Discovery and characterization of cellulose-active lytic polysaccharide monooxygenases

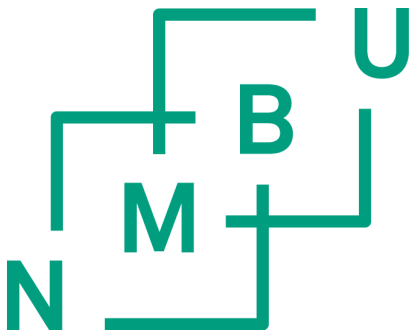
Upptäckt och karaktärisering av cellulosa-aktiva lytiska polysackarid monooxygenaser

Philosophiae Doctor (PhD) Thesis

Zarah Kristina Forsberg

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

Ås 2014



Thesis number 2014:62
ISSN 1894-6402
ISBN 978-82-575-1225-5

Table of Contents

ACKNOWLEDGEMENTS

The present work was carried out during the period of 2010-2014 in the Protein Engineering and Proteomics group (PEP), Department of Chemistry, Biotechnology and Food Science at The Norwegian University of Life Sciences (NMBU). The project was funded through a grant from the VISTA programme, which is funded by Statoil and managed by the Norwegian Academy of Science and Letters (DNVA), with grant number 6505.

First of all, I would like to express my gratitude to my supervisor Professor **Vincent Eijsink** for offering me a PhD position in this new and exciting research field. I would like to thank you for your outstanding supervision, brilliant ideas, supportive and inspiring talks, and for the amount of work you have put into helping me achieve today's result. It has been a pleasure working with you. I would also like to express my gratefulness to my co-supervisor Dr. **Gustav Vaaje-Kolstad**. Thank you for being so patient, and for always sharing your valuable time with me. Your motivation and positive attitude have been able to neutralize my sometimes overly negative vibes.

I would also like to thank Prof**. Morten Sørlie**, whose door has always been open for me and who has patiently answered my questions over and over again. My special thanks go to Dr. **Al Mackenzie** (the ÄKTA man!) and Dr. **Bjørge Westereng** for helpful discussions and support in the lab, and to Dr. **Åsmund Røhr** for coaching me in the work performed at the University of Oslo.

Thanks also to **Mats Sandgren** and **Jerry Ståhlberg** at the Swedish University of Agricultural Sciences, who endeared carbohydrate active enzymes to me and introduced me to this PhD position.

Special thanks to **Jenny Loose**, **Anikó Várnai** and **Sophanit Mekasha** for making the evenings spent in the lab a lot more fun, for their friendship and the times we enjoyed together outside the lab. Thanks to **Tina Rise Tuveng** for sharing the office with me during the challenging phase of writing this thesis. I would also like to thank **Anne-Cath** & **Ellen** and my colleagues in the PEP group for helping out in the lab and making work fun.

A special thanks to my family. Words cannot express how grateful I am to my parents, **Karin Forsberg** and **Bo Friberg**, for all the sacrifices that you have made on my behalf and for your support throughout life, not just these last few years. Thanks for listening to all ups and downs in my research, even though you claim not to have understood a single word of it. Last but by no means least, I would like express my gratitude to my very best friend **Ida Sarapik**, for simply being my friend and for opening up my eyes to other things equally important in life.

Finally, my appreciation goes to the innumerable people who supported me during my education, and incented me to strive towards my goals.

Zarah Forsberg

Ås, June 2014

SUMMARY

The efficient depolymerization of lignocellulosic biomass to fermentable sugars by enzymatic hydrolysis is a key step in the transition towards a more environmentally friendly and sustainable bio-economy. However, the complexity and recalcitrant nature of the substrate limit enzyme performance on lignocellulosic plant biomass, and at present the enzyme cocktails required for depolymerization represent a major cost in the production of biomass-based chemicals and fuels. The recent discovery of lytic polysaccharide monooxygenases (LPMOs) has changed our general understanding of polysaccharide deconstruction, and given rise to high expectations for further development of enzyme tools for biomass processing, since LPMOs enhance the activity of glycoside hydrolases.

LPMOs are copper-dependent enzymes that oxidize recalcitrant polysaccharides such as chitin and cellulose in the presence of dioxygen, and an external electron donor. Before the discovery of their enzymatic function, in 2010, LPMOs were classified as either family 33 of carbohydrate-binding modules (now family 10 of auxiliary activities, AA10, LPMO10) or family 61 of glycoside hydrolases (now AA9, LPMO9). Prior to the studies presented here, catalytic activity had just been demonstrated for a chitin-active bacterial AA10-type of LPMO from *Serratia marcescens* called CBP21. The work on CBP21 formed the basis for the first goal of this study, namely finding or engineering an LPMO targeting cellulose substrates. Paper I describes CelS2, a naturally occurring AA10-type LPMO from the Gram-positive bacterium *Streptomyces coelicolor* that cleaves crystalline cellulose and produces C1-oxidized cello-oligosaccharides appearing in solution as aldonic acids. The generation of oxidized products was demonstrated using both mass spectrometry and chromatographic methods. CelS2, which comprises an N-terminal AA10 and a C-terminal cellulose-binding carbohydrate-binding module classified as CBM2, represents the first described LPMO that is active on cellulose. It was shown that CelS2 stimulates the release of soluble sugars from filter paper by Celluclast® (a commercial enzyme cocktail).

Papers II and III of this study describe structure-function studies of cellulose-active AA10-type LPMOs with the purpose of unraveling the basic characteristics of these proteins and perhaps identify factors determining substrate specificity and the regioselectivity of hydroxylation (C1 versus C4 oxidation). Paper II describes a comparative study of four C1-oxidizing LPMOs, two of which are active on chitin and two on cellulose, and includes the description of one novel chitin-active LPMO10 (*Bl*LPMO10A from *Bacillus licheniformis*) and one novel cellulose-active LPMO10 (E8 from *Thermobifida fusca*). Sequence analysis showed that all residues in the immediate copper coordination sphere were conserved in these four LPMOs. Conversely, electron paramagnetic resonance spectroscopy (EPR) analyses indicated that the electronic environments of the copper differed between the chitin- and cellulose-active LPMOs. The differences in the EPR spectra are thus likely to reflect variation in residues outside the direct copper coordination sphere, where the chitin-active and cellulose-active AA10-type of LPMOs indeed show considerable variation.

Paper III presents the first crystal structures of cellulose-active AA10-type LPMOs, which allowed for the first time a structural comparison of LPMOs with different substrate specificities. The two *S. coelicolor* LPMO for which the structures were determined, CelS2 and *Sc*LPMO10B, represent a conserved pair of LPMOs found in cellulolytic actinomycetes. The two enzymes are upregulated during growth on cellulose substrates and we show that they act synergistically when degrading cellulose. CelS2 shows strict C1-oxidation on cellulose substrates, whereas *Sc*LPMO10B catalyzes oxidation of C1 and C4 in cellulose, as well as C1-oxidation on β-chitin. A structural comparison of the two cellulose-active LPMO10s revealed a difference in the copper coordination sphere that may relate to the (in)ability to oxidize C4. Structural comparisons of chitin-active and cellulose-active LPMO10s revealed a potential binding-pocket for a C2 acetamido group in chitin-active LPMO10s only. All LPMO10s had similar redox potentials and copper binding affinities, but showed a substrate-dependent difference in EPR spectra, as discussed in Paper II. Substrate-specificity thus seems to be determined by variation in substrate-binding and –positioning combined with variation in the electronic structure of the copper site.

In conclusion, this study represents the discovery and first in-depth characterization of LPMOs from family 10 of auxiliary activities that are active on cellulose. The work presented here has provided fundamental insight into how these enzymes work and contributed to method development, thereby constituting an important basis for future LPMO research.

SAMMANFATTNING

Effektiv enzymatisk depolymerisation av lignocellulosisk biomassa till fermenterbara sockerarter är ett viktigt steg i övergången till en mer miljövänlig och hållbar bioekonomi. Komplexiteten och resistensen av substratet begränsar dock enzym sammansättningarna som finns tillgängliga idag, vilket leder till att dessa preparat utgör en av de största kostnaderna vid produktion av biomassa-baserade kemikalier och bränslen. De nyligen upptäckta lytiska polysackarid-monooxygenaserna (LPMO-enzymer) har förändrat vår generella bild av polysackaridnedbrytning då dessa enzymer ökar aktiviteten till glykosidhydrolaser. LPMO-enzymer har till följd av detta gett höga förväntningar för vidare utveckling av enzym-verktyg för behandling av biomassa.

LPMO-enzymer är koppar-beroende enzymer som i närvaro av syrgas och en extern elektrondonator oxiderar svårnedbrytbara polysackarider såsom kitin och cellulosa. Innan upptäckten av dess enzymatiska funktion, år 2010, var LPMO-enzymer klassificerade som familj 33 kolhydratsbindande moduler (numera familj 10 av auxiliär-aktiviteter, AA10, LPMO10) eller familj 61 av glykosidhydrolaser (numera AA9, LPMO9). Innan studierna som presenteras här påbörjades hade katalytisk aktivitet precis visats för en kitin-aktiv bakteriell LPMO10, kallad CBP21, från *Serratia marcescens*. Denna upptäckt lade grunden till det första målet för denna studie, nämligen att hitta eller konstruera en LPMO aktiv på cellulosa-substrat. Artikel I beskriver CelS2, en naturligt förekommande LPMO10 från den Gram-positiva bakterien *Streptomyces coelicolor*. Detta enzym klyver kristallinsk cellulosa och producerar C1-oxiderade cello-oligosackarider som i lösning ses i form av aldonsyror. Genereringen av oxiderade produkter demonstrerades med både masspektrometri och kromatografiska metoder. CelS2 består av en N-terminal AA10 och en C-terminal cellulosa-bindande kolhydrats-bindande modul, klassificerad som en CBM2, och representerar den första beskrivna LPMO som är aktiv på cellulosa. Det visades även att CelS2 stimulerade Celluclast® (ett kommersiellt enzym preparat) vid frigörandet av lösliga socker från hydrolys av filterpapper.

Artikel II och III beskriver struktur-funktionsstudier av cellulosa-aktiva LPMO10-enzymer med syfte att ta reda på grundläggande egenskaper tillhörande dessa proteiner, samt att om möjligt identifiera faktorer som påverkar substratspecificitet och regioselektiv hydroxylering (d.v.s. C1- kontra C4-oxidation). Artikel II beskriver en jämförande studie av fyra C1-oxiderande LPMO10-enzymer, varav två är aktiva på kitin och två är aktiva på cellulosa. Aktiviteten hos två av dessa har tidigare inte dokumenterats, en visade aktivitet på kitin (*Bl*LPMO10A från *Bacillus licheniformis*) och en på cellulosa (E8 från *Thermobifida fusca*). Sekvensanalys visade att alla aminosyror i den direkt koppar-koordinerande sfären är konserverade i de fyra LPMO10-enzymerna, men elektronparamagnetisk resonans (EPR) spektroskopi indikerade att den elektroniska miljön kring kopparatomen skiljde sig mellan kitin-aktiva och cellulosa-aktiva LPMO-enzymer. Det är därför troligt att skillnaden i EPR-spektra mellan kitin-aktiva och cellulosa-aktiva

LPMO-enzymer reflekterar variation i aminosyror utanför den direkt koppar-koordinerande sfären.

I Artikel III presenteras de första kristallstrukturerna på cellulosa-aktiva LPMO10-enzymer, vilket för fösta gången gjorde det möjligt att utföra en strukturell jämförelse av LPMO-enzymer med olika substratspecificiteter. De två *S. coelicolor* LPMO-enzymerna vars struktur bestämdes, CelS2 och *Sc*LPMO10B, utgör ett konserverat par av LPMO10-enzymer som förekommer i cellulolytiska aktinomyceter. De två enzymerna är uppreglerade under bakteriell tillväxt på cellulosa och i denna studie visar vi att detta par agerar synergistiskt vid nedbrytning av cellulosa. CelS2 visar strikt C1-oxidation av cellulosa-substrat medan *Sc*LPMO10B katalyserar oxidation på både C1- och C4-kolet i cellulosa samt C1-oxidation av β-kitin. En strukturell jämförelse av de två cellulosa-aktiva LPMO10-enzymerna avslöjade en skillnad i koppar-koordinationssfären som kan förklara (o)förmågan att oxidera C4-kolet. Strukturella jämförelser av kitin-aktiva och cellulosa-aktiva LPMO10-enzymer visade en potentiell bindningsficka för en C2 acetamido-grupp i kitin-aktiva LPMO10-enzymer. Alla LPMO10-enzymer har likvärdiga redox-potentialer och bindningsaffiniteter för koppar, men visar en substrat-beroende skillnad i EPR-spektra som diskuterat i Artikel II. Till följd av detta verkar substratspecificitet vara beroende av variation i substrat-bindning och -positionering samt dessa i kombination med variation i den elektroniska strukturen för kopparsätet.

Sammantaget utgör denna studie upptäckten av det första cellulosa-aktiva LPMO-enzymet tillhörande familj 10 av auxiliär-aktiviteter, samt det första djupgående karaktäriseringsarbetet kring dessa typer av enzymer. Det föreliggande arbetet har på så sätt bidragit till grundläggande kännedom om dessa enzymer, samt till metodutveckling och utgör därmed en viktig bas för framtida LPMO-forskning.

ABBREVIATIONS

AA – Auxiliary Activity

Ao – Aspergillus oryzae

Ba – Bacillus amyloliquefaciens

Bl – Bacillus licheniformis

CAZyme – Carbohydrate Active enZyme

CBM – Carbohydrate-Binding Module

CDH – Cellobiose dehydrogenase

DP – Degree of Polymerization

Ef – Enterococcus faecalis

EPR – Electron Paramagnetic Resonance

GH – Glycoside Hydrolase

Glc – Glucose (D-glucose)

Glc1A – Gluconic acid (D-gluconic acid)

Glc4Gem – Gemdiol glucose (4-hydroxy-β-D-xylo-hexopyranosyl)

GlcNAc – *N*-acetylglucosamine (N-Acetyl-D-Glucosamine)

Hj – Hypocrea jecorina (also known as *Tr - Trichoderma reesei*)

ITC – Isothermal Titration Calorimetry

K_d – Dissociation constant

LPMO – Lytic polysaccharide monooxygenase

MALDI-TOF – Matrix-Assisted Laser Desorption/Ionization Time of Flight

MS - Mass Spectrometry

Nc – Neurospora crassa

NMR – Nuclear Magnetic Resonance

Pc – Phanerochaete chrysosporium

Sc – Streptomyces coelicolor

Sm – Serratia marcescens

Ta – Thermoascus aurantiacus

Tf – Thermobifida fusca

Tt – Thielavia terrestris

LIST OF PAPERS

Paper I

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

Paper II

Forsberg, Z., Røhr, Å. K., Mekasha, S., Andersson, K. K., Eijsink, V. G. H., 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.

Paper III

Forsberg, Z., Mackenzie, A. K., Sørlie, M., Røhr, Å. K., Helland, R., Arvai, A. S., Vaaje-Kolstad, G., and Eijsink, V. G. H., 2014, Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences of the United States of America* published ahead of print May 27, 2014, doi:10.1073/pnas.1402771111.

Other publications by the author

Vasur, J., Kawai, R., Jonsson, K. H., Widmalm, G., Engström, A., Frank, M., Andersson, E., Hansson, H., **Forsberg, Z.**, Igarashi, K., Samejima, M., Sandgren, M. and Ståhlberg, J. 2010. Synthesis of cyclic beta-glucan using laminarinase 16A glycosynthase mutant from the basidiomycete *Phanerochaete chrysosporium*. *Journal of the American Chemical Society* 132:1724-1730.

Vaaje-Kolstad, G., Westereng, B., Eijsink, V.G.H., Horn, S.J., Sørlie, M., **Forsberg, Z.** 2011. Methods of Degrading or Hydrolyzing a Polysaccharide; PCT/US2011/046838.

Aachmann, F. L., Vaaje-Kolstad, G., **Forsberg, Z.**, Røhr, Å., Eijsink, V. G. H., Sørlie, M. 2014. Lytic polysaccharide monooxygenase. *Encyclopedia of Inorganic and Bioinorganic Chemistry*; accepted provided minor revision.

Courtade, G., Balzer, S., **Forsberg, Z**., Vaaje-Kolstad., Eijsink, V. G. H., Aachmann, F. L. 2014. ^{1}H, ^{13}C, ^{15}N resonance assignment of the chitin-active lytic polysaccharide monooxygenase *Bl*LPMO10A from *Bacillus licheniformis;* submitted to *Biomolecular NMR Assignments*.

1 INTRODUCTION

1.1 The transition towards a sustainable bio-economy

The ever-increasing global wealth and population is generating challenges of providing sufficient energy and food to sustain our current lifestyle. Currently, the world economy is fueled by hydrocarbons, which have been formed over millions of years through decomposition of plants and animals. Oil reservoirs are being depleted at an increasing rate resulting in increased emission of CO_2, one of the key drivers of the global climate change (Vanholme et al., 2013). This situation has led to an increased interest in the use of renewable and sustainable energy forms, such as solar, wind, geothermal, hydroelectric and wave energy. However, none of these alternative energy sources can alone fulfill contemporary energy needs, nor do they provide a renewable source of materials, which is another prerequisite for a sustainable economy (Lund, 2007, Chang et al., 2010). Therefore, another alternative, namely the use of plant biomass, has come into focus (Figure 1). Plant biomass is easily available worldwide and cellulose, the main component in the plant cell wall, is the most abundant biopolymer on earth. Plant biomass is produced from photosynthesis, where atmospheric carbon dioxide is converted to carbohydrates using solar energy.

The bio-based economy is today mainly using easily accessible sugars, derived from so called “first generation” feedstocks such as corn and sugarcane carbohydrates, for ethanol production and rapeseed oil for the production of biodiesel (Himmel et al., 2007, Solomon, 2010). The use of first generation resources leads to a direct negative effect as it competes with food and feed industry (Williams, 2008). For that reason, research efforts in the past decade have focused on exploiting lignocellulosic plant biomass, also known as “second generation” biomass, which mainly derives from agricultural and forestry sources. The annual global production of plant biomass is estimated to 170-200 $\times$ 10^9 tons (Lieth and Whittaker, 1975), whereof about 70% consists of plant cell walls with about three quarters of these plant cell walls being polysaccharides (Duchesne and Larson, 1989). The central mono-sugar in the plant cell wall is glucose, which is an ideal carbon source that easily can be converted to ethanol and a variety of other chemical compounds by microorganisms and enzymes. Glucose occurs in (homopolymeric) cellulose as well as in certain heteropolymeric hemicelluloses such as xyloglucan. Other abundant mono-sugars in plant cell walls are xylose and mannose, which occur in the hemicellulose fraction.

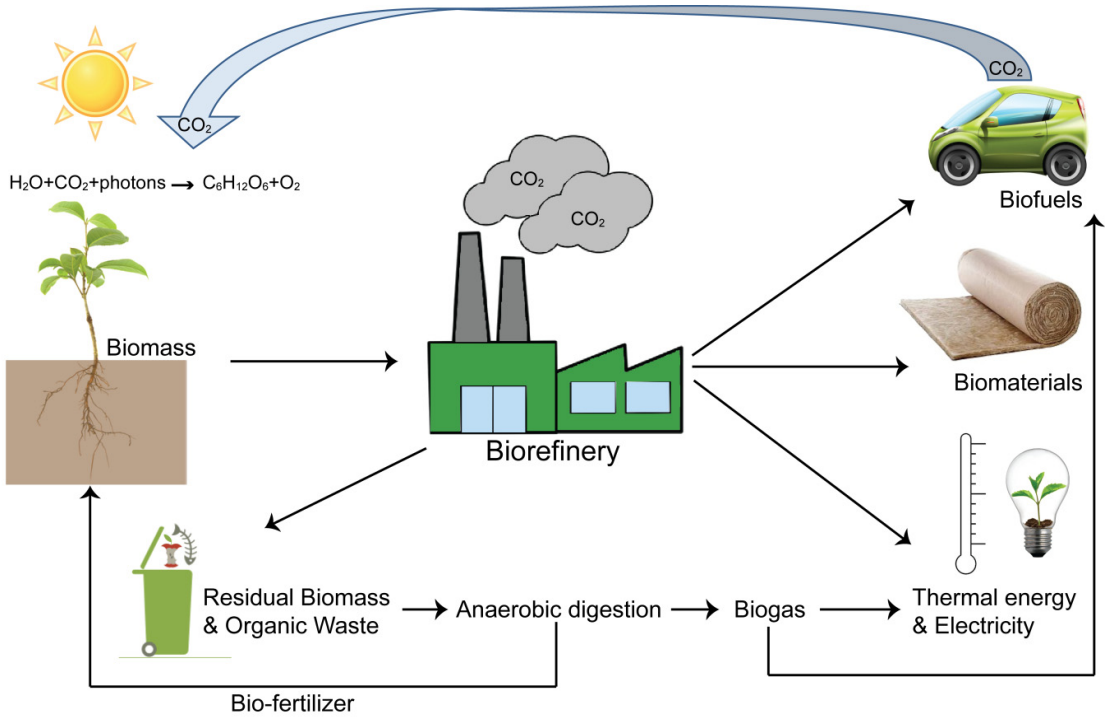


Figure 1. The concept of a biorefinery based on lignocellulosic feedstocks.

The future prospect for saccharification of lignocellulosic biomass to replace first generation biomass entails both technical and economic challenges that are mainly due to the complexity of the lignocellulosic substrate and the performance of the digestive enzymes used for its depolymerization (Himmel et al., 2007, Jørgensen et al., 2007, Banerjee et al., 2010, Chundawat et al., 2011, Cannella and Jørgensen, 2014). Therefore, efforts to improve the efficiency of biomass pretreatment and subsequent enzymatic hydrolysis methods are in focus. There are many ways to improve the enzyme cocktails, for instance by searching for new enzyme activities, engineering critical enzyme properties, or by optimizing enzyme compositions for different substrates. The goal of these efforts is to establish more cost-effective technologies for exploitation of lignocellulosic feedstocks as a renewable source of energy, chemicals and materials (Figure 1).

1.2 Carbohydrates

Carbohydrates (i.e. sugars) make up one of the four major classes of biomolecules, along with proteins, nucleic acids and lipids. Historically, the word carbohydrate derives from the fact that glucose, the first carbohydrate obtained in pure form, has the chemical formula

$C_6H_{12}O_6$ and it was therefore thought to be a "hydrate of carbon," $C_6(H_2O)_6$. This interpretation was later abandoned as carbohydrates are aldehydes or ketone compounds with hydroxyl groups, but the name persisted. Monosaccharides are the basic units of carbohydrates. The monosaccharide building blocks can be joined together to form disaccharides or longer oligo- or polysaccharides. The chemistry and structures of oligo- and polysaccharides are diverse as the formation of glycosidic bonds, which is catalyzed by enzymes called glycoside transferases, may yield a large variety of arrangements (various monosaccharides, various types of glycosidic linkages, modifications such as acetylations). Carbohydrates make up most of the organic matter on earth due to their extensive roles in all forms of life, with functions ranging from cell-cell communication to being structural elements in cells walls (e.g. cellulose and chitin, discussed below). Carbohydrates serve as nutrients for organisms from all kingdoms and are also used as energy storage (e.g. starch and glycogen) in many organisms.

