., Alasepp, K., Kari, J., Cruys-Bagger, N., Windahl, M. S., Olsen, J. P., Borch, K., and Westh, P. (2015) Probing substrate interactions in the active tunnel of a catalytically deficient cellobiohydrolase (Cel7). *J. Biol. Chem.* **290,** 2444–2454
53. Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M. H., Vaaje-Kolstad, G., Lundberg, D., Royant, A., Ståhlberg, J., Eijsink, V. G., Beckham, G. T., and Sandgren, M. (2014) Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by x-ray photoreduction. *J. Biol. Chem.* **289,** 18782–18792
54. Cannella, D., Hsieh, C. W., Felby, C., and Jørgensen, H. (2012) Production

and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol. Biofuels* **5,** 26

55. Vermaas, J. V., Crowley, M. F., Beckham, G. T., and Payne, C. M. (2015) Effects of lytic polysaccharide monooxygenase oxidation on cellulose structure and binding of oxidized cellulose oligomers to cellulases. *J. Phys. Chem. B* **119,** 6129–6143
56. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637
57. Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G., Igarashi, K., Samejima, M., Ståhlberg, J., Horn, S. J., and Sandgren, M. (2011) The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE* **6,** e27807