1.2.1 The plant cell wall

The plant cell wall is a cell organelle that provides mechanical strength and is responsible for diverse functions such as regulating transport, providing protection, signaling, and storage of energy reserves (Brett and Waldron, 1990). The plant cell wall consists of a network of cross-linked polysaccharides (cellulose and hemicellulose) and lignin (Figure 2), as well as minor fractions of pectin, glycoproteins, lipids, minerals, soluble sugars and extractives. The plant cell wall is built up in layers: the middle lamella, the primary cell wall, the secondary cell wall and the warty layer (Sjöström, 1993). The primary cell wall contains an irregular network of cellulose, hemicellulose and pectin that is supported by the secondary cell wall, which is produced after the cell has stopped growing. The secondary cell wall is rich in organized cellulose microfibrils, with a higher degree of polymerization (DP) than the cellulose found in the primary cell wall (Brett, 2000). The secondary cell wall also contains hemicelluloses, which are heteropolymeric polysaccharides and which composition varies between plant types. Hemicelluloses are cross-linked to lignin, a high molecular weight phenolic heteropolymer that provides the cells with mechanical support and strength (Higuchi, 1990). The ratio between cellulose, hemicellulose and lignin in the cell wall differs between different plant species and tissues. For example hard- and softwoods are generally rich in cellulose and lignin, which favors

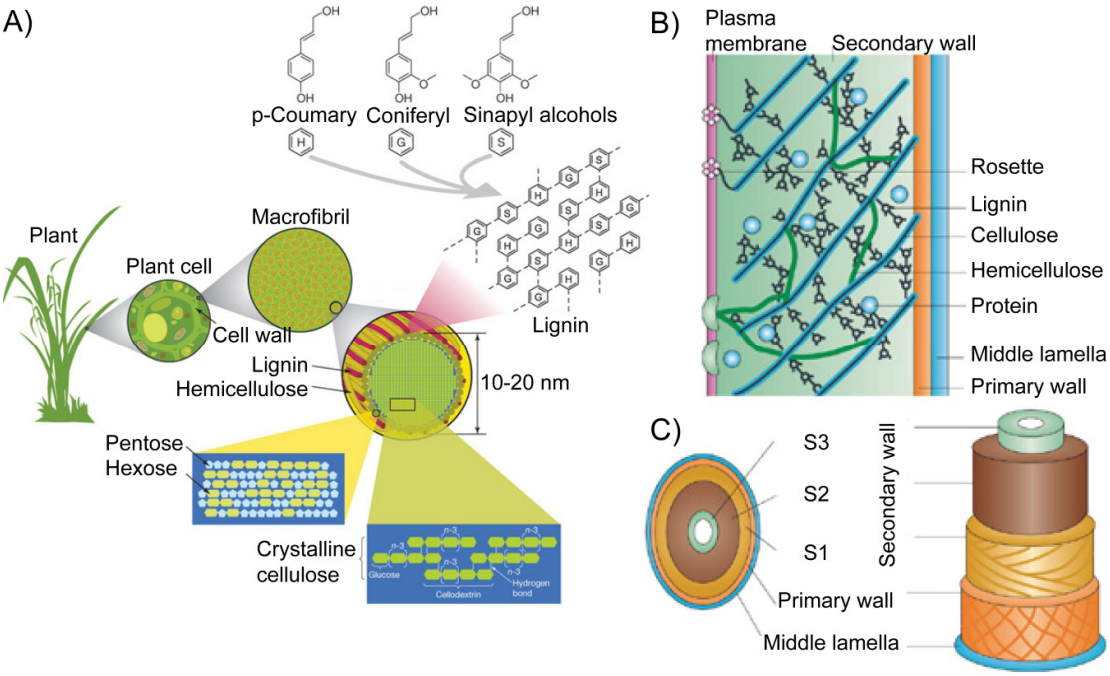


Figure 2. Structural organization of the plant cell wall. Panel A) visualizes the arrangement and composition of the main components, cellulose, hemicellulose and lignin. In panel B) the cellulose synthase rosette complexes are shown in the plasma membrane producing cellulose ending up in the secondary wall together with hemicellulose, lignin and proteins. Panel C) shows the different layers of the plant cell wall with focus on the secondary wall containing the more organized cellulose microfibrils. Picture A was taken from (Rubin, 2008) and pictures B and C from (Sticklen, 2008).

rigidity, whereas agricultural plants have a relatively high content of hemicelluloses, making their tissues more flexible than wood (Olsson et al., 2004).

1.2.1.1 Cellulose

Cellulose, the main component of plant cell walls (35-50%), is a linear non-branched homopolymer made up of D-glucopyranose units that are linked by β-1,4-glycosidic bonds (Figure 3). The glucose units in cellulose chains are rotated by 180° relative to each other, making the disaccharide, cellobiose, the repeating unit (Cocinero et al., 2009). In higher plants, cellulose is synthesized at the plasma membrane by large cellulose synthase complexes (Figure 2B). The complexes appear as six hexameric rosettes and presumably comprise 36 individual cellulose synthase proteins (Somerville, 2006, Mutwil et al., 2008). Shortly after biosynthesis, the cellulose chains aggregate into microfibrils that are stabilized

through inter- and intramolecular hydrogen bonds and van der Waals forces (Somerville et al., 2004, Parthasarathi et al., 2011). Due to the huge potential for hydrogen bond formation, cello-oligosaccharides longer than DP8 are insoluble as they have greater affinity for another glucan chain than for the aqueous solvent (Brown, 2004). Based on X-ray scattering data (Fernandes et al., 2011) and information about the cellulose synthase (Endler and Persson, 2011), elementary cellulose fibrils are thought to consist of 24 to 36 cellulose chains, where all the hydroxyl groups are positioned equatorially while the hydrogens are in axial positions. As a consequence of this arrangement, the elementary fibrils possess polar sides and a hydrophobic top and bottom (Beckham et al., 2011). Elementary fibrils associate into highly ordered structures called microfibrils. Microfibrils vary in crystallinity: the inner core of a fibril is highly organized (i.e. crystalline), whereas fibril chains located on or closer to the surface are sub-crystalline (Ding and Himmel, 2006). These latter regions show a lower degree of structural order mainly due to a lower number of intramolecular bonds and are sometimes referred to as amorphous. Altogether, the features of the fibril make cellulose highly resistant to chemical and enzymatic hydrolysis (Nishiyama et al., 2002, Himmel et al., 2007).

There are seven polymorphs of cellulose (Iα, Iβ, II, III_I, III_{II}, IV_I and IV_{II}) that differ in symmetry and chain geometry (O'Sullivan, 1997). Native cellulose is found in two crystalline forms, cellulose Iα and Iβ, in which the latter is the more thermodynamically stable. Cellulose Iβ is the most abundant form found in nature, since it is the main cellulose form in higher plants (Brown, 2004). Several Gram-positive and Gram-negative bacteria are capable of synthesizing cellulose used to form a protective envelope around their cells. The predominant polymorph isolated from bacteria is cellulose Iα (Ross et al., 1991). Cellulose Iα and Iβ have the same conformation, with a parallel chain arrangement but differ in hydrogen bonding patterns. Most pretreatment methods of cellulose retain the cellulose I polymorph, although certain treatments, e.g. with ionic liquids or liquid ammonia (or other amines), may lead to formation of cellulose II and cellulose III, respectively. Ionic liquids can fully solubilize cellulose and recrystallization often forms cellulose II which is arranged as antiparallel sheets (Swatloski et al., 2002). Pretreatment of cellulose I (or II) with ammonia forms cellulose III_I (or III_{II}) that has an expanded surface area with loss of microfibril orientation (Wada et al., 2004). Cellulose III has been shown to have significantly lower recalcitrance to enzyme deconstruction than native cellulose I

(Igarashi et al., 2007). Heating of cellulose III_I or III_{II} forms the cellulose polymorphs IV_I and IV_{II} (O'Sullivan, 1997).

1.2.1.2 Hemicellulose

Hemicelluloses are non-cellulose plant polysaccharides commonly with a β-1,4-linked backbone. Hemicelluloses include xyloglucans, xylans, mannans, glucomannans and mixed β-1,3(4)-linked glucans and their composition is highly variable between different plants and different plant cells (Scheller and Ulvskov, 2010). These soluble polysaccharides coat the cellulose microfibrils and consist of aldohexoses (glucose, galactose or mannose), pentoses (xylose, arabinose) and uronic acid building blocks. Some hemicellulose polymers are linear, whereas others are more complex, containing branches and chemical modifications of certain sugars. The hemicelluloses coat the cellulose microfibrils in the plant cell wall and may be covalently attached to lignin through ester bonds. Xylans and glucomannans constitute the two most abundant hemicelluloses, as xylans are the major hemicellulose in hardwoods and agricultural plants and glucomannans dominate in softwoods.

1.2.2 Chitin

After cellulose, the second most abundant polymer in nature is chitin, which appears as a structural component in yeast and fungal cell walls as well as in the exoskeletons of arthropods (e.g. crabs and shrimps) and insects. Chitin is similar to cellulose (Figure 3), consisting of linear β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) units that are rotated by 180° relative to each other, making *N,N'*-diacetylchitobiose the repeating unit (Gooday, 1990). There are two crystalline polymorphs of chitin found in nature and the most common form is α-chitin. In α-chitin the polysaccharide chains are arranged in an antiparallel fashion making this chitin form more dense and rigid than β-chitin, which has a more open structure due to a parallel chain arrangement (Gardner and Blackwell, 1975, Minke and Blackwell, 1978). Sufficient deacetylation [i.e. minimally approximately 35%; (Vårum et al., 1994)] of the acetamido groups in GlcNAc to form D-glucosamine (GlcN) leads to formation of a soluble polysaccharide that is referred to as chitosan. Chitosan and chitooligosaccharides are valuable biodegradable and biocompatible products with a

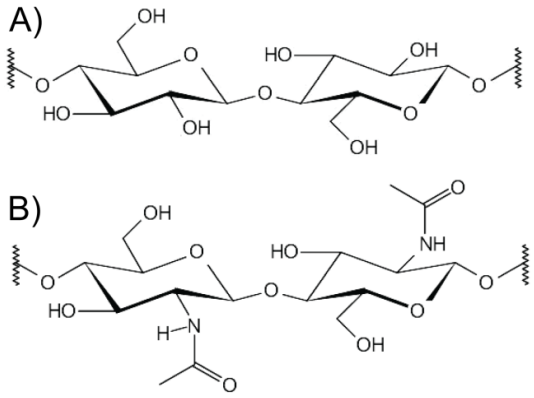


Figure 3. Repeating unit of A) cellulose and B) chitin.

number of applications in agriculture, cosmetics, wastewater treatment, and medicine (Aam et al., 2010).

1.3 Microbial degradation of structural polysaccharides

As described above, polysaccharides such as cellulose and chitin are insoluble and contain a high degree of structural order, which makes them resistant to microbial degradation. Nevertheless, microbes have evolved enzyme systems to deconstruct these polysaccharides, and such enzyme systems provide a good starting point for designing industrial enzyme cocktails for biomass saccharification. Cellulolytic microorganisms include bacteria, protozoa and fungi, and these microbes occupy a number of ecosystems that include soils, composts, swamps, rivers, lakes, seawater sediments, decaying vegetable matters and more (Ljungdahl and Eriksson, 1985). Cellulose-degrading bacteria can also be found in cellulolytic animals for example in the digestive tract of ruminants (Leschine, 1995) or in the guts of wood-degrading termites (Warnecke et al., 2007). Cellulose utilization by bacteria occurs in aerobic, anaerobic, mesophilic and thermophilic environments (Lynd et al., 2002).

Aerobic bacteria and fungi secrete their enzymes which then act in the surroundings of the microbe. These secreted free enzyme systems tend to contain a variety of enzyme types targeting cellulose, hemicellulose and in some cases also lignin. The wood-decay fungus *Trichoderma reesei* (also known as *Hypocrea jecorina*) has for long been used as a model organism for the study of free enzyme systems (Martinez et al., 2008).

Anaerobic bacteria that utilize lignocellulosic biomass as carbon source have evolved another enzyme system wherein the enzymes are organized into large extracellular multienzyme complexes called cellulosomes, first described for the thermophilic bacterium *Clostridium thermocellum* (Bayer et al., 1998). In cellulosomes various lignocellulolytic enzymes, each containing a dockerin domain, assemble on a scaffoldin protein containing multiple cohesin domains through cohesin-dockerin interactions. The scaffoldin may be anchored to the bacterial cell, but evidence for free extracellular cellulosomes exist (Fontes and Gilbert, 2010, Zhou et al., 2014). The enzymes found in these macromolecular assemblies contain multiple specificities and act in close proximity to each other, which is thought to be beneficial when degrading the complex polysaccharides of the plant cell wall (Bayer et al., 1998, Bayer et al., 2008, Fontes and Gilbert, 2010). Interestingly, it was recently shown that cellulosomes and free enzyme systems act in synergy in the degradation of cellulose and pretreated corn stover (Resch et al., 2013, Resch et al., 2014). A recent analysis of a complex microbial cellulolytic community indicates that this community exploits the combined action of cellulosome and free enzyme systems (Zhou et al., 2014).

In this thesis, aerobic free enzyme systems for cellulose deconstruction will be in focus. The following sections deal with the typical enzymes that are part of such systems.

1.4 Carbohydrate active enzymes

The term Carbohydrate Active enZymes (CAZymes) refers to enzymes that act on carbohydrate substrates, i.e. in the synthesis and breakdown of glycoconjugates, oligo- and polysaccharides. As discussed above (Section 1.2), the variety of monosaccharides combined with the variety of sugar linkages, and not to mention the fact that many molecules in biological systems can be glycosylated (e.g. glycoproteins), makes carbohydrate substrates among the most structurally diverse in nature. Thus, nature needs a large variety of enzymes acting on these substrates, which are involved in a wide range of biological processes (Cantarel et al., 2009). CAZymes usually have to perform their function with high specificity, which is a challenging task, considering the only very minor variations between some of the carbohydrates. As the number of known CAZymes has increased during the last decades, a database has been built up for such enzymes, called CAZy [www.cazy.org (Cantarel et al., 2009, Lombard et al., 2014)]. The CAZy classification system is primarily based on amino acid sequence similarities, which implies

that the classification correlates better with structural fold and catalytic mechanism than with enzyme specificity. Consequently, enzymes with the same specificity can be found in several families and vice versa.

The CAZy database current holds more than 300 protein families divided into six enzyme classes, including one class that comprise non-catalytic, associated modules (carbohydrate-binding modules, CBMs). The biggest and most well studied CAZyme class is represented by the glycoside hydrolases (GHs), which currently are divided into more than 130 families, and are responsible for hydrolysis of glycosidic bonds (Henrissat and Davies, 1997). Some GHs can also perform transglycosylation, meaning that another sugar molecule is used as acceptor instead of water (as in hydrolysis), leading to the formation of a new glycosidic bond. About 40% of the GH families contain enzymes that contribute to plant cell wall deconstruction (Gilbert, 2010).

The other classes of CAZymes are the glycosyl transferases (GTs) that are involved in the biosynthesis of glycosidic bonds from phospho-activated sugar donors, the polysaccharide lysases (PLs) that cleave glycosidic bonds in uronic acid containing polysaccharides using a β-elimination mechanism and the carbohydrate esterases (CEs) that remove ester-based modifications (Cantarel et al., 2009). The latest addition to CAZy is a rather diverse collection of redox enzymes that goes under the name "auxiliary activities" (AAs). Unlike the other enzyme classes, enzymes in the AA class do not exclusively act on carbohydrates. Enzymes acting on lignin (i.e. laccases, manganese peroxidases and lignin peroxidases) have been included since lignin is found in conjunction with polysaccharides in the plant cell wall (Levasseur et al., 2013). One major reason for this recent expansion of the CAZy database was the discovery of enzymes that today are called lytic polysaccharide monooxygenases (LPMOs) and that carry out oxidative cleavage of polysaccharides [(Vaaje-Kolstad et al., 2010, Horn et al., 2012b) and Paper I in this thesis]. These enzymes were previously classified as GH61 or CBM33 and are discussed in detail in section 1.5, below.

The final CAZy class comprises the carbohydrate-binding modules (CBMs) which are non-catalytic modules that are associated to other CAZymes. By binding to polysaccharides, CBMs bring the catalytic domains into close proximity to target substrates thus potentiating catalysis (Bolam et al., 1998, Boraston et al., 2004). CBMs display large variation in binding specificity to the extent that cellulose-binding CBMs, which occur in several families, seem optimized to bind different types and faces of cellulose (Blake et al.,

2006). Interestingly, several studies indicate discrepancies between the binding preferences of CBMs and the substrate specificity of the catalytic domains they are appended too (Hervé et al., 2010, Cuskin et al., 2012). The most likely explanation for these observations is referred to as the "proximity effect", i.e. by binding to one polysaccharide type in the plant cell wall, the CBM brings the catalytic domain in proximity of its substrate, which is another type of polysaccharide (Hervé et al., 2010). Notably, a recent study has shown that the beneficial effect of CBMs on enzyme efficiency diminishes at high substrate concentration, which is a finding with important practical applications for the biorefinery (Várnai et al., 2013).

1.4.1 Synergism in cellulose degradation

Some enzymes are known to perform more efficiently when acting together than alone, a phenomenon known as synergism. Synergism is said to occur when the action of two enzymes combined is higher than the sum of actions of the individual enzymes (Wood and Garcia-Campayo, 1990). The efficiency of secreted cellulolytic enzyme systems, found in fungi and bacteria, is thought to be in part due to the synergistic action of three different types of glycoside hydrolases (Merino and Cherry, 2007) (Figure 4). These enzymes are known as cellobiohydrolases, acting from the non-reducing end (EC 3.2.1.91) or the reducing end (EC 3.2.1.176), endoglucanases (EC 3.2.1.4), and β-glucosidases (EC 3.2.1.21). Two types of synergism are well-described in literature: synergy between endoglucanases and cellobiohydrolases, termed the endo-exo-synergism, and synergy between different cellobiohydrolases termed exo-exo-synergism (Henrissat et al., 1985, Nidetzky et al., 1993, Väljamäe et al., 1999).

Endoglucanases (EGs) are mostly non-processive enzymes and are thought to randomly hydrolyze polysaccharide chains, probably in the more amorphous regions of the cellulose polymer. Enzymes with EG activity are found in 15 of the GH families. Some are single domain enzymes, whereas some have additional CBMs. Endoglucanase activity results in generation of new chain ends, one reducing and one non-reducing end, thus generating an increased number of access points for the cellobiohydrolases (CBHs). CBHs are processive enzymes, meaning that the enzymes thread a single carbohydrate chain through their active sites and repetitively cleave off cellobiose using a stepwise sliding movement. CBHs attacking the substrate from the reducing end include the fungal CBHs of GH family 7 and

the bacterial CBHs of GH family 48. CBHs attacking the substrate from the non-reducing end include cellulases of GH family 6, occurring in both fungi and bacteria.

Processive enzymes remain closely associated to the substrate in between subsequent hydrolytic reactions, which is an energetically favorable trait especially for enzymes acting on crystalline surfaces (Teeri, 1997, Horn et al., 2006, Eijsink et al., 2008, Zakariassen et al., 2010). The basic idea behind this reasoning is that once the decrystallization penalty for extracting a polysaccharide chain from its insoluble and possibly crystalline context has been paid, it is favorable that this chain stays attached to the enzyme in between catalytic steps rather than (partly) re-associating with the insoluble material (Teeri, 1997, Horn et al., 2006, Beckham et al., 2011, Beckham et al., 2014). Detailed studies on processive chitinases (Horn et al., 2006, Zakariassen et al., 2009, Horn et al., 2012a, Vaaje-Kolstad et al., 2013) have shown that aromatic residues lining the substrate binding clefts are crucial for processivity. Importantly, by analyzing non-processive mutants, these studies have also shown that, while processivity indeed improves efficiency on crystalline substrates, the structural features necessary to achieve processivity slow down the enzymes when acting on soluble substrates (Horn et al., 2006).

The cellobiose released by the CBHs is the substrate for the third group of glycoside hydrolases involved in degradation of cellulose, namely the β-glucosidases (found in family GH1 and GH3). These enzymes convert solubilized cello-oligomers and cellobiose to monomeric glucose, mitigating CBH product inhibition arising from cellobiose (Henrissat et al., 1985, Wilson, 2009).

Because of the homopolymeric and unbranched nature of cellulose, only a few enzyme activities are needed for its depolymerization. The situation for the heteropolymeric hemicelluloses is quite different. Xylans and glucomannans are often branched and contain various substitutions and modifications (see section 1.2.1.2). Due to this complexity, depolymerziation of hemicelluloses requires a more diverse portfolio of CAZyme activities, making synergism in such systems more difficult to study (de Vries and Visser, 2001).

1.4.2 The C_1, C_x hypothesis and accessory proteins

In 1950, Reese *et al.* suggested that cellulolytic organisms may have two systems to convert cellulose into smaller sugars that then can diffuse into the cell. The hypothesis was based on their observations that many microorganisms possessing enzymes that are able to hydrolyze β-(1→4)-glycosidic bonds, showed a very limited ability to utilize native

crystalline cellulose. Based on this observation, it was suggested that the microbes that can utilize native cellulose have at least two enzyme systems, the C_1 system and the C_x system. The C_1 system was postulated to transform native crystalline cellulose into shorter and more accessible linear polysaccharides, which would then be accessible for degradation by the C_x system that is not restricted to cellulolytic microbes only. The C_x system comprises hydrolytic enzymes that cleave the β-(1→4)-glycosidic linkage, but at the time, the C_1 system remained unknown (Reese et al., 1950).

Since 1950, a limited number of non-hydrolytic proteins have been discovered that enhance cellulose degradation. One type of accessory protein is known as expansins, which occur in plants and are used to mechanically destabilize the plant cell wall during cell growth (Cosgrove, 2000). This type of protein is also found in certain bacteria and fungi, that probably acquired this by horizontal gene transfer (Nikolaidis et al., 2014). An expansin-like secreted protein from *Bacillus subtilis* was shown to bind plant cell walls, cellulose and peptidoglycan and promote plant cell wall extension similar to plant-expansins. Deletion of the encoding gene reduced the ability of the bacterium to colonize maize roots (Kerff et al., 2008). One fungal expansin homologue, called swollenin, has received particular attention. It comprises a CBM1 cellulose-binding domain and an expansin-like domain and has been shown to be involved in fiber disruption (Saloheimo et al., 2002). Another accessory protein was discovered in 2005, when Vaaje-Kolstad *et al.* showed that a chitin-binding protein from family 33 of CBMs greatly enhanced the efficiency of hydrolytic enzymes in the degradation of chitin (Vaaje-Kolstad et al., 2005a, Vaaje-Kolstad et al., 2005b). Five years later it was shown that this CBM33 protein in fact is an enzyme that uses an oxidative mechanism to cleave crystalline chitin (Vaaje-Kolstad et al., 2010). Today, these enzymes go under the name lytic polysaccharide monooxygenases (LPMOs) (Figure 4) and are recognized as being important novel tools for biotechnology and biorefining (Horn et al., 2012b).

1.5 Lytic polysaccharide monooxygenases (LPMOs)

In 2010 and 2011, it was discovered that proteins/enzymes belonging to family 33 of CBMs (CBM33) and family 61 of GHs (GH61) in fact are oxidative enzymes capable of cleaving chitin and cellulose chains using an unprecedented oxidative mechanism (Vaaje-Kolstad et al., 2010, Forsberg et al., 2011, Phillips et al., 2011, Quinlan et al., 2011). These enzymes are today classified in auxiliary activity families 10 and 9 (AA10 and AA9),

respectively, and are collectively referred to as lytic polysaccharide monooxygenases (LPMOs). LPMO are copper-enzymes and their activity depends on the presence of dioxygen and an external electron donor.

CBM33 proteins have been isolated from a number of chitinolytic bacteria and generally show high affinity for chitin substrates (Schnellmann et al., 1994, Kolbe et al., 1998, Suzuki et al., 1998, Schrempf et al., 1999, Folders et al., 2000, Chu et al., 2001). This affinity combined with the presence of conserved aromatic residues (Zeltins and Schrempf, 1997), which is a common feature of CBMs (Gilbert et al., 2002, Boraston et al., 2004), was the reason for classifying these proteins as CBMs. The first crystal structure of a CBM33 revealed that the majority of the conserved aromatic amino acids in fact are located internally and led to the suggestion that the binding surface is made up of mainly polar residues (Vaaje-Kolstad et al., 2005a). This first CBM33 structure was the structure of a 21 kDa Chitin-Binding Protein (CBP21) produced by the chitinolytic bacterium *Serratia marcescens.* It was originally thought to be a non-catalytic chitin binding protein, hence the name, and it represents one of the major proteins in the supernatant of *S. marcescens* when grown in the presence of chitin (Suzuki et al., 1998). As alluded to above, CBP21 was shown to enhance chitinase activity in the degradation of β-chitin as early as in 2005 (Vaaje-Kolstad et al., 2005b). Back then, this effect that was ascribed to CBP21 interfering with the structure of crystalline chitin, possibly causing partial decrystallization, which would increase substrate accessibility and, thus increased chitinase efficiency (Vaaje-Kolstad et al., 2005b, Eijsink et al., 2008).

In 1997, a *T. reesei* gene was cloned, expressed and shown to have weak endoglucanase activity on cellulose substrates (Saloheimo et al., 1997). The new enzyme laid the basis for the new family 61 of GHs (GH61). The GH61 enzyme (previous name EG-IV) is induced together with cellulases when *T. reesei* grows on cellulose, but its cellulase activity was found to be several orders of magnitude lower compared to the other cellulases. Due to the low activity on cellulose, a variety of plant cell wall polysaccharides was tested as potential substrates, but no activity was detected (Karlsson et al., 2001). Determination of the first structure of a GH61 revealed structural similarities to the CBP21 protein (Karkehabadi et al., 2008). In 2010, Harris *et al.* showed that a GH61 enzyme isolated from the fungus *Thielavia terrestris* enhanced cellulase activity, thus revealing clear functional similarity to CBP21 (Vaaje-Kolstad et al., 2005a, Vaaje-Kolstad et al., 2005b). One year after the

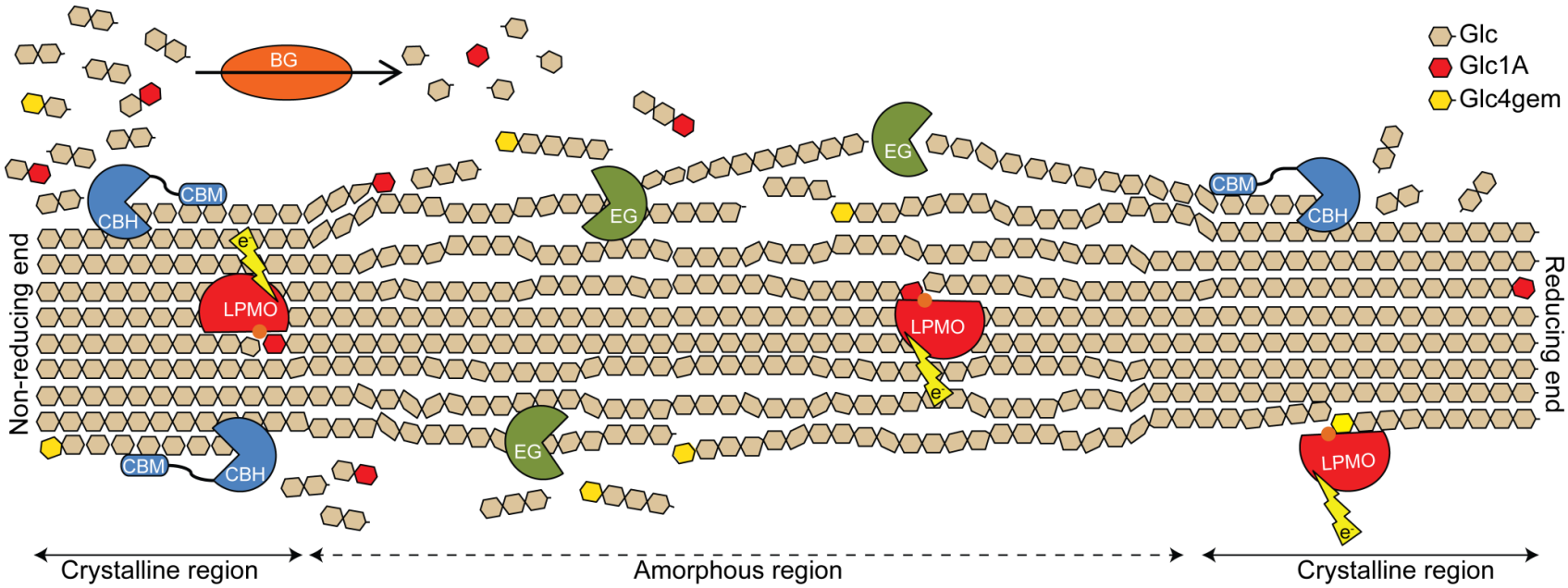


Figure 4. Cellulose degradation in aerobic microorganisms using the secreted free enzyme system. The enzymes include the lytic polysaccharide monooxygenases (LPMOs) shown in red, which can either oxidize C1 and/or C4, leading to a gluconic acid (Glc1A; red) or gemdiol (Glc4gem; yellow) chain ends. The LPMOs are copper dependent (orange dots) and their activity depends on electrons that can originate from a variety of sources (e.g. ascorbate, cellobiose dehydrogenase or lignin). Endoglucanases (green) hydrolyze the cellulose leading to formation of new chain ends, one reducing end and one non-reducing, which serve as substrates for the processive cellobiohydrolases (blue). Finally, β-glucosidases (orange) convert solubilized cello-oligomers and dimers to monomeric glucose. Cellobionic acid (C1-oxidized dimer) has been shown to be hydrolyzed to glucose and gluconic acid by β-glucosidases with 10-fold reduced efficiency compared to cellobiose (Cannella et al., 2012). There are currently no available data on the activity of beta-glucosidases on C4-oxidized cello-oligomers, but since the enzymes cleave from the non-reducing end, β-glucosidase activity on these compounds is not likely.

discovery that CBP21 as an oxidative enzyme (Vaaje-Kolstad et al., 2010), papers describing oxidative cleave of cellulose were published [Paper I of this thesis and (Phillips et al., 2011, Quinlan et al., 2011, Beeson et al., 2012)].

Today, it is widely accepted that the proteins previously classified as GH61 or CBM33 and today classified as AA9 or AA10 are enzymes that carry out oxidative cleavage of glycosidic bonds in crystalline substrates such as cellulose and chitin (Horn et al., 2012b, Hemsworth et al., 2013b).

1.5.1 Occurrence of LPMOs

LPMOs are found in families 9-11 of auxiliary activities in the CAZy database (Levasseur et al., 2013). Family AA9 consists of fungal proteins only and its members were previously classified as family 61 glycoside hydrolases (GH61). Family AA10 proteins are found in all domains of life (Figure 5), but only bacterial members have been characterized so far. As mentioned above, AA10-type LPMOs were originally thought to be non-catalytic carbohydrate binding proteins and were therefore classified as family 33 carbohydrate-binding modules (CBM33). The last LPMO family is AA11, which was very recently described for the first time (Hemsworth et al., 2014). Currently, this family includes fungal sequences only. Notably, as discussed below, all LPMOs have similar structures, while sequence diversity is huge.

The genomes of biomass degrading fungi encode a plethora of LPMOs (up to 40 genes in one fungus) several of which are upregulated when grown on biomass substrates (Berka et al., 2011, Eastwood et al., 2011, Hori et al., 2011, Adav et al., 2012, Yakovlev et al., 2012, Poidevin et al., 2014). The genomes of LPMO-containing bacteria usually encode only one or two LPMOs, but some species, such as the soil bacterium *Streptomyces coelicolor* A3(2), contain up to seven LPMO genes (Bentley et al., 2002). Some bacterial LPMOs have been shown to be co-regulated with cellulases (Garda et al., 1997, Ramachandran et al., 2000), and some are known to be secreted upon growth on various types of biomass such as chitin (Suzuki et al., 1998) and cellulose (Adav et al., 2010, Takasuka et al., 2013, Book et al., 2014).

LPMOs often seem related to biomass degradation, but family AA10 proteins, which are the main focus of this study, also occur in viral genomes, as well as in the genomes of pathogenic bacterial species whose roles in biomass conversion are not obvious. As an example, the pathogen *Vibrio cholera*, which causes diarrhea in humans, possesses a

GlcNAc-binding protein A (GbpA), which has an N-terminal LPMO10 domain. It has been shown that this protein is a colonization factor for bacterial attachment to both marine chitin and mammalian intestinal mucin, a *N*-acetyl-glucosamine contain polymer (Wong et al., 2012). Knock-out studies have shown that bacterial virulence is significantly reduced in the absence of GbpA (Kirn et al., 2005). Interestingly, several studies suggest that AA10 proteins may be involved in bacterial virulence, for example in *Listeria monocytogenes*, which causes severe food-borne infections, as well as in opportunistic pathogens such as *Pseudomonas aeruginosa* (Tran et al., 2011) and *Enterococcus faecalis* (Vaaje-Kolstad et al., 2012). Also probiotic bacteria such as *Lactococcus lactis* ssp. *lactis* and *Lactobacillus plantarum* have AA10 proteins, which bind *N*-acetyl-glucosamine containing polymers (Vaaje-Kolstad et al., 2009, Sanchez et al., 2011), but whose biological function is unclear.

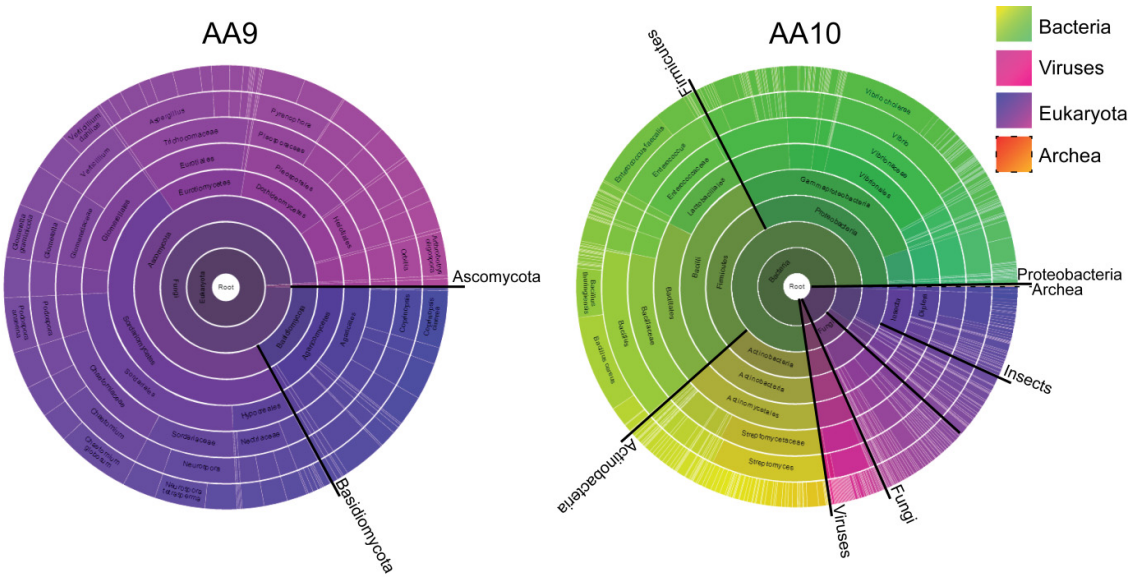


Figure 5. **Species distribution of proteins belonging to families AA9 and AA10.** The picture is constructed from information found in the Pfam database (AA9s, Pfam ID: PF03443; AA10s, Pfam ID: PF03067). The AA9 family contains fungal protein domains only, manly from the Ascomycota phylum, but also from the Basidiomycota phylum. Family AA10 on the other hand contains protein domains from all four kingdoms, with an abundance of bacterial proteins from the Proteobacteria, Firmicutes and Actinobacteria. The total number of entries is 827 for AA9 and 1556 for AA10.

1.5.2 LPMO substrates

The first LPMO to be characterized was CBP21 from *S. marcescens* (*Sm*LPMO10A), which was shown to be active on crystalline β-chitin and to a lesser extent on α-chitin (Vaaje-Kolstad et al., 2010). In another study of the same LPMO (Nakagawa et al., 2013), it was shown that the synergy between CBP21 and the *S. marcescens* chitinases declines as the crystallinity index of the substrate was reduced, indicating that this LPMO is active on crystalline materials. Activity on cellulose has been described for both LPMO9s (previously GH61s) and LPMO10s [see Paper I of this thesis and (Quinlan et al., 2011)]. Very recently, it has been shown that an LPMO9 from *Neurospora crassa* (*Nc*LPMO9C) is active on both crystalline cellulose and soluble cello-oligosaccharides (Isaksen et al., 2014). Importantly, it was also shown that this enzyme is active on hemicelluloses with a β-(1→4)-glucan backbone, such as xyloglucan, and that it accepts various substitutions on the glucan backbone (Agger et al., 2014). These latter discoveries indicate that LPMO substrates may be far more diverse than first anticipated.

According to the Pfam database (Finn et al., 2010), proteins from the LPMO containing AA families appear as single domains as well as in combination with CBMs or other domains. AA10s (Pfam ID: PF03067) most commonly occur as single domain proteins (Figure 6) but are also found with CBMs from family 5/12 (chitin binding) or family 2 (cellulose binding). Some fungal AA10s contain a CBM20, which is known to bind starch substrates (Boraston et al., 2004, Christiansen et al., 2009). Several other combinations occur (Figure 6), including combinations with glycoside hydrolases from family 18 and 19 (chitinases) and with a GH5 mannanase (Sunna et al., 2000). The fungal AA9s (Pfam ID: PF03443) show a less diverse architecture. Approximately 75% occur as single domain proteins, whereas approximately 20% have a family 1 cellulose binding CBM. The remaining AA9s have additional modules with indistinct and uncertain annotations.

The new family AA11 LPMO from *Aspergillus oryzae* was discovered by "module walking". A module with unknown function found in AA9s, called X278, was used to search for other proteins with the same module. A large number of secreted proteins, other than AA9, were found, suggesting that some could be CAZymes. Some of these had an N-terminal histidine that is characteristic for LPMOs. A truncated version of the *Ao*AA11 protein (without the X278 domain) was then expressed and shown to possess LPMO

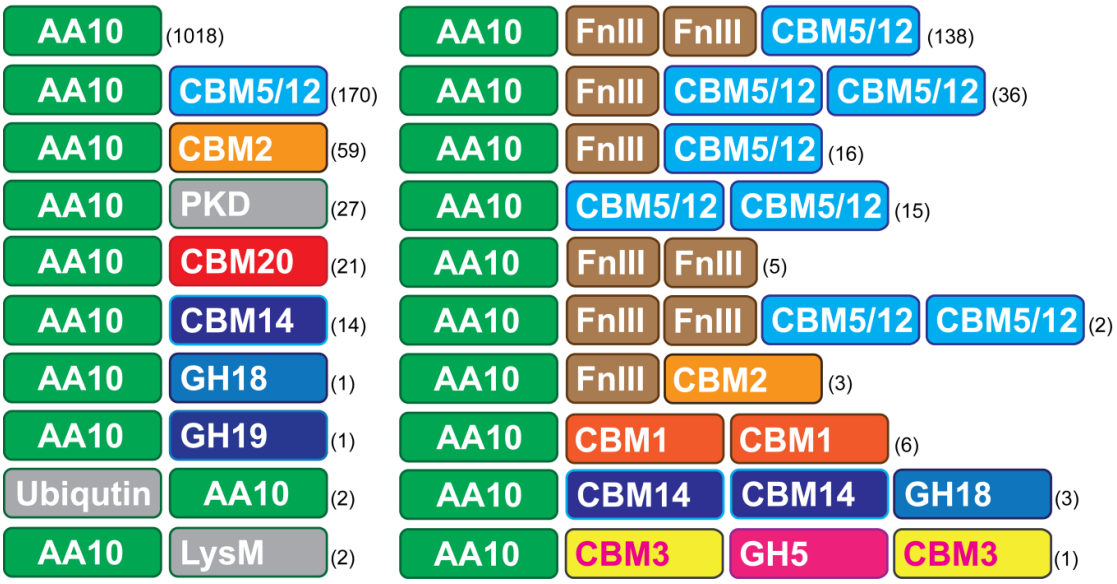


Figure 6. Domain structure of naturally occurring AA10 proteins. Twenty of the 34 different AA10 domain architectures, described in the Pfam database [(PF03067) http://pfam.sanger.ac.uk/]. The blue colored domains are related to chitin binding or degradation whereas orange colored domains are related to cellulose binding. The total number of entries is given in parenthesis.

activity towards chitin (Hemsworth et al., 2014). At present, there are about 50 AA11 sequences in the CAZy database.

1.5.3 Cellulose oxidation by LPMOs

The first LPMO to show activity towards cellulose substrates was the *S. coelicolor* enzyme CelS2, which produces C1-oxidized products (see Paper I of this thesis). The firstly described LPMO9 active on cellulosic substrates [*Ta*LPMO9A, (Quinlan et al., 2011)] generated a more complex product profile than CelS2, suggesting that this LPMO9 enzyme could oxidize the C1 carbon as well as the C4 or C6 carbon. Shortly after, Phillips *et al.* (2011) described the activity of three of the 14 LPMO9s in the fungus *Neurospora crassa,* all with a different oxidation pattern. One was shown to produce C1 oxidized products only, one was shown to oxidize C4 only, and the third showed a mixed activity similar to *Ta*LPMO9A described by Quinlan *et al* (2011). Phillips *et al.* (2011) rejected the possibility of C6 oxidation as this will not lead to cleavage of the glycosidic bond, in contrast to C1

and C4 oxidation. Thus, C4 oxidation, which results in formation of ketoaldoses, was suggested as the second major oxidative activity of LPMOs (next to C1 oxidation). Evidence for C4 oxidation, has since then been provided in several studies (Beeson et al., 2012, Isaksen et al., 2014, Vu et al., 2014).

Phylogenetic analysis of AA9s shows four major clusters (Figure 7) and three of these correlate to the regioselectivity of substrate oxidation (Li et al., 2012, Vu et al., 2014). The four groups constitute PMO-1 (PMO for polysaccharide monooxygenase) that oxidizes the C1-carbon, PMO-2 that oxidizes the C4-carbon, and PMO-3 that can oxidize both the C1- and C4-carbon. The fourth group has been named PMO-3* as the sequences are similar to PMO-3s, but activities are equal those found in the PMO-1 cluster (Vu et al., 2014).

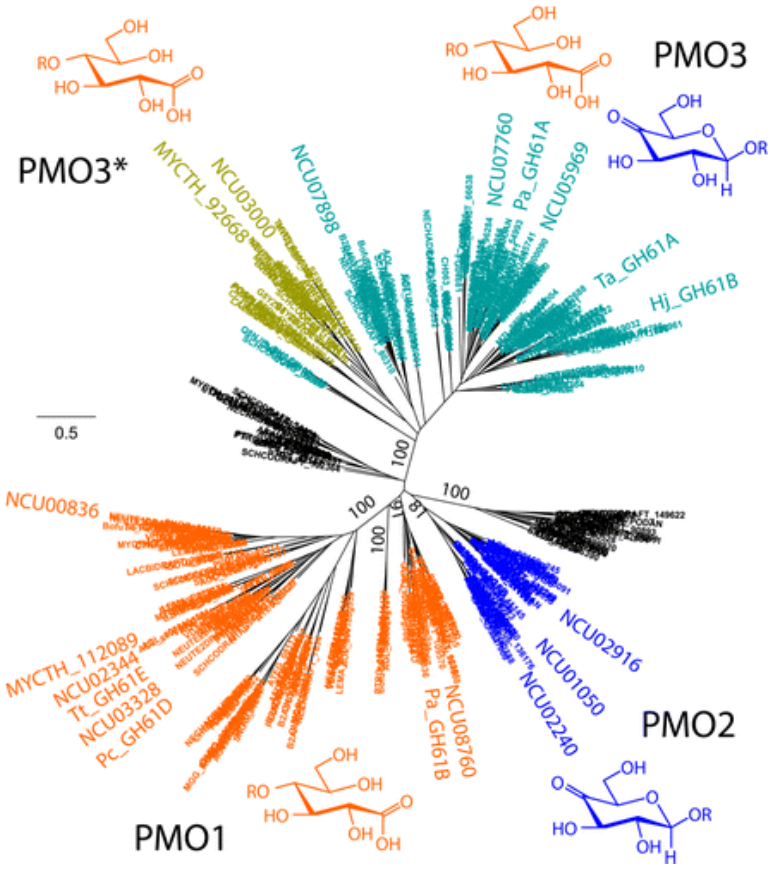


Figure 7. Phylogenetic tree of 497 AA9 sequences. Respective oxidation products are shown next to the LPMO group. Black colored groups are unclassified, due to lack of characterized representatives. The magnified entities represent enzymes for which regioselective hydroxylation has been demonstrated or for which structural data are available. This picture was taken from (Vu et al., 2014).

Figure 8 shows the various products formed by the three types of LPMOs. Native oligosaccharides are released by all LPMOs as a result of oxidation close to a chain end. C1 oxidation initially leads to the formation of δ-1,5-lactones, which at physiological pH will be hydrolyzed to aldonic acids (Vaaje-Kolstad et al., 2010). Oxidation of the non-reducing end (C4-oxidation) results in formation of 4-ketoaldoses. The hydrated forms of 4-ketoaldoses are referred to as gemdiols. LPMOs with a mixed C1/C4-oxidizing activity can also produce "double oxidized" products (i.e. an oligosaccharide oxidized at both the reducing and the non-reducing end), which are formed when a polysaccharide chain is cleaved twice, once with C1 oxidation and once with C4 oxidation.

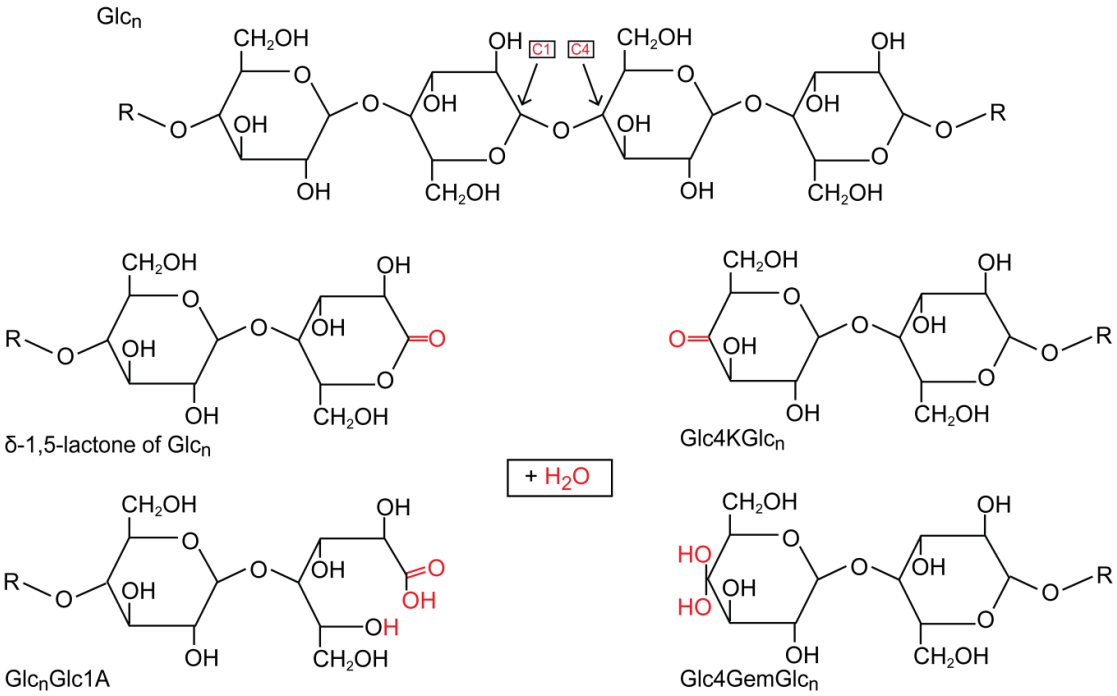


Figure 8. Oxidized products formed upon C1 (left) or C4 (right) oxidation of a cellulose chain. C1 oxidation results in the formation of a lactone that spontaneously will become hydrolyzed to the more stable aldonic acid form. C4 oxidation leads to formation of a ketoaldose and that is in equilibrium with its gemdiol form.

1.5.4 Three-dimensional structures of LPMOs

At the time of writing, April 2014, there existed six 3D-structures of LPMO9s, eight structures of LPMO10s and one incomplete structure of an AA11-type LPMO in the Protein Data Bank database (http://www.pdb.org/), including two AA10 structures that are presented and discussed in Paper III of this thesis. The enzymes in the three different LPMO families share a similar structural fold (Figure 9), but have low sequence similarity (> 10%). The structure of CBP21 was the first LPMO structure to be determined (Vaaje-Kolstad et al., 2005a). The structure of CBP21 was also later determined by NMR spectroscopy, which confirmed the crystal structure and revealed that this LPMO is a compact and rigid molecule (Aachmann et al., 2012). In 2008, Karkehabadi et al. determined the first structure of an AA9-type of LPMO, namely *Hj*LPMO9B (GH61B), from the filamentous fungi *Hypocrea jecorina*. LPMOs have small and globular immunoglobulin-like structures with antiparallel β-stands in a distorted β-sandwich fold. The strands are connected by loops with a varying number of α-helix insertions [Figure 9 (Karkehabadi et al., 2008, Harris et al., 2010, Quinlan et al., 2011, Li et al., 2012, Vaaje-Kolstad et al., 2012, Wong et al., 2012, Wu et al., 2013, Hemsworth et al., 2013a, Hemsworth et al., 2014) and Paper III of this thesis]. Most of the variable structural elements in LPMOs are found in a loop referred to as loop 2 (L2), which differs in size and conformation and which may play a role in determining substrate specificity and perhaps even the regioselectivitiy of substrate hydroxylation [(Li et al., 2012, Wu et al., 2013, Vu et al., 2014) and Paper III of this thesis]. In LPMO9s that have mixed C1/C4 activity, the L2 loop is extended which leads to an increased surface area (shown in yellow, Figure 9). Outside the L2 loop, other differences occur that may explain functional variation. For example, the structure of C4-oxidizing *Nc*LPMO9D (a PMO-2), shows an additional α-helix (see dotted box in Figure 9) that is conserved in LPMOs with predicted C4-oxidation activity (Vu et al., 2014).

LPMOs possess flat binding surfaces (Vaaje-Kolstad et al., 2010, Aachmann et al., 2012) that harbors a type-2 copper center (Quinlan et al., 2011, Hemsworth et al., 2013a). Aromatic residues are commonly involved in enzyme-carbohydrate interactions (Boraston et al., 2004) and the binding surface of LPMO9s indeed possess two to three conserved tyrosine residues with rings parallel to the binding surface (Figure 9) that suggest a role in substrate binding (Li et al., 2012, Wu et al., 2013, Hemsworth et al., 2013b). The binding surface in LPMO10s is dominated by polar residues and only one aromatic residue

(Tyr/Trp) is presented in an orientation that could be parallel to the substrate. Mutation of this residue (Tyr54) in CBP21 to alanine resulted in reduced substrate binding (Vaaje-Kolstad et al., 2005a). The substrate binding surface of CBP21 has been studied in detail by

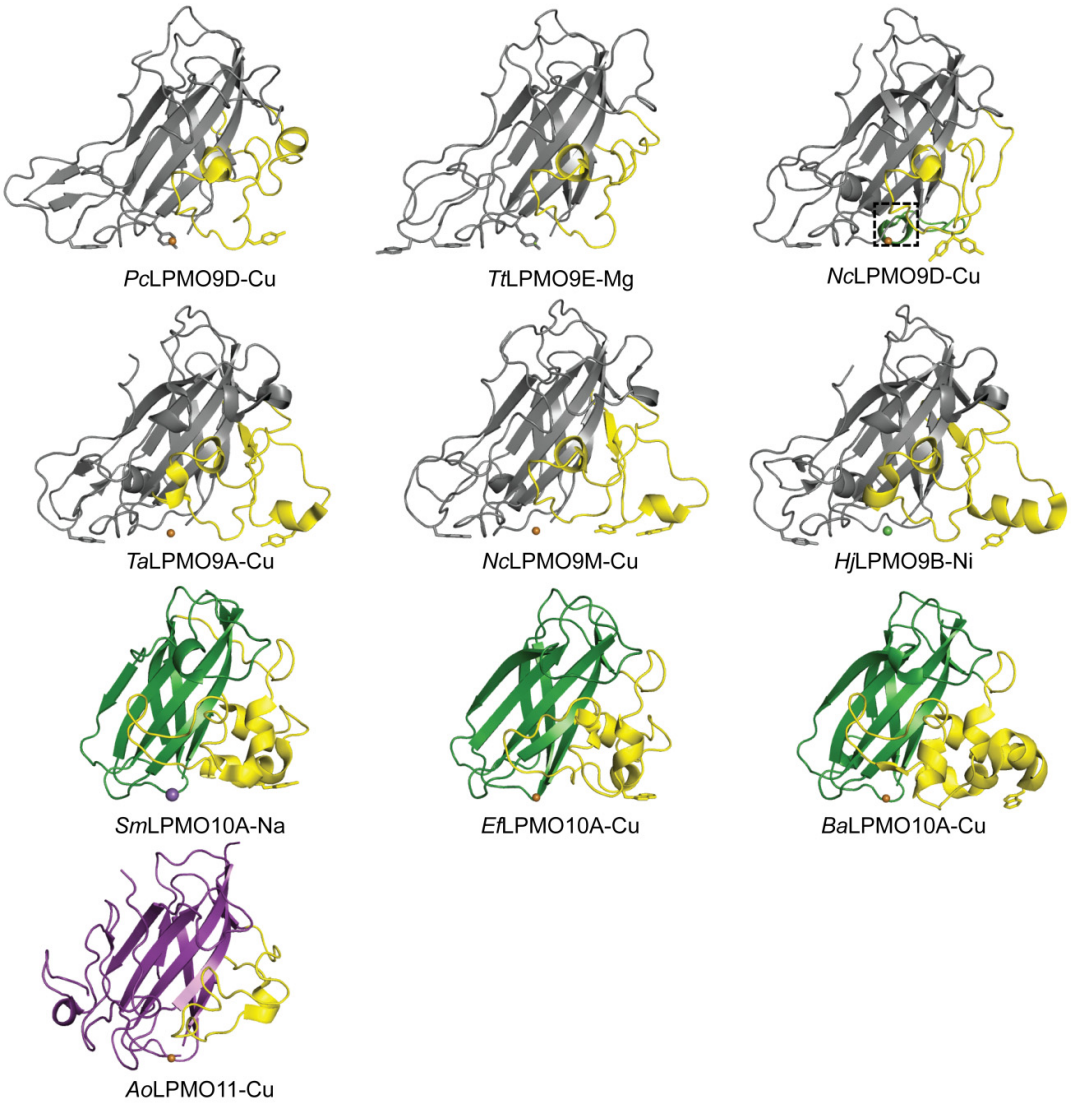


Figure 9. Cartoon representation of six cellulose active LPMO9s (grey), three chitin active LPMO10s (green) and one chitin active LPMO11 (purple). The box in *Nc*LPMO9D shows an additional helix that, as suggested by sequence alignments, is conserved in LPMO9s with predicted C4-oxidizing activity (PMO-2). The L2 loop is colored yellow and tends to be extended in the LPMO9s on the second row which are classified as C1/C4-oxidizing LPMO9s (PMO-3s). Residues with aromatic rings lying parallel to the putative substrate-binding surface are shown as sticks. Various metal ions (Cu, Mg, Ni, and Na) refined in the catalytic centers are shown as spheres. Figures were made with PyMOL (DeLano and Lam, 2005).

NMR, which revealed that a major part of the surface residues is involved in substrate binding (Aachmann et al., 2012).

1.5.5 The copper active site

The first LPMO structures revealed a solvent exposed metal-binding site formed by the δ-nitrogen of the N-terminal histidine, its main chain amino group, and the ε-nitrogen of another highly conserved histidine (Vaaje-Kolstad et al., 2005a, Karkehabadi et al., 2008) (Figure 10). Later, studies have shown that this metal binding site is a type 2 copper site that is essential for the LPMO activity (Quinlan et al., 2011, Westereng et al., 2011, Vaaje-Kolstad et al., 2012, Aachmann et al., 2012). The structure and binding affinity of the copper site has been studied by X-ray crystallography, electron paramagnetic resonance (EPR) spectroscopy, nuclear magnetic resonance (NMR) spectroscopy and isothermal titration calorimetry (ITC). Crystal structures have been obtained for apo-LPMOs (free from metal ligands) as well as LPMOs binding a variety of divalent metal ions (Ni^{2+}, Zn^{2+}, Mg^{2+}).

Early work studying the effects of CBM33 and GH61 proteins on chitin and cellulose degradation (Vaaje-Kolstad et al., 2005a, Harris et al., 2010, Vaaje-Kolstad et al., 2010) left an impression that LPMOs could employ various metal ions, but today it is clear that LPMOs are copper-enzymes. In 2011, Quinlan *et al.* showed by various methods that *Thermoascus aurantiacus* LPMO9A (*Ta*LPMO9A) is a Cu-dependent enzyme, thus providing the first solid evidence for the role of this transition metal in LPMO activity. The structure of *Ta*LPMO9A was determined with a copper ion in the active site (Figure 9 and Figure 10), and the ligands were shown to include the two histidines described above that form a T-shaped geometry referred to as the 'histidine brace' (Quinlan et al., 2011). Another finding was that the N-terminal histidine was methylated on the ε-nitrogen, a post transcriptional modification with unknown function that later has been identified in all fungal LPMO structures where the proteins have been produced in their native hosts (Quinlan et al., 2011, Li et al., 2012). The overall copper coordination in LPMO9s displays an octahedral geometry with Jahn-Teller distortion (i.e. with six ligands). In addition to the histidine brace ligands, the LPMO9s have a conserved glutamine residue (Gln173 in *Ta*LPMO9A; Figure 10) that coordinates a water molecule that represents the fourth equatorial ligand. The axial positions are occupied by a highly conserved tyrosine (most likely a tyrosinate; Tyr175 in *Ta*LPMO9A), and a free water molecule on the solvent-facing

side. In the two *N. crassa* structures determined by Li *et al.* (2012), density possibly representing dioxygen species was found in the solvent-facing axial region, albeit with a distance that is significantly longer (~ 2.9 Å) then than expected for a copper-oxygen bond [~ 2.0 Å (Hemsworth et al., 2013b)].

The T-shaped histidine brace is the key structural element conserved in all LPMO active sites (Figure 10). The axial positions differ between the LPMO families. In LPMO10s, these positions are usually occupied by hydrophobic residues: a phenylalanine (Phe187 in CBP21 vs Tyr175 in *Ta*LPMO9A) on the protein side and an alanine (Ala112 in CBP21) on the solvent side. The alanine is conserved in LPMO10s and has been proposed to restrict axial access to the copper site in LPMO10s (Hemsworth et al., 2013b). LPMO10s, which have all been expressed heterogeneously, show no electron density reflecting His methylation and neither do LPMO9s from heterogeneous expression in *Pichia pastoris* (Wu et al., 2013) that are still active (Westereng et al., 2011, Kittl et al., 2012). The LPMO11 structure (Figure 9 and Figure 10) shows features from both LPMO9s and LPMO10s. The LPMO11 has the alanine in the solvent-facing axial region, while it contains a Tyr in the other axial position, albeit quite far from the copper (Tyr140 in *Ao*LPMO11 with a Cu-O distance of 3.2 Å).

In the study by Quinlan *et al.* (2011) an EPR spectrum for *Ta*LPMO9A-Cu^{2+} was presented showing the characteristics of a type 2 copper site according to the Peisach-Blumberg classification of type 1 and type 2 copper proteins (Peisach and Blumberg, 1974). In 2013, Hemsworth *et al.* (2013a) showed an EPR spectrum for *Bacillus amyloliquefaciens* LPMO10A (*Ba*LPMO10A), a protein that has been reported to have high affinity for chitin substrates (Chu et al., 2001), but for which no catalytic activity has been reported so far. The EPR spectrum showed that *Ba*LPMO10A falls between the typical type 1 and 2 copper proteins although a copper type 2 classification was suggested to be appropriate for LPMO10s based on the overall axial envelope visible in the crystal structure (Hemsworth et al., 2013a). EPR analysis of *Ao*LPMO11, active on β-chitin (Hemsworth et al., 2014), showed a spectrum similar to that of cellulose active *Ta*LPMO9A. A comparative EPR study of LPMO10s is presented in Paper II of this thesis.

Isothermal titration calorimetry has been used to measure the dissociation constant (K_d) for LPMO-Cu^{2+}. In the study of *Ta*LPMO9A, the K_d could not be determined due to too

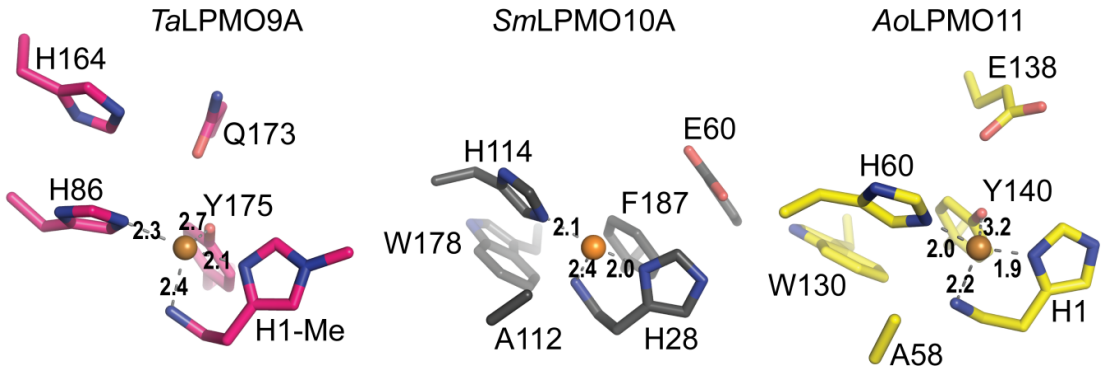


Figure 10. LPMO active sites. The pictures are stick representations of the active sites and include residues within 6 Å from the copper (golden spheres) of LPMOs. Magenta, *Ta*LPMO9A, an LPMO9 oxidizing C1 and C4 in cellulose; grey, *Sm*LPMO10A (CBP21), which oxidizes C1 on chitin substrates; yellow, *Ao*LPMO11, which also oxidizes C1 in chitin substrates. Figures were made with PyMOL (DeLano and Lam, 2005).

high affinity and was therefore suggested to be lower than 1 nM (Quinlan et al., 2011). For CBP21 the K_d was estimated to be 55 nM at pH 6.5 and by determining the redox potential for the LPMO-Cu^{2+} + e^- → LPMO-Cu^+ redox couple, a theoretical K_d for CBP21-Cu^+ of 1 nM was calculated (Aachmann et al., 2012). In the latter study, the fact that reduced copper binds stronger to CBP21 than oxidized copper was demonstrated by several methods. ITC was also used by Hemsworth *et al.* (2013a) to determine the Cu^{2+} dissociation constant for *Ba*LPMO10A and *Ao*LPMO11. The K_d was estimated to range from 80 nM at pH 8.0 to 6 nM at pH 5.0 for *Ba*LPMO10A and to be 0.8 nM at pH 5.0 for *Ao*LPMO11 (Hemsworth et al., 2013a, Hemsworth et al., 2014).

1.5.6 Reaction mechanism

In 2010, when it became evident that LPMOs are enzymes that cleave glycosidic bonds in an oxidative manner, it was shown by oxygen isotope labelling ($H_2{}^{18}O$ and $^{18}O_2$) that CBP21 (*Sm*LPMO10A) introduces one oxygen atom from molecular oxygen into the product (Figure 11). It was also shown that the reaction was boosted by the presence of an external electron donor such as ascorbic acid or reduced glutathione (Vaaje-Kolstad et al., 2010).

Interestingly, it has been shown that the electrons needed for LPMO action can be supplied by an enzyme known as cellobiose dehydrogenase (CDH; EC 1.1.99.18), which is found in some fungal secretomes (Langston et al., 2011, Phillips et al., 2011). CDHs are

oxidoreductases found in family AA3 and AA8 (Levasseur et al., 2013) that catalyze the oxidation of cellobiose or longer cellodextrins to δ-1,5-lactones that then are spontaneously hydrolyzed in solution to form aldonic acids. Even though the function of CDHs has been known for a long time, these enzymes have no clearly defined biological roles. In a knockout study, it was shown that deletion of the major *cdh* gene in the filamentous ascomycete *N. crassa* caused substantially reduced overall cellulose-degrading activity, an effect that could be ascribed to the lack of electron input to LPMOs (Phillips et al., 2011). CDHs have a heme and a flavin domain and some CDHs have an additional cellulose binding module belonging to family 1 CBMs. Cellobiose oxidation leads to reduction of the flavin group that subsequently transfers electrons to the heme domain. The reduced heme can further reduce a variety of compounds, including copper in the active site of LPMO9s. There is ample data in the literature showing that fungal LPMO9s may use both small molecule electron donors and CDH as electron donor (Langston et al., 2011, Phillips et al., 2011, Beeson et al., 2012, Kittl et al., 2012). Whether CDH is the natural electron donor for LPMOs remains uncertain as CDH is not found in all LPMO-containing fungal genomes and neither in bacterial genomes (Phillips et al., 2011).

In line with initial observations by Harris *et al.* (2010), it has been proposed in several studies that lignin may supply the electrons and stimulate LPMO activity (Cannella et al., 2012, Dimarogona et al., 2012, Hu et al., 2014). In the very recent study by Hu el al., it was shown that the activity of an enzyme cocktail supplemented with an LPMO9 on substrates with residual lignin content did not depend on the addition of gallate, a small molecule reducing-agent. Addition of gallate was only necessary to boost the hydrolysis of "pure" cellulosic substrates. In the same study, a boosting effect of hydrolysis was seen when hemicelluloses were present, and, more specifically, when xylan was supplemented to the reaction.

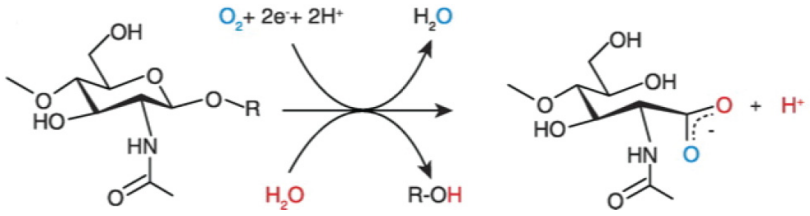


Figure 11. Chitin cleavage and oxidation catalyzed by CBP21. In the final oxidized product, an aldonic acid, one oxygen atom has been incorporated from molecular oxygen (blue) and one from water (red). The figure was taken from (Vaaje-Kolstad et al., 2010).

How the electrons are delivered to the catalytic center of LPMOs is unclear, but a cluster of conserved aromatic residues in proximity to the metal binding site has received attention in this respect, as it could potentially be involved in electron tunneling (Vaaje-Kolstad et al., 2005a, Li et al., 2012, Hemsworth et al., 2013a).

Since the discovery of LPMOs, putative reaction mechanisms for LPMO9s have been described (Phillips et al., 2011, Beeson et al., 2012, Li et al., 2012). The basic scenario, described for the first time by the Marletta group, is that the reduced copper activates molecular oxygen, which then abstracts a hydrogen atom from either C1 or C4. This is followed by hydroxylation of the same carbon, followed by a spontaneous elimination reaction leading to scission of the glycosidic bond (Phillips et al., 2011). In a recent study by Kim *et al.* (2014), quantum mechanical calculations were used in an attempt to identify the reactive oxygen species used by LPMOs as well as the key steps in the oxidative mechanism. These researchers used density functional theory calculations on an active site model based on the *Ta*LPMO9A structure [a C1/C4-active LPMO9 acting on cellulose substrates (Quinlan et al., 2011)] to compare two hypothesized reaction pathways for hydrogen abstraction and subsequent polysaccharide hydroxylation. One of the pathways involves the formation of a Cu(II)-superoxo intermediate, as originally proposed by the Marletta group (Phillips et al., 2011). The other tested mechanism, which came out as the most energetically preferred, involves formation of a Cu(II)-oxyl radical which has a stronger oxidative character than a Cu(II)-superoxo intermediate. The proposed "copper-oxyl, oxygen rebound" mechanism (Kim et al., 2014) is depicted in Figure 12. The study by Kim *et al* (2014) also addressed the effect of the N-terminal methylation found in LPMO9s with the outcome that this modification had minor effects on the reaction cycle (Kim et al., 2014).

The mechanism suggested by Kim *et al.* (2014) implies oxygen activation from the axial position. In LPMO10 structures, access to the solvent-facing axial position seems to be restricted by a highly conserved alanine residue (Hemsworth et al., 2013b). Thus, it is possible that oxygen activation differs between LPMO families. Additional data addressing this issue is presented in paper III in this thesis.

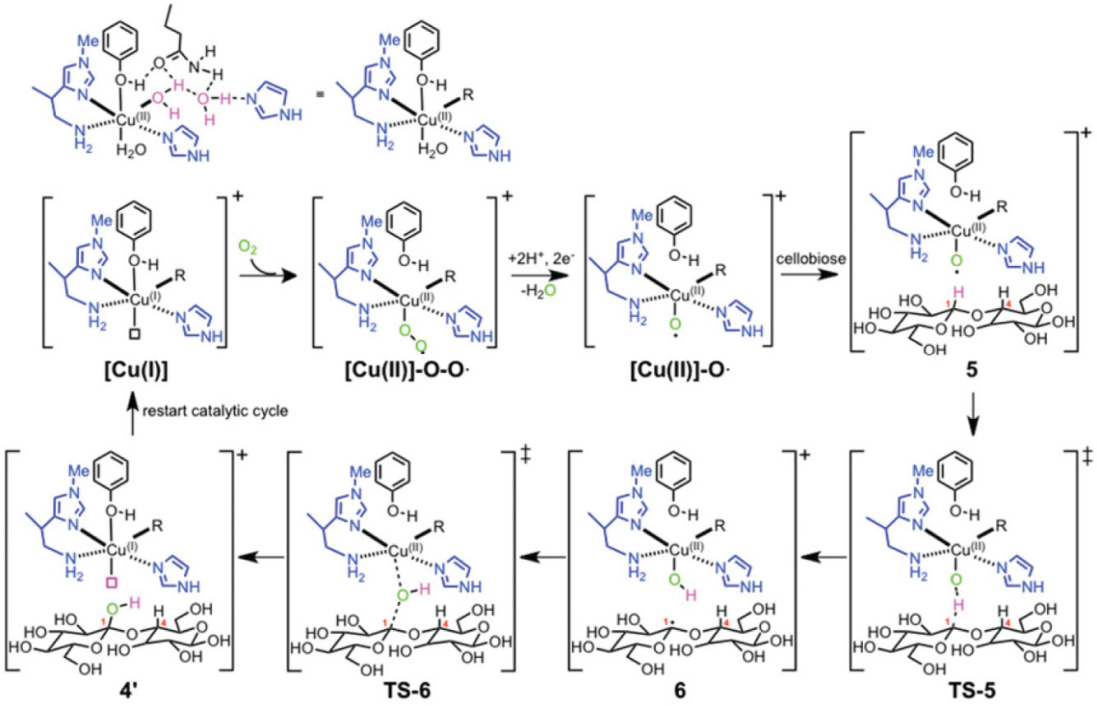


Figure 12. Proposed copper-oxyl mechanism for C1-hydroxylation by *Ta*LPMO9A based on quantum mechanical calculations. In brief, dioxygen binds end-on to copper to form a copper-superoxo intermediate, Cu(II)-O-O·. Next, the intermediate is reduced by a two electron transfer to a copper-oxyl radical, Cu(II)-O·. The radical subsequently abstracts hydrogen from the substrate, and the resulting [Cu(II)]-OH complex transfers the OH group to the sugar radical via an oxygen-rebound mechanism, resulting in hydroxylation of the substrate and reformation to the reduced copper state. The picture was taken from (Kim et al., 2014).

2 OUTLINE OF THE THESIS

During the last four years, LPMOs have become enzymes of great interest, which have profoundly changed our understanding of the enzymatic conversion of recalcitrant polysaccharides. At the time this work was initiated, in spring 2010, it was not known that proteins within the former family 33 of carbohydrate-binding modules (today family 10 of the auxiliary activities) were oxidative enzymes. This knowledge was first gained in October 2010 for an enzyme active on chitin. Since enzymes responsible for chitin degradation often have counterparts related to cellulose degradation, the first goal of the project was to identify a cellulose-targeting LPMO. This was not only of fundamental scientific interest, since new cellulose-active enzymes that may boost the efficiency of "classical" cellulases are of major importance for the growing biorefining industry, where lignocellulosic biomass is a major feedstock for production of biofuels and chemicals.

Due to the presence of multiple LPMO encoding genes, including genes encoding two-domain LPMOs containing cellulose binding modules, *Streptomyces* species were identified as a starting point for the search for cellulose-active AA10-type LPMOs. A protein referred to as CelS2 from *S. coelicolor* was targeted based on the observation that it contained a cellulose-binding module and that its expression was co-regulated with a cellulase. The subsequent study led to the identification of the first cellulose-targeting AA10-type LPMO which, at the time, in fact was the first cellulose targeting LPMO ever described. This work is described in the first paper of the thesis. CelS2 was shown to generate C1-oxidized cello-oligosaccharides from Avicel (a cellulose substrate) in the presence of dioxygen and an external electron donor. The activity of CelS2 was also shown to enhance the performance on Celluclast®, a commercial enzyme cocktail, in the degradation of filter paper.

As discussed in the Introduction of the thesis, LPMO enzymes have surface-exposed copper-containing active sites that cleave crystalline substrates in the presence of dioxygen and an external electron donor. Although reactants and some products of LPMO catalysis have been identified, not much is known about the reaction mechanism and enzyme properties governing substrate specificity and catalysis. In order to obtain a deeper understanding of the fundamental properties governing LPMO mechanism, research

following the work described in Paper I was devoted to structural and functional characterization of bacterial (AA10-type) LPMOs. Paper II describes the activity of two novel bacterial LPMOs, one active on cellulose and one active on chitin. This paper also describes a comparative study of the two novel and two well-known LPMOs, CBP21 and CelS2, combining EPR spectroscopy with in-depth sequence analysis. The EPR analysis showed that the electronic environment of the copper in the active site varies within the AA10 family and that this variation is coupled to substrate specificity.

Paper III describes the crystal structures of two cellulose targeting LPMO10s from *S. coelicolor*, CelS2 (*Sc*LPMO10C) and *Sc*LPMO10B. These are the first cellulose-oxidizing LPMO10 structures ever to be determined, permitting, for the first time, structural comparisons of LPMOs within one family with different substrate specificities (chitin- and cellulose-active LPMO10s) and of LPMOs in different families with similar substrate specificities (cellulose-active LPMO9s and LPMO10s). This pair of LPMOs is conserved in several cellulolytic actinomycetes. The data presented in Paper III show that the two enzymes from *S. coelicolor* and two homologous enzymes from *T. fusca* oxidize the substrate differently and act synergistically when degrading cellulose. Paper III also addresses possible differences in active site architectures (for LPMO9s and LPMO10s) that could determine the regioselectivity of the oxidation.

3 MAIN RESULTS AND DISCUSSION

3.1 Discovery of cellulose-active LPMO10s (Paper I)

The prelude to the work presented in this thesis was the discovery of a new enzyme activity that boosted the enzymatic solubilization of chitin (Vaaje-Kolstad et al., 2005a, Vaaje-Kolstad et al., 2010). Most enzyme activities related to chitin degradation also exist in the enzymatic machineries that target cellulose. Thus the primary objective of research in the initial phase of the thesis work was to identify a bacterial LPMO that could cleave the glycosidic bonds of cellulose chains. At the planning stage of the first study (Paper I), several leads indicating the existence of a cellulose-active LPMO were identified through studying the existing literature. The existence of cellulose targeting "stimulating proteins" was suggested by work on the fungal "GH61" proteins (now classified as family AA9 LPMOs) from *T. reesei* (Karlsson et al., 2001, Karkehabadi et al., 2008, Harris et al., 2010). In addition, at the time, it was known that genomes of cellulolytic bacteria found in the CAZy database encoded putative LPMO10s (Bentley et al., 2002, Moser et al., 2008) and that some bacterial LPMOs were co-regulated with cellulases and upregulated upon growth on cellulose substrates (Garda et al., 1997, Ramachandran et al., 2000, Adav et al., 2010). These studies and observations collectively indicated that cellulose oxidizing LPMO10s homologous to CBP21 existed in nature and motivated the study presented in Paper I.

The bacterial genome selected for screening for putative cellulose-active LPMOs was that of *Streptomyces coelicolor* A3(2). This actinomycete is widely represented in the ecological niche of soil (Hodgson, 2000). *S. coelicolor* A3(2) is genetically the most well-studied species in the *Streptomyces* genus, and its genome encodes more than 220 CAZymes whereof 13 chitinases and eight cellulases/β-glucosidases (Hopwood, 1999, Bentley et al., 2002). According to the CAZy database, the *S. coelicolor* genome contains genes encoding seven putative secreted LPMOs, which is among the highest number of LPMOs found in a bacterium. One of these genes contains a cell-wall anchoring motif (Walter and Schrempf, 2008). Three of the other six LPMOs contain an additional CBM, two having a CBM5/12 chitin binding module and one having a CBM2 cellulose binding module. The latter LPMO, CelS2, is a 34.5 kDa protein with a N-terminal LPMO domain and a well-defined C-terminal CBM2a cellulose binding domain (Simpson et al., 2000). In the genome, it is clustered and co-regulated with a GH12 β-(1→4)-endoglucanase (Garda et

al., 1997, Ramachandran et al., 2000). Thus, at the time, CelS2 was an obvious candidate for cloning, expression and characterization.

CelS2 was produced with a native N-terminus (i.e. the catalytically crucial N-terminal histidine) and degradation experiments showed that CelS2 produced oxidized cello-oligosaccharides in reactions containing Avicel or filter paper (Whatman #1) and an external electron donor (ascorbate or reduced glutathione). A lower quantity of soluble sugars was generated from reactions with filter paper compared to Avicel. This difference is likely due to a difference in the degree of polymerization of the substrate [2000 vs. 200, respectively (Zhang and Lynd, 2005)]. Since cello-oligosaccharides have limited solubility in water and must be below a certain DP (~DP8) before dissociating from the cellulose crystal, more bond cleavages must take place on a high DP substrate compared to a low DP substrate to reach the same level of accumulated soluble products (Figure 13).

Two new chain ends are generated when CelS2-catalyzed hydroxylation occurs at C1, one native non-reducing end and an "oxidized reducing end". Since it makes no sense to refer to an "oxidized reducing end", the terms "downstream end" (the reducing side of the saccharide chain) and "upstream end" (the non-reducing end of the saccharide chain) were introduced by Westereng *et al.* (Westereng et al., 2013). Oxidation of the C1 carbon on the downstream end of a cello-oligosaccharide yields an aldonic acid (gluconic acid), which is in a pH dependent equilibrium with the corresponding δ-1,5-lactone (δ-glucono-lactone). Indeed Figure 1 in Paper I, showing oxidized products detected by MALDI-TOF MS, contains minor peaks that correspond to the masses of the δ-1,5-lactone forms of various oligomeric products, but is dominated by signals representing aldonic acids.

Notably, product analysis itself was a major achievement of the work described in Paper I, as no chromatographic protocols or standards existed for the analysis of cello-oligosaccharides with a gluconic acid moiety at the downstream end. Although only lightly touched upon in the manuscript text, the method development was an essential constituent of the study that enabled verification of the identity of the oxidized products. This work included the generation of standards of cello-oligosaccharides containing a downstream end gluconic acid, which was produced using a two-step protocol involving production of cello-oligosaccharides by chemical hydrolysis (Wing and Freer, 1984) followed by a mild oxidation using iodine as an oxidant (Kobayashi et al., 1985, Kobayashi et al., 1996). The standards produced were verified by ^{13}C NMR spectroscopy and thereafter used to identify

CelS2-generated products by high-performance ion-exchange chromatography (see Figure S1 of Paper I).

Oxidized products generated by CelS2 showed a DP ranging from 2-7 with a dominance of even numbered cello-oligosaccharides. A similar product profile has been observed for the chitin active CBP21 (Vaaje-Kolstad et al., 2010). Vaaje-Kolstad *et al.* (2010) hypothesized that the dominance of even numbered products may indicate that the hydroxylation occurs on a solvent exposed polysaccharide chain that is embedded in a crystalline structure (Figure 13). The repeating unit of the cellulose chain is cellobiose since the monosaccharides are rotated by 180° relative to each other (Figure 3). Thus, if productive binding of CelS2 (or the LPMO in general) is dictated by the molecular topology of the cellulose crystal surface, only positions located at an even number of sugar monomers from an initial point of cleavage in the same chain will be accessible for productive binding by the LPMO. Interestingly, recent experiments by the author on a variety of LPMOs have shown that the product profiles of LPMOs change with the concentration of reductant used in the reaction; the products become shorter, with less dominance of even-numbered oligomers in reactions with higher reductant concentration (unpublished observations). It is conceivable that the increasing reducing power in the reaction may activate the LPMO to act on soluble substrates. In favor of this speculation, it should be noted that an LPMO9 was recently shown to cleave soluble oligosaccharides (Agger et al., 2014, Isaksen et al., 2014).

Prior to and after the discovery of LPMOs in the former CBM33 family (now AA10), it has been repeatedly shown that these proteins act in synergy with hydrolytic enzymes (Vaaje-Kolstad et al., 2005b, Moser et al., 2008, Vaaje-Kolstad et al., 2009, Vaaje-Kolstad et al., 2012, Kostylev and Wilson, 2013, Nakagawa et al., 2013). LPMOs release few soluble products themselves due their low catalytic rates (Vaaje-Kolstad et al., 2010, Agger et al., 2014). While LPMOs by themselves can cause some degree of substrate depolymerization, their key function is believed to be the disruption of crystalline regions of the substrate, making these regions more accessible to glycoside hydrolases (Vaaje-Kolstad et al., 2005b, Horn et al., 2012b). Although evidence supporting this hypothesis is limited, it can be envisioned that the negative charge (aldonic acid) introduced on the substrate surface by LPMOs can have a disruptive effect on the crystal, as has been observed for chemically oxidized cellulose (Bondeson et al., 2006).

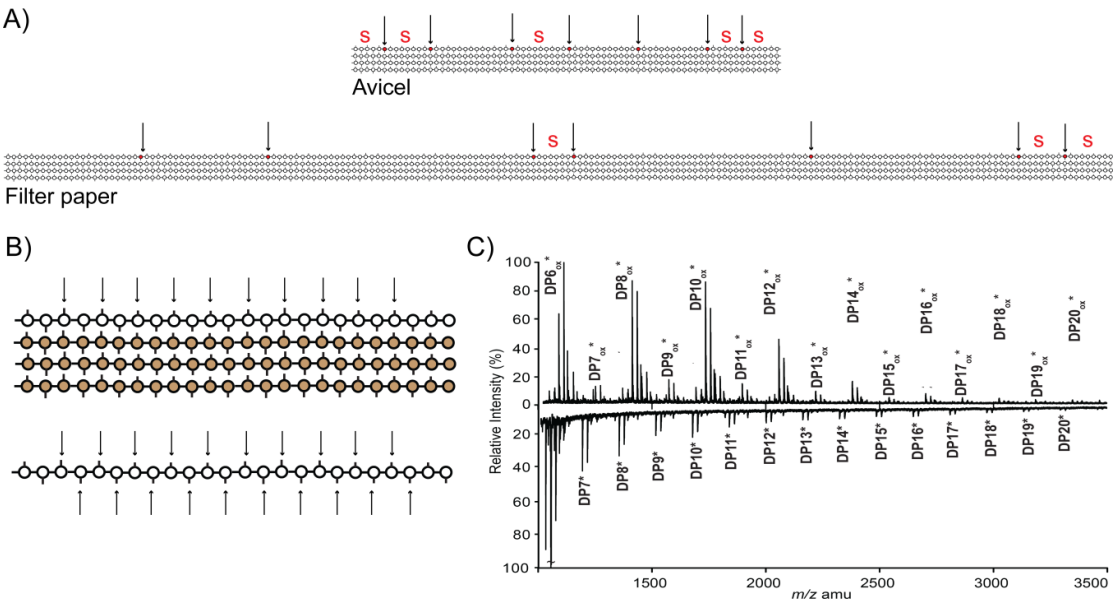


Figure 13. Schematic illustration of LPMO action on crystalline polysaccharides. Panel A) illustrates how the same number of hydroxylations causes more soluble (S) products to be released from substrates with lower DP (Avicel) compared to substrates with higher DP (filter paper). Panel B) illustrates the periodicity of even numbered cello/ chito-oligosaccharides released from a crystalline substrate region. The white sugar moieties in the cross-section represent surface exposed areas whereas the brown are embedded and not solvent exposed. The arrows indicate where productive binding can occur in a crystalline context (left) and in a soluble context (right). Panel C) shows MALDI-TOF MS spectra of soluble products released from β-chitin by combined treatment with a chitin-deacetylase (to increase product solubility) and CBP21 (a chitin-active LPMO; top) or chitinase C (an endochitinase; bottom). Only CBP1 generates a dominance of even-number products, suggesting that it acts on polysaccharide chains in a crystalline context, in contrast to chitinase C. Panel C) was taken from (Vaaje-Kolstad et al., 2010).

By analogy to previous findings for chitin-active LPMO10s, it was expected that CelS2 would be able to act in synergy with cellulases. This hypothesis was explored by incubating CelS2 with cellulolytic enzymes from the commercial enzyme cocktail Celluclast® (Novozymes) using high molecular weight filter paper as substrate, in the presence or absence of reduced glutathione (GSH) as electron donor. The activity of CelS2 did indeed enhance the hydrolytic enzymes, both in presence and absence of GSH. As expected, the highest product yield was obtained with reducing agent present in the reaction mixture, but the effect of adding the reductant was surprisingly small compared to what had been observed for chitin-active LPMO10s. When CelS2 was incubated with filter paper in the

absence of the cellulases, no detectable native products could be observed. The fact that CelS2 had an effect in the absence of reducing agent may be explained by the presence of compounds in the Celluclast mixture being able to provide electrons to the LPMO reaction. Synergistic affects were also observed with single hydrolytic enzymes (Figure 14), namely with a GH7 cellobiohydrolase from *Trichoderma reesei* (*Tr*Cel7A) and a GH5 β-(1→4)-endoglucanase from *Thermobifida fusca* (*Tf*Cel5A) (Vaaje-Kolstad et al., 2011). The hydrolytic enzymes also seemed to stay active over a longer time in the presence of CelS2, a result that might be due to the LPMO removing "obstacles" that otherwise would lead to unproductive cellulase binding (Igarashi et al., 2011, Kurasin and Väljamäe, 2011).

At the time of studying the synergy between CelS2 and various cellulases, not much was known about the optimal conditions for LPMO activity, not to mention the fact that the catalytic copper ligand had not yet been identified. With that information in mind, it is perhaps not surprising that high concentrations of CelS2, relative to the concentration of Celluclast (Figure 2 in Paper I), was needed to obtain a detectable boosting effect of CelS2. In early work, before the catalytic power of LPMOs had been discovered, Moser *et al.* (2008) showed that it was necessary to use low concentrations of the hydrolytic enzymes in order to detect effects of the *T. fusca* LPMOs (E7 and E8, which are further described in Paper III), and that the hydrolytic enzyme with the lowest activity (*Tf*Cel48A cellobiohydrolase) benefited the most from the presence of LPMOs (Moser et al., 2008).

The synergy observed between cellulose active LPMOs [from family AA9 and AA10; (Moser et al., 2008, Harris et al., 2010) and Paper I] and cellulases is much weaker than the LPMO-hydrolase synergies observed for chitinolytic enzyme systems (Vaaje-Kolstad et al., 2010, Vaaje-Kolstad et al., 2012, Nakagawa et al., 2013, Vaaje-Kolstad et al., 2013). Even though cellulose and chitin are similar, differences in the crystalline packing, which is less rigid and dense in the most LPMO-susceptible chitin form, β-chitin, could be one of the reasons for the stronger effects of chitin-active LPMOs. One might speculate that differences in the crystalline structure could affect enzyme-substrate interactions as well as, perhaps, oxygen binding.

In the particular case of CelS2 and the experiments described in Paper I, it must be noted that the nature of the metal ion essential for catalysis by LPMOs was not yet known and that it was assumed that the enzymes were relatively promiscuous in their metal preferences. Based on the few existing papers published on LPMOs (Karkehabadi et al.,

2008, Harris et al., 2010, Vaaje-Kolstad et al., 2010) Mg^{2+} was supplemented to the LPMO reactions, as this metal gave optimal activity for CBP21 (Vaaje-Kolstad et al., 2010). Shortly after Paper I was published, Quinlan *et al.* (2011) and Phillips *et al.* (2011) convincingly demonstrated that LPMOs from family AA9 (previously GH61) were copper-dependent enzymes. Soon after the copper-dependency of bacterial LPMOs was also demonstrated (Vaaje-Kolstad et al., 2011, Vaaje-Kolstad et al., 2012, Aachmann et al., 2012). LPMOs turned out to have very high affinities for copper, which enables LPMOs to scavenge metals that are found in the substrate (Quinlan et al., 2011, Aachmann et al., 2012), perhaps even more so in the presence of added bivalent metal ions that can compete for "metal-binding sites" in the substrate. This explains why CelS2 activity could be observed at all, even though it was not supplemented with copper prior to the experiments. Still, the putative lack of copper in the reactions may have caused an underestimation of the potential synergy of the system. On the other hand all LPMO-cellulase synergies reported so far in the literature (Langston et al., 2011, Westereng et al., 2011, Kostylev and Wilson, 2013, Hu et al., 2014) are modest compared to the effects seen in chitin degradation.

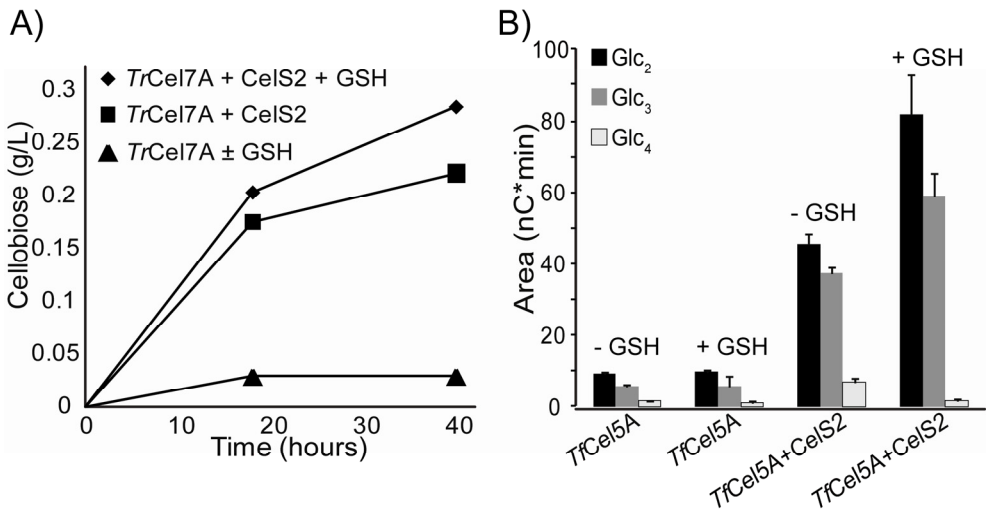


Figure 14. Synergism between CelS2 and cellulases. Synergy between CelS2 and a cellobiohydrolase (*Tr*Cel7A; A) or an endoglucanase (*Tf*Cel5A; B) in the degradation of filter paper. The reactions were set up with 5 µg/mL cellulase and 40 µg/mL LPMO in the presence or absence of 1 mM reduced glutathione (GSH) as electron donor in 20 mM sodium acetate buffer pH 5.5. CelS2 alone produced no detectable amounts of native cello-oligosaccharides and the curve for this reaction is not included in the figures. The figure is taken from (Vaaje-Kolstad et al., 2011).

The synergy between LPMOs and glycoside hydrolases is still poorly understood, one reason being that the substrates are variable and not fully understood. Even with one substrate type, large variations may occur. A recent study of an AA9-type LPMO from *T. aurantiacus* showed that specific substrate properties affected the level of synergy obtained when combining glycoside hydrolase and LPMOs in a degradation reaction. The most critical parameter determined was the relative amount of accessible crystalline cellulose. The same study showed that synergy occurred in reactions with cellulose I, but not in reactions with the more amorphous polymorphs, cellulose II and III, where the substrate is already easily available for the GHs (Hu et al., 2014). Likewise, it has been shown that the synergy between CBP21 and chitinases is reduced as the chitin substrate crystallinity decreases (Nakagawa et al., 2013). Importantly, it should be kept in mind that almost all studies are done with 'model' substrates, such as filter paper in the CelS2 study, which are (almost) pure cellulose substrates that have been modified and may not reflect how the enzymes would behave in the presence of natural substrates. Notably, it has been shown that CelS2 also enhances Celluclast® in the degradation of steam exploded saw dust from poplar (Vaaje-Kolstad et al., 2011).

3.2 Comparative functional and structural studies of LPMOs (Papers II & III)

In paper II, the activity of two new LPMOs is presented, *Bl*LPMO10A from *Bacillus licheniformis* that has C1 activity on chitin substrates, similar to CBP21, and *Tf*LPMO10B (E8) from *Thermobifida fusca* that has C1 activity on cellulose substrates, similar to CelS2. Paper II further presents a comparative study of two chitin-active (CBP21 & *Bl*LPMO10A) and two cellulose-active (CelS2 and E8) C1 oxidizing LPMO10s with the aim of finding correlations between substrate specificity, primary sequence, active site architectures and copper coordination as reflected in EPR spectra. Such correlations could be used as predictive tools for classifying LPMOs and also to provide new insight into LPMO mechanism.

In order to guide the interpretation of LPMO substrate preferences and EPR signatures, a structure-based multiple sequence alignment was constructed. All LPMO10 structures determined were used to set up a structure-based alignment using PyMOD and PyMOL (DeLano and Lam, 2005, Bramucci et al., 2012). At the time, no structural data were available for cellulose active LPMO10s nor for *Bl*LPMO10A, so the sequences for CelS2,

E8 and *Bl*LPMO10A had to be added to the structural alignment without the structural knowledge of these enzymes. The sequence identity between LPMOs with the same substrate specificity is in the order of 50%, whereas it is in the order of 30 % between LPMOs with different substrate specificity. The active site residues appeared to be highly conserved between all LPMOs considered in this study. These residues include the two histidines that form the T-shaped 'histidine brace' as well as the hydrophobic phenylalanine and alanine in the axial positions. It was noted that there are differences in the outer coordination sphere that could relate to both substrate binding and electron transport, as discussed further below and in Paper III.

Copper coordination in the four LPMOs (i.e. CelS2, E8, CBP21 and *Bl*LPMO10A) was analyzed by EPR spectroscopy. Prior to the study described in Paper II, EPR spectroscopy had been used to examine the copper active site in *Ta*LPMO9A (Quinlan et al., 2011) and *Ba*LPMO10A (Hemsworth et al., 2013a). The data generated by these two studies showed that the LPMO9 contains a typical type 2 copper center whereas the copper center geometry in the LPMO10 was concluded to fall between the typical type 1 and type 2 copper center geometries. This conclusion was reached by comparing obtained spin Hamiltonian parameters with a plot of typical type 1 and 2 proteins made by Peisach and Blumberg (1974). The overall shape of the EPR signal suggests that a type 2 classification for LPMO10s is appropriate (Hemsworth et al., 2013a). Hemsworth *et al.* (2013a) have suggested that the different copper site geometries reflect inherent differences between LPMOs in families AA9 and AA10 (Hemsworth et al., 2013a). This is, however, not in agreement with the observations made in Paper II, which shows that the EPR spectra recorded for cellulose-active LPMO10s (Paper II) are very similar to the EPR spectrum of (cellulose-active) *Ta*LPMO9A. The chitin active LPMOs (*Bl*LPMO10A and CBP21) gave EPR spectra that were similar to that observed for *Ba*LPMO10A. The observed differences within the LPMO10 family imply that the EPR signals and, thus, the electronic environment of the copper ion, are correlated with substrate-specificity rather than the type of LPMO family.

All LPMO9s characterized so far have a conserved tyrosine/tyrosinate in the active site and the hydroxyl group acts as an axial copper ligand in an octahedral geometry (Figure 10). In the LPMO10s analyzed in Paper II, this position is occupied by a phenylalanine. In order to investigate the role of this residue in copper coordination and LPMO mechanism,

CelS2 mutants F219A and F219Y were analyzed for relative activity and copper coordination was studied by EPR spectroscopy. The F219A mutant showed a reduced activity whereas the F219Y mutant was inactive. Based on analysis of the CelS2 structure, which was determined later (Paper III), it seems that a hydroxyl group added to the Cζ of the phenylalanine side chain may be positioned too close to the copper ion (1.7 Å; Figure 15). In LPMO9s, which contain a Tyr at this position and that coordinate copper in an octahedral geometry, the distance between the copper atom and the Tyr hydroxyl group is ~2.8 Å. This distance is compatible with a Jahn-Teller distorted geometry, where the axial ligands are positioned further away from the copper atom than the ligands in the equatorial positions. Clearly, positioning of the Tyr hydroxyl group "too close" to the copper ion might prevent correct coordination and possibly give a coordination geometry that is incompatible with catalysis. Interestingly, both CelS2 mutants showed only small variations in EPR spectra compared to wild type CelS2 with slightly lower g_z values and higher A_z

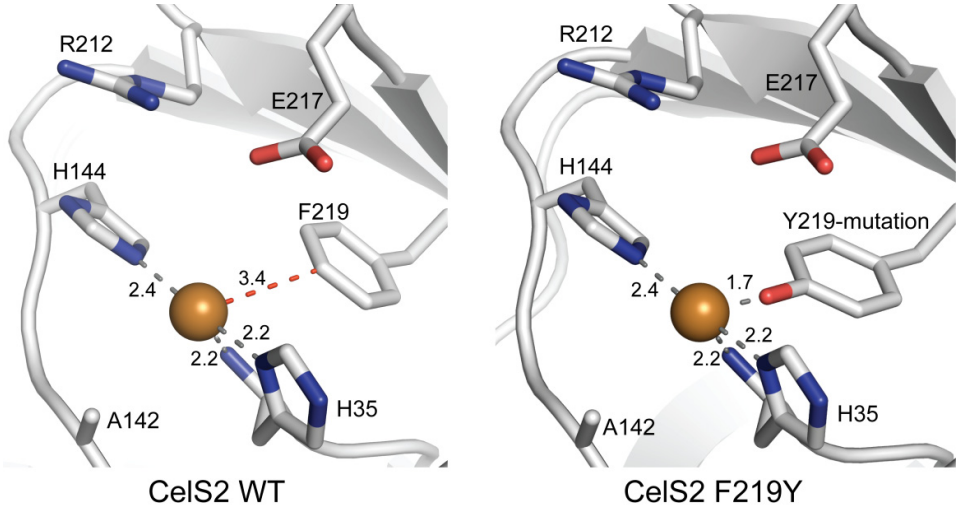


Figure 15. Active site of CelS2 in the wild type (WT) and the F219Y mutant. The F219Y structure represents a model that was generated using the mutagenesis function in PyMOL. The distances are taken from the CelS2 structure chain A (presented in Paper III) and illustrate that the tyrosine hydroxyl group is so close to the copper that structural adjustments will be needed. Such adjustments may cause a change in copper coordination, but could also entail displacement of the tyrosine side and main chain. The latter could affect close neighbors such as Glu217.

values. This shows that copper is still bound to the active site and that the electronics of the copper-site are "wildtype-like". Since the *g* and *A* Cu tensors reflect the active site copper coordination environment, the small g_z and A_z value changes may reflect variations in the copper geometry caused by the mutations relative to the wild type protein. To address this issue properly, the X-ray crystallographic structure of the CelS2 F219Y mutant could provide a better understanding as to the nature of the mechanisms of inactivation. A "reverse" mutation has earlier been made in LPMO9E from *T. terrestris* (*Tt*LPMO9E), which, judged by phylogenetic clustering (Vu et al., 2014), probably oxidizes C1. This mutation, Y153F, resulted in reduced cellulose conversion when the enzyme was combined with a cellulase preparation from *T. reesei* (Harris et al., 2010). The conserved active site aromatic residue (Tyr/Phe) is in close proximity to a conserved cluster of aromatic residues in the core of the LPMO10s (Figure 16), comprising three Trps in the chitin-active LPMOs

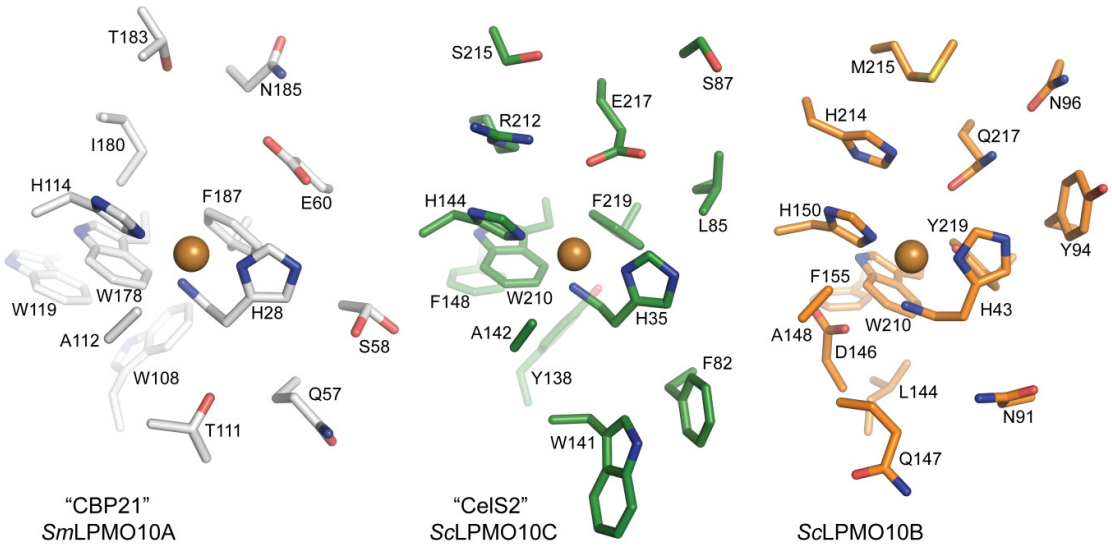


Figure 16. Expanded view of the active site of three LPMO10s with different substrate specificities. From the left: CBP21 (white) that oxidizes C1 in chitin, CelS2 (green) that oxidizes C1 in cellulose, and *Sc*LPMO10B (orange) that oxidizes C1 and C4 in cellulose and C1 in β-chitin. Shown is a top view of the copper active site including side chain residues that are protruding from the imaginary substrate binding surface and may be determinants of substrate specificity. The figures show the side chains of a conserved aromatic cluster that is located beneath the surface and that differs between the three presented LPMOs (W119, W108, W178; Y138, F148, W210; F155, W210, respectively).

and Tyr-Phe-Trp and Tyr-Trp-Trp in CelS2 and E8, respectively [(Vaaje-Kolstad et al., 2005a, Hemsworth et al., 2013a) and Paper II]. It is possible that the Phe/Tyr residue is an extension of this aromatic cluster and that it is involved in electron transfer. The Phe → Try mutation in CelS2 is likely to displace this residue, which again could affect a putative electron-transfer chain. Variation in the cluster of aromatic residues could perhaps also explain why the EPR signals of the chitin-active and cellulose-active LPMPO10s differ, despite the highly conserved direct coordination spheres of the copper ion (Figure 16).

The data discussed so far indicate that the substrate specificity of LPMO10s somehow is related to the electronics of the copper site, possibly combined with variation in substrate positioning that is brought about by the quite extensive sequence variation in surface areas outside the direct copper coordination sphere. The final experimental section of Paper II describes assays of the affinity of full length CelS2 and CelS2 lacking the cellulose binding CBM2 domain towards crystalline cellulose and chitin. Both the full length and truncated CelS2 variant showed stronger binding to chitin than to cellulose. This result is remarkable in the sense that CelS2 (full length or truncated form) is incapable of oxidizing chitin (this has been determined in experiments repeated several times for both α- and β-chitin). The binding experiment also shows that one cannot judge substrate preference based on binding experiment alone, which in accordance with findings on the "proximity effect" of CBMs discussed in section 1.4 (Hervé et al., 2010). It is known the CBM2s can bind chitin (Boraston et al., 2004), but it was unexpected that the truncated CelS2 also showed strong affinity to both chitin substrates and almost no binding to filter paper. It should be noted that the CBM2 is of importance for activity on cellulose, as the activity of the full length enzyme is substantially higher than the activity of the truncated variant (see Figure 2 and Figure S5 in Paper III). These unexpected findings regarding binding affinities add to the idea that fine-tuning of the geometry of the enzyme-substrate complex, perhaps affecting electron transfer paths, determines LPMO specificity.

In the final study of the thesis, the main goal was to obtain a more detailed structural and mechanistic understanding of LPMO10s. The focus was still on cellulose-active LPMOs, as these enzymes remained greatly unexplored at the time. In summary, Paper III describes the structures of CelS2 (*Sc*LPMO10C) and a second LPMO from *S. coelicolor*, *Sc*LPMO10B, as well as functional studies of these two enzymes addressing activity, metal-binding and redox properties. The CelS2 and *Sc*LPMO10B structures were the first to be solved for

cellulose-active LPMO10s. Activity experiments showed that the new enzyme, *Sc*LPMO10B, has a mixed C1 and C4 activity on cellulose substrates, demonstrating that C4 oxidation is not confined to LPMO9s. *Sc*LPMO10B also showed activity on β-chitin, with a preference for deacetylated regions of the substrate, but no products were obtained from α-chitin.

Homologues of CelS2 and *Sc*LPMO10B are found in other cellulolytic actinomycetes, such as *T. fusca*, where the corresponding LPMO10s are known as E7 and E8. The conservation of these two genes/ proteins indicates that they possibly have co-evolved and constitute a pair of enzymes that functionally complement each other. Analysis of the E7 and E8 pair from *T. fusca* showed activities that were very similar to the corresponding proteins from *S. coelicolor* (*Sc*LPMO10B and CelS2, respectively). In other words, E7 was active on cellulose and β-chitin whereas E8, the activity of which was already described in Paper II, showed strict C1-oxidation on cellulose, like CelS2. When the two activities (CelS2 and *Sc*LPMO10B or E7 and E8) where combined in a reaction, a clear synergy was observed. It does indeed seem like this pair of LPMOs exist to support each other's activity in cellulose degradation, much like what is observed for exo-processive cellobiohydrolases and endoglucanases (Henrissat et al., 1985, Nidetzky et al., 1993). There is no obvious explanation for the observed synergy. Possibly, the activity of one enzyme provides attachment sites for the other enzyme. Another option is that the two enzymes attack different parts/sides of the substrate. LPMO substrate hydroxylation generates both soluble products and products with higher DP that remain associated with the insoluble substrate. Only full degradation of LPMO- treated substrate with GHs can reveal the number of oxidative cleavages by the LPMO. In the study performed in Paper III, only the soluble fractions were analyzed, and therefore, a third possibility for the observed synergy is that the combined activity yields a higher degree of released (soluble) products. A final hypothesis could be that the LPMO enzymes themselves act as electron donors for each other, thus increasing the overall reducing power in the reaction and thereby the amount of released sugars.

The synergy/complementarity observed for the LPMO pair studied in Paper III may help explain the high number of LPMO genes in some microbial genomes. The most obvious reason for LPMO multiplicity is that these enzymes act on several substrates. Today, chitin, cellulose and β-(1→4)-glucan hemicelluloses have been identified, whereas other substrates

are likely to be discovered in the future. Paper III shows that microbes may use several LPMOs to simultaneously act on the same substrate.

Phylogenetic analysis of LPMOs (see Figure 4 in Paper III) revealed three major clusters of LPMO10s, which seem correlated to substrate specificity and regioselectivity of the hydroxylation they catalyze. The three clusters comprise chitin-active LPMOs, cellulose-active C1 oxidizing LPMOs and cellulose-active mixed C1/C4 oxidizing (and β-chitin oxidizing) LPMOs. The chitin active LPMOs further divide into two sub-groups where bacteria from Proteobacteria and Actinobacteria cluster differently (see Figure 4 in Paper III).

Importantly, the two structures of cellulose oxidizing LPMO10s allowed for the first time structural comparison of LPMO10s with different substrate specificities and also comparison of LPMOs from different families having the same substrate specificity. The overall structures of both CelS2 and *Sc*LPMO10B are similar to those of chitin-active LPMO10s, but with variation in loop 2 (L2) that comprises almost half of the substrate binding surface. In LPMO9s, several aromatic residues are located on the substrate binding surface and have been suggested to be involved in substrate binding (Li et al., 2012, Wu et al., 2013, Hemsworth et al., 2013b). In the structures determined for LPMO10s (cellulose- as well as chitin-active), only one aromatic amino acid (Tyr or Trp) is found on the substrate binding surface. This residue has been shown to be important for both binding and activity of the chitin active LPMO CBP21 (Vaaje-Kolstad et al., 2005a, Vaaje-Kolstad et al., 2005b).

The active sites of CelS2 and *Sc*LPMO10B both contain the conserved histidines that bind copper in the so called 'histidine brace' (Figure 17). The histidine brace is conserved in all LPMOs, the only variation being methylation of the Nε of the N-terminal histidine in LPMO9s expressed in their fungal hosts. The importance of this post-translational modification is unclear. Recent modelling studies of the catalytic mechanism [see section 1.5.6 (Kim et al., 2014)] showed that the methylation plays a minor role. In CBP21 (a bacterial LPMO10) isolated from its natural producer organism, *S. marcescens* and in all heterologously expressed LPMOs this posttranscriptional modification is absent (Paper II).

Some of the other active site residues do differ between CelS2 and *Sc*LPMO10B and all other structurally characterized LPMO10s (which all are likely to be active on chitin, based

on sequence alignments and phylogenetic clustering) (Figure 3 in Paper III). For the C1/C4 oxidizing *Sc*LPMO10B, the active site resembles the active site of LPMO9s that, in addition to the histidine brace, includes an axial tyrosine (3.3 Å from the copper in *Sc*LPMO10B) and a glutamine that interacts with a water molecule occupying the fourth equatorial position in the copper-site in some LPMO9 structures. No water molecule ligand was present in the *Sc*LPMO10B structure. *Sc*LPMO10B was crystallized in a high acetate concentration and an acetate ion close to the copper occupies the solvent facing axial position, while its other oxygen extends partially into the fourth equatorial position. The CelS2 active site is more similar to those of chitin active LPMO10s (Vaaje-Kolstad et al., 2005a, Vaaje-Kolstad et al., 2012, Hemsworth et al., 2013a), although differences are found in residues that are located close to the active site (Figure 16). For instance, the glutamic acid in CelS2 is located on strand nine, but structurally it has the same position as the glutamic acid located on strand two in CBP21. Glu217 in CelS2 forms a hydrogen bond with Arg212 that fills a surface cavity that is found in the previously determined LPMO10 structures. This cavity has been suggested to be the dioxygen binding pocket (Hemsworth et al., 2013a), but as it is absent in cellulose-active LPMOs it is more likely that the cavity in chitin-active LPMO10s interacts with an acetamido group in the substrate during chitin oxidation (Figure S8, Paper III). The absence of the cavity in *Sc*LPMO10B, which also

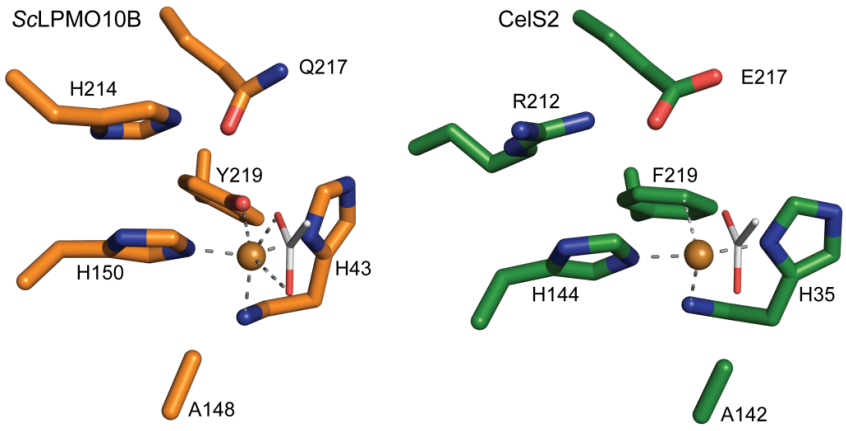


Figure 17. Active site of *Sc*LPMO10B (orange) and CelS2 (green). Note the position of the acetate ion (white) from the *Sc*LPMO10B structure, which makes up both an axial and a skewed equatorial ligand. When overlaid, the CelS2 alanine (Ala142) clearly is too close to the acetate ion indicating that the alanine restricts axial access to the copper ion.

showed activity on β-chitin, may explain the release of partially deacetylated oxidized products, as glucosamine resembles glucose rather than *N*-acetylglucosamine (the β-chitin used in the study had approximately eight percent deacetylated sugars).

CelS2 crystallized with eight molecules in the asymmetric unit showing variation in copper site bond lengths and geometry. All chains showed density in the equatorial position (refined as a water molecule) with a 1.6-2.0 Å distance from the copper ion and in proximity to the glutamate residue describe above. The average distance of 1.8 Å observed is most likely too short for a Cu-OH_2 bond, and it is thus possible that the density refined as a water molecule may represent a different ligand. The distances resulting from simulation of the Cu(II)-oxyl complex (Kim et al., 2014) indicate a bond length in the same range (1.8 Å), and this distance has also been reported from simulations on other enzymatic and synthetic copper hydroxylation reactions (Crespo et al., 2006, Yoshizawa et al., 2006, Huber et al., 2009, Kim et al., 2014). However, a Cu(II)-oxyl complex has never been experimentally observed in a protein crystal structure and it is thus unlikely that this complex is observed in the CelS2 structure.

As discussed in Paper II, the axial positions in LPMO10s are most commonly occupied by two hydrophobic residues, a phenylalanine and an alanine. The alanine has been postulated to prevent ligands from binding in the in the solvent-facing axial position, rendering dioxygen activation more likely in the equatorial position (Hemsworth et al., 2013a, Hemsworth et al., 2013b). Dioxygen binding has been proposed to occur from the solvent-facing axial position in *Ta*LPMO9A that oxidizes both C1 and C4, and lacks the alanine (Kim et al., 2014). In cellulose- oxidizing LPMO10s, this alanine is conserved in the active site, but the structural studies described in Paper III revealed that its position differs significantly between LPMO10s that are known to oxidize C1 (in cellulose and chitin) and *Sc*LPMO10B, which also can oxidize the C4 position (in cellulose). Moreover the structure shows that *Sc*LPMO10B coordinates a ligand (acetate) in the solvent-facing axial position, despite the presence of the alanine. Comparison of the two structures of cellulose-oxidizing LPMO10 shows that the loop hosting the conserved alanine adopts different conformations. In *Sc*LPMO10B, Asp146 forces the main chain of the loop into a conformation that positions Ala148 2.5 Å away from the position of the corresponding alanine (Ala142) in CelS2, allowing sufficient space for a ligand to bind in the solvent-facing axial position. Asp146 is conserved in LPMO10s that, according to phylogenetic

grouping, are likely to exhibit C1/C4 oxidation activity. Importantly, closer inspection of available LPMO9 structures revealed a similar scenario where LPMO9s with C4 and C1/C4 oxidizing activity have an open solvent-facing axial coordination site, whereas LPMO9s known to be strict C1 oxidizers have a tyrosine preventing optimal axial access to the copper ion (see Figure S15 in Paper III). Thus, the ability to bind a ligand in the axial position could be a determinant of C4 oxidizing activity.

The affinity of LPMOs for copper has been subject of debate especially since Quinlan *et al.* (2011) indicated that LPMO9s have a K_d well below1 nM for this metal ion. Phillips *et al.* (2011) also reported high affinity for the *N. crassa* LPMOs, but no dissociation constants were reported. The first well documented dissociation constant for Cu^{2+} was determined for CBP21, and estimated to be 55 nM at pH 6.5 (Aachmann et al., 2012). Subsequently, Hemsworth *et al.* (2013a) reported a K_d of 6-80 nM for *Ba*LPMO10A in the pH range of 5-7. In Paper III, Cu^{2+} dissociation constants for the two cellulose-active LPMOs were measured with ITC at pH 5.5 and determined to be 31 nM (CelS2) and 12 nM (*Sc*LPMO10B). Thus, it seems that LPMO10s have comparable dissociation constants for Cu^{2+}, regardless of substrate specificity. The limited available data of fungal LPMO9s as well as data on the only characterized LPMO11 indicate that fungal LPMOS have a somewhat higher affinity for Cu^{2+} (Quinlan et al., 2011, Hemsworth et al., 2014). The copper-binding site of the bacterial copper binding protein CopC, involved in a copper resistance response, is very similar to that of the LPMOs. For CopC the K_d for Cu^{2+} was estimated to be in the picomolar range at pH 7.0, using equilibrium ligand competition reactions (Zhang et al., 2006). All in all, available data suggest that proteins with seemingly very similar copper binding sites may show considerable differences in copper affinity, which indicates that binding affinity in not only determined by the copper ligands, but also by other macromolecular properties.

In order to determine the dissociation constants for Cu^{+}, the cell potentials for the LPMO10-Cu^{2+}/LPMO10-Cu^{+} redox couple were determined. The data obtained showed a cell potential of 242 mV and 251 mV, for CelS2 and *Sc*LPMO10B respectively. These values correspond to an average increase of ~90 mV compared to the redox potential for Cu^{2+}/Cu^{+} in water. By combining the redox potentials and the Cu^{2+} dissociation constant in three thermodynamic relationships the dissociation constant for LPMO-Cu^{+} could be estimated, resulting in values of 1.1 nM for CelS2 and 0.3 nM for *Sc*LPMO10B,

respectively. The values for the redox potential and the binding constants for Cu^{2+} and Cu^{+} are similar to values previously determined for chitin-active CBP21 (Aachmann et al., 2012).

Finally, copper coordination in CelS2 and *Sc*LPMO10B was also investigated by EPR spectroscopy as a follow up study to Paper II. The EPR spectrum for *Sc*LPMO10B was similar to the EPR spectra obtained for other cellulose active LPMOs (AA9 & AA10) even though it also can hydroxylate β-chitin. Interestingly, in a Peisach-Blumberg plot containing all LPMOs with known EPR parameters and a selection of typical type 1 and type 2 copper proteins, the chitin-active LPMO10s grouped differently, displaying a lower A_z^{Cu} tensor than cellulose-active LPMOs (from family AA9 and AA10). A common feature for chitin-active LPMO10s is the presence of a cavity close to the active site that potentially accommodates an *N*-acetylgroup from the chitin substrate. In the cellulose-active LPMOs this cavity is "filled" with a positively charged amino acid side chain (R212 in CelS2 and H214 in *Sc*LPMO10B) that contributes to a flat surface topology. This difference in electronic environment could perhaps cause the variations that are observed in the ERP spectra of chitin- and cellulose-active LPMOs. This hypothesis is only valid for LPMO10s since the structure of (chitin-active) *Ao*LPMO11 also shows a cavity adjacent to the active site (Hemsworth et al., 2014), but has an EPR signal that resembles those observed for cellulose-active LPMOs. It should be noted, however, that the structure of *Ao*LPMO11 is incomplete as two loops are missing, including one that likely contributes to the substrate-binding surface close to the active site. Thus the crystal structure may not reflect the real active site structure.

All in all the data in Paper III provide important novel information on LPMO structure and function, and one of its most important conclusions is that there do not seem to be fundamental differences between fungal and bacterial LPMOs. LPMOs display a continuum of active site configurations, but the observed variations cannot easily be coupled to variation in substrate specificities and/or oxidative power. Together, Papers II and III suggest that LPMO functionality is determined by variation outside the direct copper coordination sphere.

4 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The discovery of LPMOs, which are able to cleave glycosidic bonds in a crystalline context, has changed the paradigm of enzymatic conversion of polysaccharides and thereby provided new enzyme tools for biotechnology and biorefining. At this early stage of LPMO research, there is a need for fundamental work to unveil the molecular mechanism of these enzymes and how they can be exploited in biotechnological processes. The work presented in this thesis comprises the discovery and the first characterization of LPMOs from the family 10 of auxiliary activities that are active on cellulose substrates.

In terms of their function in cellulose degradation, this study sheds light on how LPMOs influence the activity of GHs and how different LPMOs may work together. In Paper I, the first cellulose-active LPMO is described and this enzyme showed synergistic action when combined with various cellulolytic glycoside hydrolases. Paper III describes another type of synergy observed within a conserved pair of LPMO10s with different regioselectivities of hydroxylation. What underlays these synergies is still poorly understood, despite the high value of LPMOs for industrial purposes, which is due to the fact that they tend to significantly increase the enzymatic hydrolysis rates in the conversion of lignocellulosic biomass. Experimental data indicate that synergy is dependent on the type of enzyme, enzyme ratios, as well as the type of the substrate, but, for reasons discussed above, it is clear that more studies are needed to understand and maximize synergistic effects. One key issue, discussed above, will be to create conditions in which all factors potentially limiting LPMO activity (e.g. copper concentration, electron availability, substrate concentration, damage by H_2O_2 production) are under control.

For better design of enzyme cocktail compositions, a good starting point is to study the secretomes/transcriptomes of microorganism grown on different biomasses. Such studies will increase our understanding of how LPMOs are regulated in nature, i.e. in which phase of the degradation process they are expressed, in what quantities, and with which other CAZymes, which may give clues to optimal enzyme combinations. A problem that limits the value of much of the current knowledge on LPMOs is the use of heavily processed substrates. This is a critical aspect, as enzymes may perform differently on model substrates than on realistic substrates. Clearly, more applied research on industrially realistic biomass substrates is needed.

The recent finding of a xyloglucan active LPMO9 and the variation in domain composition seen for AA10-type LPMOs (Figure 6) indicate that the substrate spectrum of LPMOs may be wider then first anticipated. Such widening may have implications beyond biomass processing, as suggested by the fact that AA10 sequences are also found in the genomes of organisms that are not obviously connected to biomass deconstruction (e.g. in viruses and probiotic bacteria).

Papers II and III provide a wealth of in-depth structure-function data for cellulose-active LPMO10s, without revealing differences that clearly can be correlated to LPMO substrate specificity and performance. From available structure data, it seems that LPMOs use a continuum of active site configurations, with the histidine brace being the key conserved element. Sequence diversity is found outside the direct copper coordination sphere, which is likely to affect both substrate affinity and the positioning of bound substrates relative to the catalytic center. The latter, combined with variation in the electronic structure of the copper site revealed by EPR, may determine LPMO substrate specificity and/or the site of oxidation. It is conceivable that precise aligning of a particular substrate affects the copper-site (geometry, redox potential, electronic structure) and that this is essential for enzyme activity. The cavity that is unique for chitin-active LPMO10s and that potentially accommodate an *N*-actetylgroup may be important in this respect. A final, and potentially very important, finding described in Paper III concerns the variation in the position of the active site alanine, which might determine regioselectivity (i.e. C1 versus C4 oxidation). The experimental data and considerations described in Papers II and III provide an excellent basis for further work. Most ideas coming out of these papers may be tested experimentally by performing site directed mutagenesis experiments and such work is currently in progress.

This thesis represents a step forward in the understanding of how LPMOs operate, provides new leads for designing enzyme cocktails, and puts forward novel ideas concerning the LPMO reaction mechanism. The current increase in the number of researchers pursuing the secrets of LPMOs will surely have a dramatic impact on our understanding of this captivating enzyme family in the years to come.

5 REFERENCES

ADAV, S. S., NG, C. S., ARULMANI, M. & SZE, S. K. (2010) Quantitative iTRAQ Secretome Analysis of Cellulolytic *Thermobifida fusca*. *J Proteome Res,* 9, 3016-3024.

ADAV, S. S., RAVINDRAN, A. & SZE, S. K. (2012) Quantitative proteomic analysis of lignocellulolytic enzymes by *Phanerochaete chrysosporium* on different lignocellulosic biomass. *J Proteomics,* 75, 1493-1504.

AGGER, J. W., ISAKSEN, T., VÁRNAI, A., VIDAL-MELGOSA, S., WILLATS, W. G., 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. *Proc Natl Acad Sci U S A,* 111, 6287-6292.

BANERJEE, G., CAR, S., SCOTT-CRAIG, J. S., BORRUSCH, M. S. & WALTON, J. D. (2010) Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnol Biofuels,* 3, 22.

BAYER, E. A., CHANZY, H., LAMED, R. & SHOHAM, Y. (1998) Cellulose, cellulases and cellulosomes. *Curr Opin Struct Biol,* 8, 548-557.

BAYER, E. A., LAMED, R., WHITE, B. A. & FLINT, H. J. (2008) From cellulosomes to cellulosomics. *Chem Rec,* 8, 364-377.

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. *J Phys Chem B,* 115, 4118-4127.

BECKHAM, G. T., STÅHLBERG, J., KNOTT, B. C., HIMMEL, M. E., CROWLEY, M. F., SANDGREN, M., SØRLIE, M. & PAYNE, C. M. (2014) Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases. *Curr Opin Biotech,* 27, 96-106.

BEESON, W. T., PHILLIPS, C. M., CATE, J. H. & MARLETTA, M. A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc,* 134, 890-892.

BENTLEY, S. D., CHATER, K. F., CERDENO-TARRAGA, A. M., CHALLIS, G. L., THOMSON, N. R., JAMES, K. D., HARRIS, D. E., QUAIL, M. A., KIESER, H., HARPER, D., BATEMAN, A., BROWN, S., CHANDRA, G., CHEN, C. W., COLLINS, M., CRONIN, A., FRASER, A., GOBLE, A., HIDALGO, J., HORNSBY, T., HOWARTH, S., HUANG, C. H., KIESER, T., LARKE, L., MURPHY, L., OLIVER, K., O'NEIL, S., RABBINOWITSCH, E., RAJANDREAM, M. A., RUTHERFORD, K., RUTTER, S., SEEGER, K., SAUNDERS, D., SHARP, S., SQUARES, R., SQUARES, S., TAYLOR, K., WARREN, T., WIETZORREK, A., WOODWARD, J., BARRELL, B. G., PARKHILL, J. & HOPWOOD, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature,* 417, 141-147.

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., 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*. *Nat Biotechnol,* 29, 922-927.

BLAKE, A. W., MCCARTNEY, L., FLINT, J. E., BOLAM, D. N., BORASTON, A. B., GILBERT, H. J. & KNOX, J. P. (2006) Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes. *J Biol Chem,* 281, 29321-29329.

BOLAM, D. N., CIRUELA, A., MCQUEEN-MASON, S., SIMPSON, P., WILLIAMSON, M. P., RIXON, J. E., BORASTON, A., HAZLEWOOD, G. P. & GILBERT, H. J. (1998) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochem J,* 331, 775-781.

BONDESON, D., MATHEW, A. & OKSMAN, K. (2006) Optimization of the isolation of nanocrystals from microcrystalline cellulose by acid hydrolysis. *Cellulose,* 13, 171-180.

BOOK, A. J., LEWIN, G. R., MCDONALD, B. R., TAKASUKA, T. E., DOERING, D. T., ADAMS, A. S., BLODGETT, J. A. V., CLARDY, J., RAFFA, K. F., FOX, B. G. & CURRIE, C. R. (2014) Cellulolytic *Streptomyces* strains associated with herbivorous insects share a phylogenetically-linked capacity for the degradation of lignocellulose. *Appl Environl Microb*, doi:10.1128/aem.01133-01114.

BORASTON, A. B., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J,* 382, 769-781.

BRAMUCCI, E., PAIARDINI, A., BOSSA, F. & PASCARELLA, S. (2012) PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics,* 13 Suppl 4, S2.

BRETT, C. T. (2000) Cellulose microfibrils in plants: Biosynthesis, deposition, and integration into the cell wall. *Int Rev Cytol,* 199, 161-199.

BRETT, C. T. & WALDRON, K. (1990) *Physiology and biochemistry of plant cell walls,* London ; Boston, Unwin Hyman.

BROWN, R. M. (2004) Cellulose structure and biosynthesis: What is in store for the 21st century? *J Polym Sci A1,* 42, 487-495.

CANNELLA, D., HSIEH, C. W. C., FELBY, C. & JØRGENSEN, H. (2012) Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol Biofuels,* 5, 26.

CANNELLA, D. & JØRGENSEN, H. (2014) Do new cellulolytic enzyme preparations affect the industrial strategies for high solids lignocellulosic ethanol production? *Biotechnol Bioeng,* 111, 59-68.

CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V. & HENRISSAT, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res,* 37, D233-238.

CHANG, H. N., KIM, N. J., KANG, J. & JEONG, C. M. (2010) Biomass-derived Volatile Fatty Acid Platform for Fuels and Chemicals. *Biotechnol Bioproc E,* 15, 1-10.

CHRISTIANSEN, C., ABOU HACHEM, M., JANECEK, S., VIKSO-NIELSEN, A., BLENNOW, A. & SVENSSON, B. (2009) The carbohydrate-binding module family 20-diversity, structure, and function. *FEBS J,* 276, 5006-5029.

CHU, H. H., HOANG, V., HOFEMEISTER, J. & SCHREMPF, H. (2001) A *Bacillus amyloliquefaciens* ChbB protein binds beta- and alpha-chitin and has homologues in related strains. *Microbiology,* 147, 1793-1803.

CHUNDAWAT, S. P., BECKHAM, G. T., HIMMEL, M. E. & DALE, B. E. (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu Rev Chem Biomol Eng,* 2, 121-145.

COCINERO, E. J., GAMBLIN, D. P., DAVIS, B. G. & SIMONS, J. P. (2009) The building blocks of cellulose: the intrinsic conformational structures of cellobiose, its epimer, lactose, and their singly hydrated complexes. *J Am Chem Soc,* 131, 11117-11123.

COSGROVE, D. J. (2000) Loosening of plant cell walls by expansins. *Nature,* 407, 321-326.

CRESPO, A., MARTI, M. A., ROITBERG, A. E., AMZEL, L. M. & ESTRIN, D. A. (2006) The catalytic mechanism of peptidylglycine alpha-hydroxylating monooxygenase investigated by computer simulation. *J Am Chem Soc,* 128, 12817-12828.

CUSKIN, F., FLINT, J. E., GLOSTER, T. M., MORLAND, C., BASLE, A., HENRISSAT, B., COUTINHO, P. M., STRAZZULLI, A., SOLOVYOVA, A. S., DAVIES, G. J. & GILBERT, H. J. (2012) How nature can exploit nonspecific catalytic and carbohydrate binding modules to create enzymatic specificity. *Proc Natl Acad Sci U S A,* 109, 20889-20894.

DE VRIES, R. P. & VISSER, J. (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev,* 65, 497-522.

DELANO, W. L. & LAM, J. W. (2005) PyMOL: A communications tool for computational models. *Abstr Pap Am Chem S,* 230, U1371-U1372.

DIMAROGONA, M., TOPAKAS, E., OLSSON, L. & CHRISTAKOPOULOS, P. (2012) Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresource Technol,* 110, 480-487.

DING, S. Y. & HIMMEL, M. E. (2006) The maize primary cell wall microfibril: a new model derived from direct visualization. *J Agric Food Chem,* 54, 597-606.

DUCHESNE, L. C. & LARSON, D. W. (1989) Cellulose and the Evolution of Plant Life. *Bioscience,* 39, 238-241.

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., KUES, U., HIBBETT, D. S., HOFFMEISTER, D., HOGBERG, 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, 762-765.

EIJSINK, V. G. H., VAAJE-KOLSTAD, G., VÅRUM, K. M. & HORN, S. J. (2008) Towards new enzymes for biofuels: lessons from chitinase research. *Trends Biotechnol,* 26, 228-235.

ENDLER, A. & PERSSON, S. (2011) Cellulose synthases and synthesis in *Arabidopsis*. *Mol Plant,* 4, 199-211.

FERNANDES, A. N., THOMAS, L. H., ALTANER, C. M., CALLOW, P., FORSYTH, V. T., APPERLEY, D. C., KENNEDY, C. J. & JARVIS, M. C. (2011) Nanostructure of cellulose microfibrils in spruce wood. *Proc Natl Acad Sci U S A,* 108, 1195-1203.

FINN, R. D., MISTRY, J., TATE, J., COGGILL, P., HEGER, A., POLLINGTON, J. E., GAVIN, O. L., GUNASEKARAN, P., CERIC, G., FORSLUND, K., HOLM, L., SONNHAMMER, E. L., EDDY, S. R. & BATEMAN, A. (2010) The Pfam protein families database. *Nucleic Acids Res,* 38, D211-222.

FOLDERS, J., TOMMASSEN, J., VAN LOON, L. C. & BITTER, W. (2000) Identification of a chitin-binding protein secreted by *Pseudomonas aeruginosa*. *J Bacteriol,* 182, 1257-1263.

FONTES, C. M. & GILBERT, H. J. (2010) Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu Rev Biochem,* 79, 655-681.

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 Sci,* 20, 1479-1483.

GARDA, A. L., FERNANDEZ-ABALOS, J. M., SANCHEZ, P., RUIZ-ARRIBAS, A. & SANTAMARIA, R. I. (1997) Two genes encoding an endoglucanase and a cellulose-binding protein are clustered and co-regulated by a TTA codon in *Streptomyces halstedii* JM8. *Biochem J,* 324 403-411.

GARDNER, K. H. & BLACKWELL, J. (1975) Refinement of the structure of beta-chitin. *Biopolymers,* 14, 1581-1595.

GILBERT, H. J. (2010) The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol,* 153, 444-455.

GILBERT, H. J., BOLAM, D. N., SZABO, L., XIE, H., WILLIAMSON, M. P., SIMPSON, P. J., JAMAL, S., BORASTON, A. B., KILBURN, D. G., ANTHONY, R. & WARREN, J. (2002) An update on carbohydrate binding modules. *Roy Soc Ch,* 275, 89-98.

GOODAY, G. (1990) The Ecology of Chitin Degradation. *In:* MARSHALL, K. C. (ed.) *Advances in Microbial Ecology.* Springer US.

HARRIS, P. V., WELNER, D., MCFARLAND, K. C., RE, E., POULSEN, J. C. N., BROWN, K., SALBO, R., DING, H. S., VLASENKO, E., MERINO, S., 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, 3305-3316.

HEMSWORTH, G. R., DAVIES, G. J. & WALTON, P. H. (2013b) Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol,* 23, 660-668.

HEMSWORTH, G. R., HENRISSAT, B., DAVIES, G. J. & WALTON, P. H. (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol,* 10, 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. (2013a) The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc,* 135, 6069-6077.

HENRISSAT, B. & DAVIES, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol,* 7, 637-644.

HENRISSAT, B., DRIGUEZ, H., VIET, C. & SCHULEIN, M. (1985) Synergism of Cellulases from *Trichoderma reesei* in the Degradation of Cellulose. *Bio-Technol,* 3, 722-726.

HERVÉ, C., ROGOWSKI, A., BLAKE, A. W., MARCUS, S. E., GILBERT, H. J. & KNOX, J. P. (2010) Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci U S A,* 107, 15293-15298.

HIGUCHI, T. (1990) Lignin Biochemistry - Biosynthesis and Biodegradation. *Wood Sci Technol,* 24, 23-63.

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, 804-807.

HODGSON, D. A. (2000) Primary metabolism and its control in streptomycetes: A most unusual group of bacteria. *Adv Microb Physiol,* 42, 47-238.

HOPWOOD, D. A. (1999) Forty years of genetics with *Streptomyces*: from in vivo through in vitro to in silico. *Microbiol-Uk,* 145, 2183-2202.

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, 14-23.

HORN, S. J., SIKORSKI, P., CEDERKVIST, J. B., VAAJE-KOLSTAD, G., SØRLIE, M., SYNSTAD, B., VRIEND, G., VÅRUM, K. M. & EIJSINK, V. G. H. (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci U S A,* 103, 18089-18094.

HORN, S. J., SØRLIE, M., VÅRUM, K. M., VÄLJAMÄE, P. & EIJSINK, V. G. H. (2012a) Measuring Processivity. *Cellulases,* 510, 69-95.

HORN, S. J., VAAJE-KOLSTAD, G., WESTERENG, B. & EIJSINK, V. G. H. (2012b) Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels,* 5, 45.

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. *Energ Environ Sci*, doi:10.1039/C1034EE00891J.

HUBER, S. M., ERTEM, M. Z., AQUILANTE, F., GAGLIARDI, L., TOLMAN, W. B. & CRAMER, C. J. (2009) Generating Cu(II)-oxyl/Cu(III)-oxo species from Cu(I)-alpha-ketocarboxylate complexes and O2: in silico studies on ligand effects and C-H-activation reactivity. *Chem-Eur J,* 15, 4886-4895.

IGARASHI, K., UCHIHASHI, T., KOIVULA, A., WADA, M., KIMURA, S., OKAMOTO, T., PENTTILA, M., ANDO, T. & SAMEJIMA, M. (2011) Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science,* 333, 1279-1282.

IGARASHI, K., WADA, M. & SAMEJIMA, M. (2007) Activation of crystalline cellulose to cellulose IIII results in efficient hydrolysis by cellobiohydrolase. *FEBS J,* 274, 1785-1792.

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. *J Biol Chem,* 289, 2632-2642.

JØRGENSEN, H., KRISTENSEN, J. B. & FELBY, C. (2007) Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuel Bioprod Bior,* 1, 119-134.

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. *J Mol Biol,* 383, 144-154.

KARLSSON, J., SALOHEIMO, M., SIIKA-AHO, M., TENKANEN, M., PENTTILÄ, M. & TJERNELD, F. (2001) Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur J Biochem,* 268, 6498-6507.

KERFF, F., AMOROS, A., HERMAN, R., SAUVAGE, E., PETRELLA, S., FILEE, P., CHARLIER, P., JORIS, B., TABUCHI, A., NIKOLAIDIS, N. & COSGROVE, D. J. (2008) Crystal structure and activity of *Bacillus subtilis* YoaJ (EXLX1), a bacterial expansin that promotes root colonization. *Proc Natl Acad Sci U S A,* 105, 16876-16881.

KIM, S., STÅHLBERG, J., SANDGREN, M., PATON, R. S. & BECKHAM, G. T. (2014) Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc Natl Acad Sci U S A,* 111, 149-154.

KIRN, T. J., JUDE, B. A. & TAYLOR, R. K. (2005) A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature,* 438, 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. *Biotechnol Biofuels,* 5, 79.

KOBAYASHI, K., KAMIYA, S. & ENOMOTO, N. (1996) Amylose-carrying styrene macromonomer and its homo- and copolymers: Synthesis via enzyme-catalyzed polymerization and complex formation with iodine. *Macromolecules,* 29, 8670-8676.

KOBAYASHI, K., SUMITOMO, H. & INA, Y. (1985) Synthesis and Functions of Polystyrene Derivatives Having Pendant Oligosaccharides. *Polym J,* 17, 567-575.

KOLBE, S., FISCHER, S., BECIREVIC, A., HINZ, P. & SCHREMPF, H. (1998) The *Streptomyces reticuli* alpha-chitin-binding protein CHB2 and its gene. *Microbiology,* 144, 1291-1297.

KOSTYLEV, M. & WILSON, D. (2013) Two-parameter kinetic model based on a time-dependent activity coefficient accurately describes enzymatic cellulose digestion. *Biochemistry,* 52, 5656-5664.

KURASIN, M. & VÄLJAMÄE, P. (2011) Processivity of Cellobiohydrolases Is Limited by the Substrate. *J Biol Chem,* 286, 169-177.

LANGSTON, J. A., SHAGHASI, T., ABBATE, E., XU, F., VLASENKO, E. & SWEENEY, M. D. (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microb,* 77, 7007-7015.

LESCHINE, S. B. (1995) Cellulose Degradation in Anaerobic Environments. *Annu Rev Microbiol,* 49, 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. *Biotechnol Biofuels,* 6, 41.

LI, X., BEESON, W. T. T., PHILLIPS, C. M., MARLETTA, M. A. & CATE, J. H. (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure,* 20, 1051-1061.

LIETH, H. & WHITTAKER, R. H. (1975) *Primary productivity of the biosphere,* New York, Springer-Verlag.

LJUNGDAHL, L. G. & ERIKSSON, K. E. (1985) Ecology of Microbial Cellulose Degradation. *Adv Microb Ecol,* 8, 237-299.

LOMBARD, V., RAMULU, H. G., DRULA, E., COUTINHO, P. M. & HENRISSAT, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res,* 42, D490-D495.

LUND, H. (2007) Renewable energy strategies for sustainable development. *Energy,* 32, 912-919.

LYND, L. R., WEIMER, P. J., VAN ZYL, W. H. & PRETORIUS, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev,* 66, 506-577.

MARTINEZ, D., BERKA, R. M., HENRISSAT, B., SALOHEIMO, M., ARVAS, M., BAKER, S. E., CHAPMAN, J., CHERTKOV, O., COUTINHO, P. M., CULLEN, D., DANCHIN, E. G. J., GRIGORIEV, I. V., HARRIS, P., JACKSON, M., KUBICEK, C. P., HAN, C. S., HO, I., LARRONDO, L. F., DE LEON, A. L., MAGNUSON, J. K., MERINO, S., MISRA, M., NELSON, B., PUTNAM, N., ROBBERTSE, B., SALAMOV, A. A., SCHMOLL, M., TERRY, A., THAYER, N., WESTERHOLM-PARVINEN, A., SCHOCH, C. L., YAO, J., BARABOTE, R., NELSON, M. A., DETTER, C., BRUCE, D., KUSKE, C. R., XIE, G., RICHARDSON, P., ROKHSAR, D. S., LUCAS, S. M., RUBIN, E. M., DUNN-COLEMAN, N., WARD, M. & BRETTIN, T. S. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol,* 26, 1193-1193.

MERINO, S. T. & CHERRY, J. (2007) Progress and challenges in enzyme development for Biomass utilization. *Adv Biochem Eng Biot,* 108, 95-120.

MINKE, R. & BLACKWELL, J. (1978) The structure of alpha-chitin. *J Mol Biol,* 120, 167-181.

MOSER, F., IRWIN, D., CHEN, S. & WILSON, D. B. (2008) Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. *Biotechnol Bioeng,* 100, 1066-1077.

MUTWIL, M., DEBOLT, S. & PERSSON, S. (2008) Cellulose synthesis: a complex complex. *Curr Opin Plant Biol,* 11, 252-257.

NAKAGAWA, Y. S., EIJSINK, V. G. H., TOTANI, K. & VAAJE-KOLSTAD, G. (2013) Conversion of alpha-chitin substrates with varying particle size and crystallinity reveals substrate preferences of the chitinases and lytic polysaccharide monooxygenase of *Serratia marcescens*. *J Agric Food Chem,* 61, 11061-11066.

NIDETZKY, B., HAYN, M., MACARRON, R. & STEINER, W. (1993) Synergism of *Trichoderma reesei* Cellulases While Degrading Different Celluloses. *Biotechnol Lett,* 15, 71-76.

NIKOLAIDIS, N., DORAN, N. & COSGROVE, D. J. (2014) Plant expansins in bacteria and fungi: evolution by horizontal gene transfer and independent domain fusion. *Mol Biol Evol,* 31, 376-386.

NISHIYAMA, Y., LANGAN, P. & CHANZY, H. (2002) Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc,* 124, 9074-9082.

O'SULLIVAN, A. C. (1997) Cellulose: the structure slowly unravels. *Cellulose,* 4, 173-207.

OLSSON, L., JØRGENSEN, H., KROGH, K. B. R. & ROCA, C. (2004) *Bioethanol production from lignocellulosic material. In: Dumitriu S (ed) Polysaccharides: structural diversity and functional versatility.*, CRC Press, USA, 957-993.

PARTHASARATHI, R., BELLESIA, G., CHUNDAWAT, S. P., DALE, B. E., LANGAN, P. & GNANAKARAN, S. (2011) Insights into hydrogen bonding and stacking interactions in cellulose. *J Phys Chem A,* 115, 14191-14202.

PEISACH, J. & BLUMBERG, W. E. (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins. *Arch Biochem Biophys,* 165, 691-708.

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 Chem Biol,* 6, 1399-1406.

POIDEVIN, L., BERRIN, J. G., BENNATI-GRANIER, C., LEVASSEUR, A., HERPOËL-GIMBERT, I., CHEVRET, D., COUTINHO, P. M., HENRISSAT, B., HEISS-BLANQUET, S. & RECORD, E. (2014) Comparative analyses of *Podospora anserina* secretomes reveal a large array of lignocellulose-active enzymes. *Appl Microbiol Biot*, doi:10.1007/s00253-00014-05698-00253.

QUINLAN, R. J., SWEENEY, M. D., LO LEGGIO, L., OTTEN, H., POULSEN, J. C., JOHANSEN, K. S., KROGH, K. B., JORGENSEN, C. I., TOVBORG, M., ANTHONSEN, A., TRYFONA, T., WALTER, C. P., DUPREE, P., XU, F., DAVIES, G. J. & WALTON, P. H. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A,* 108, 15079-15084.

RAMACHANDRAN, S., MAGNUSON, T. S. & CRAWFORD, D. L. (2000) Cloning, Sequencing, and Characterization of Two Clustered Cellulase-Encoding Genes, celS1 and celS2, from *Streptomyces viridosporus* T7A and their Expression in *Escherichia coli*. *Actinomycetologica,* 14, 11-16.

REESE, E. T., SIU, R. G. & LEVINSON, H. S. (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J Bacteriol,* 59, 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. *Energ Environ Sci,* 6, 1858-1867.

RESCH, M. G., DONOHOE, B. S., CIESIELSKI, P. N., NILL, J. E., MAGNUSSON, L., HIMMEL, M. E., MITTAL, A., KATAHIRA, R., BIDDY, M. J. & BECKHAM, G. T. (2014) Clean fractionation pretreatment reduces enzyme loadings for biomass saccharification and reveals the mechanism of free and cellulosomal enzyme synergy. *ACS Sustainable Chem Eng,* 2, 1377-1387.

ROSS, P., MAYER, R. & BENZIMAN, M. (1991) Cellulose biosynthesis and function in bacteria. *Microbiol Rev,* 55, 35-58.

RUBIN, E. M. (2008) Genomics of cellulosic biofuels. *Nature,* 454, 841-845.

SALOHEIMO, M., NAKARI-SETÄLÄ, T., TENKANEN, M. & PENTTILÄ, M. (1997) cDNA cloning of a *Trichoderma reesei* cellulase and demonstration off endoglucanase activity expression in yeast. *Eur J Biochem,* 249, 584-591.

SALOHEIMO, M., PALOHEIMO, M., HAKOLA, S., PERE, J., SWANSON, B., NYYSSÖNEN, E., BHATIA, A., WARD, M. & PENTTILÄ, M. (2002) Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur J Biochem,* 269, 4202-4211.

SANCHEZ, B., GONZALEZ-TEJEDO, C., RUAS-MADIEDO, P., URDACI, M. C. & MARGOLLES, A. (2011) Lactobacillus plantarum extracellular chitin-binding

protein and its role in the interaction between chitin, Caco-2 cells, and mucin. *Appl Environ Microbiol,* 77, 1123-1126.

SCHELLER, H. V. & ULVSKOV, P. (2010) Hemicelluloses. *Annu Rev Plant Biol,* 61, 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. *Mol Microbiol,* 13, 807-819.

SCHREMPF, H., KOLBE, S., BECIREVIC, A. & ZELTINS, A. (1999) Characteristics of chitin-binding proteins from streptomycetes. *Roy Soc Ch*, 227-231.

SIMPSON, P. J., XIE, H., BOLAM, D. N., GILBERT, H. J. & WILLIAMSON, M. P. (2000) The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. *J Biol Chem,* 275, 41137-41142.

SJÖSTRÖM, E. (1993) *Wood polysaccharides: Wood chemistry, Fundamentals and Applications*, Academic Press Inc., San Diago. 1-20, 51-70.

SOLOMON, B. D. (2010) Biofuels and sustainability. *Ann N Y Acad Sci,* 1185, 119-134.

SOMERVILLE, C. (2006) Cellulose synthesis in higher plants. *Annu Rev Cell Dev Bi,* 22, 53-78.

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, 2206-2211.

STICKLEN, M. B. (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol *Nat Rev Genet,* 9, 433-443.

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. *Appl Environ Microb,* 66, 664-670.

SUZUKI, K., SUZUKI, M., TAIYOJI, M., NIKAIDOU, N. & WATANABE, T. (1998) Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosci Biotech Bioch,* 62, 128-135.

SWATLOSKI, R. P., SPEAR, S. K., HOLBREY, J. D. & ROGERS, R. D. (2002) Dissolution of cellose with ionic liquids. *J Am Chem Soc,* 124, 4974-4975.

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*. *Sci Rep,* 3, 1030.

TEERI, T. T. (1997) Crystalline cellulose degradation: New insight into the function of cellobiohydrolases. *Trends Biotechnol,* 15, 160-167.

TRAN, H. T., BARNICH, N. & MIZOGUCHI, E. (2011) Potential role of chitinases and chitin-binding proteins in host-microbial interactions during the development of intestinal inflammation. *Histol Histopathol,* 26, 1453-1464.

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. *Front Plant Sci,* 4, 174.

VÁRNAI, A., SIIKA-AHO, M. & VIIKARI, L. (2013) Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnol Biofuels,* 6, 30.

VU, V. V., BEESON, W. T., PHILLIPS, C. M., CATE, J. H. & MARLETTA, M. A. (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc,* 136, 562-565.

VÄLJAMÄE, P., SILD, V., NUTT, A., PETTERSSON, G. & JOHANSSON, G. (1999) Acid hydrolysis of bacterial cellulose reveals different modes of synergistic action between cellobiohydrolase I and endoglucanase I. *Eur J Biochem,* 266, 327-334.

VAAJE-KOLSTAD, G., BUNÆS, A. C., MATHIESEN, G. & EIJSINK, V. G. H. (2009) The chitinolytic system of *Lactococcus lactis* ssp. *lactis* comprises a nonprocessive chitinase and a chitin-binding protein that promotes the degradation of alpha- and beta-chitin. *FEBS J,* 276, 2402-2415.

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. *J Mol Biol,* 416, 239-254.

VAAJE-KOLSTAD, G., HORN, S. J., SØRLIE, M. & EIJSINK, V. G. H. (2013) The chitinolytic machinery of *Serratia marcescens*-a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS J,* 280, 3028-3049.

VAAJE-KOLSTAD, G., HORN, S. J., VAN AALTEN, D. M., SYNSTAD, B. & EIJSINK, V. G. H. (2005b) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem,* 280, 28492-28497.

VAAJE-KOLSTAD, G., HOUSTON, D. R., RIEMEN, A. H., EIJSINK, V. G. H. & VAN AALTEN, D. M. (2005a) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem,* 280, 11313-11319.

VAAJE-KOLSTAD, G., WESTERENG, B., EIJSINK, V. G. H., HORN, S. J., SØRLIE, M. & FORSBERG, Z. (2011) *Methods of Degrading or Hydrolyzing a Polysaccharide*. PCT/US2011/046838.

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, 219-222.

VÅRUM, K. M., OTTØY, M. H. & SMIDSRØD, O. (1994) Water-Solubility of Partially N-Acetylated Chitosans as a Function of Ph - Effect of Chemical-Composition and Depolymerization. *Carbohyd Polym,* 25, 65-70.

WADA, M., CHANZY, H., NISHIYAMA, Y. & LANGAN, P. (2004) Cellulose IIII crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules,* 37, 8548-8555.

WALTER, S. & SCHREMPF, H. (2008) Characteristics of the surface-located carbohydrate-binding protein CbpC from *Streptomyces coelicolor* A3(2). *Arch Microbiol,* 190, 119-127.

WARNECKE, F., LUGINBUHL, 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. N., HERNANDEZ, 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, 560-U517.

WESTERENG, B., AGGER, J. W., HORN, S. J., VAAJE-KOLSTAD, G., AACHMANN, F. L., STENSTRØM, Y. H. & EIJSINK, V. G. H. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J Chromatogr A,* 1271, 144-152.

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 *Pc*GH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One,* 6, e27807.

WILLIAMS, N. (2008) Biofuel debate deepens. *Curr Biol,* 18, R891-892.

WILSON, D. B. (2009) Cellulases and biofuels. *Curr Opin Biotech,* 20, 295-299.

WING, R. E. & FREER, S. N. (1984) Use of Trifluoroacetic-Acid to Prepare Cellodextrins. *Carbohyd Polym,* 4, 323-333.

WONG, E., VAAJE-KOLSTAD, G., GHOSH, A., HURTADO-GUERRERO, R., KONAREV, P. V., IBRAHIM, A. F., SVERGUN, D. I., EIJSINK, V. G. H., CHATTERJEE, N. S. & VAN AALTEN, D. M. (2012) The *Vibrio cholerae* colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS Pathog,* 8, e1002373.

WOOD, T. M. & GARCIA-CAMPAYO, V. (1990) Enzymology of cellulose degradation. *Biodegradation,* 1, 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*. *J Biol Chem,* 288, 12828-12839.

YAKOVLEV, I., VAAJE-KOLSTAD, G., HIETALA, A. M., STEFANCZYK, E., SOLHEIM, H. & FOSSDAL, C. G. (2012) Substrate-specific transcription of the enigmatic GH61 family of the pathogenic white-rot fungus *Heterobasidion irregulare* during growth on lignocellulose. *Appl Microbiol Biot,* 95, 979-990.

YOSHIZAWA, K., KIHARA, N., KAMACHI, T. & SHIOTA, Y. (2006) Catalytic mechanism of dopamine beta-monooxygenase mediated by Cu(III)-oxo. *Inorg Chem,* 45, 3034-3041.

ZAKARIASSEN, H., EIJSINK, V. G. H. & SØRLIE, M. (2010) Signatures of activation parameters reveal substrate-dependent rate determining steps in polysaccharide turnover by a family 18 chitinase. *Carbohyd Polym,* 81, 14-20.

ZAKARIASSEN, H., AAM, B. B., HORN, S. J., VÅRUM, K. M., SØRLIE, M. & EIJSINK, V. G. H. (2009) Aromatic Residues in the Catalytic Center of Chitinase A from *Serratia marcescens* Affect Processivity, Enzyme Activity, and Biomass Converting Efficiency. *J Biol Chem,* 284, 10610-10617.

ZELTINS, A. & SCHREMPF, H. (1997) Specific interaction of the *Streptomyces* chitin-binding protein CHB1 with alpha-chitin - The role of individual tryptophan residues. *Eur J Biochem,* 246, 557-564.

ZHANG, L., KOAY, M., MAHERT, M. J., XIAO, Z. & WEDD, A. G. (2006) Intermolecular transfer of copper ions from the CopC protein of *Pseudomonas syringae*. Crystal structures of fully loaded (CuCuII)-Cu-I forms. *J Am Chem Soc,* 128, 5834-5850.

ZHANG, Y. H. & 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.

ZHOU, Y., POPE, P. B., LI, S., WEN, B., TAN, F., CHENG, S., CHEN, J., YANG, J., LIU, F., LEI, X., SU, Q., ZHOU, C., ZHAO, J., DONG, X., JIN, T., ZHOU, X., YANG, S., ZHANG, G., YANG, H., WANG, J., YANG, R., EIJSINK, V. G. H. & WANG, J. (2014) Omics-based interpretation of synergism in a soil-derived cellulose-degrading microbial community. *Sci Rep,* revised version submitted for publication.

AACHMANN, F. L., SØRLIE, M., SKJÅK-BRAEK, G., EIJSINK, V. G. H. & VAAJE-KOLSTAD, G. (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc Natl Acad Sci U S A,* 109, 18779-18784.

AAM, B. B., HEGGSET, E. B., NORBERG, A. L., SØRLIE, M., VÅRUM, K. M. & EIJSINK, V. G. H. (2010) Production of chitooligosaccharides and their potential applications in medicine. *Mar Drugs,* 8, 1482-1517.

Paper I

Cleavage of cellulose by a CBM33 protein

Zarah Forsberg, Gustav Vaaje-Kolstad, Bjørge Westereng, Anne C. Bunæs, Yngve Stenstrøm, Alasdair K. Mackenzie, Morten Sørlie, Svein J. Horn and Vincent G. H. Eijsink, *Protein Science*, 2011, 20, 1479-1483.

ACCELERATED COMMUNICATION

Cleavage of cellulose by a CBM33 protein

Zarah Forsberg, Gustav Vaaje-Kolstad,* Bjørge Westereng, Anne C. Bunæs, Yngve Stenstrøm, Alasdair MacKenzie, Morten Sørlie, Svein J. Horn, and Vincent G.H. Eijsink

Department of Chemistry Biotechnology and Food Science, Norwegian University of Life Sciences, Aas, Norway

Received 1 July 2011; Accepted 1 July 2011
DOI: 10.1002/pro.689
Published online 11 July 2011 proteinscience.org

Abstract: Bacterial proteins categorized as family 33 carbohydrate-binding modules (CBM33) were recently shown to cleave crystalline chitin, using a mechanism that involves hydrolysis and oxidation. We show here that some members of the CBM33 family cleave crystalline cellulose as demonstrated by chromatographic and mass spectrometric analyses of soluble products released from Avicel or filter paper on incubation with CelS2, a CBM33-containing protein from *Streptomyces coelicolor* A3(2). These enzymes act synergistically with cellulases and may thus become important tools for efficient conversion of lignocellulosic biomass. Fungal proteins classified as glycoside hydrolase family 61 that are known to act synergistically with cellulases are likely to use a similar mechanism.

Keywords: CBM33; cellulose oxidation; GH61; cellulose degradation

Introduction

For long the biochemically challenging and economically important process of enzymatic cellulose degradation was thought to be achieved by the synergistic action of endo- and exo-acting cellulases.[1,2] Still there have been speculations that other factors may be involved, in particular factors that would make the crystalline and recalcitrant polysaccharide more accessible to the hydrolytic enzymes.[1,3–5] In order for cellulases to act on cellulose chains organized in a crystalline matrix, the enzyme will need to "extract" several consecutive sugars from their crystalline context and bind them in its active site. This is energetically demanding,[6,7] except perhaps for chain ends that may be present in limiting amounts in substrates with high crystallinity.

Studies on the enzymatic depolymerization of chitin, a cellulose analog used as a structural component in, for example, crustaceans, insects and fungi, have shown that, indeed, the concept of synergistically acting endo- and exo-enzymes may be incomplete. Some years ago, it was shown that chitinolytic bacteria produce proteins classified as family 33 carbohydrate-binding modules (CBM33; see Refs. 8,9) and most importantly, these proteins act synergistically with chitinases.[10,11] In a very recent study, it was shown that these proteins in fact are enzymes that cleave chitin chains while still being in their crystalline context, using an unprecedented mechanism that involves a hydrolytic and an oxidative step.[12] Thus, the CBM33 generates two new chain ends on the crystalline surface, one normal

Abbreviations: CBM, carbohydrate-binding module; DP, degree of polymerization; Glc, glucose; GlcA, gluconic acid; GlcLA, gluconolactone; HPAEC, high pressure anion exchange chromatography; MALDI-TOF MS, matrix-assisted laser desorption/ionization time of flight mass spectrometry.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: VISTA Program; Grant number: 6505; Grant sponsor: Norwegian Research Council; Grant number: 190965/S60190877/S60196885/F20.

*Correspondence to: Gustav Vaaje-Kolstad, Department of Chemistry Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Aas, Norway. E-mail: gustko@umb.no

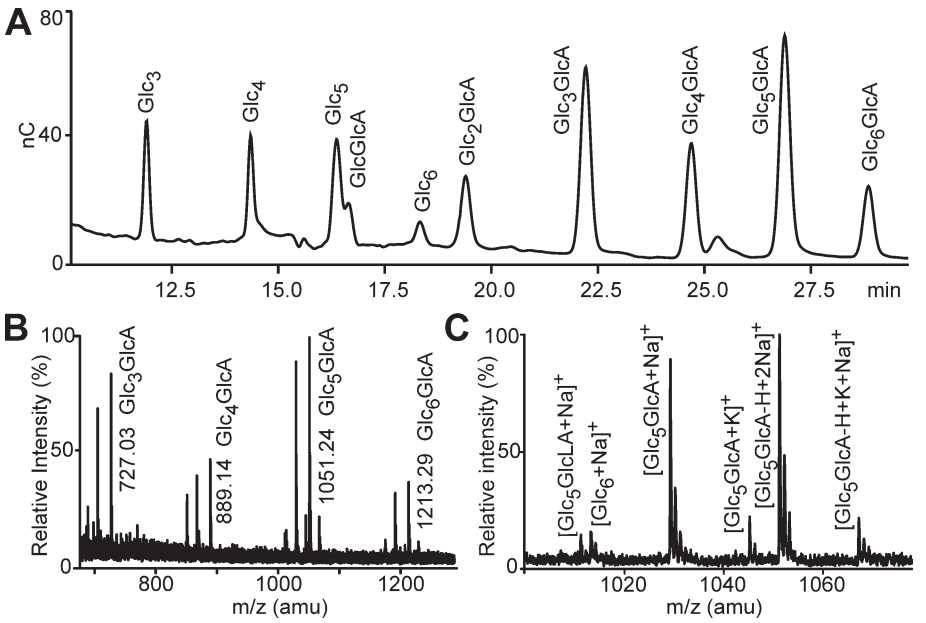


Figure 1. The action of CelS2. Panels A and B show soluble native ($Glc_{3–6}$) and oxidized ($Glc_{2–6}GlcA$) cello-oligosaccharides generated by CelS2 activity on Avicel, as detected by HPAEC (A) and MALDI-TOF MS (B). In panel B, only the major peaks in each oligosaccharide cluster are labeled. Panel (C) shows details of the mass spectrum for Glc_5GlcA. The oxidized oligosaccharide is observed as sodium and potassium adducts, and as sodium and potassium adducts of the oligosaccharide sodium/potassium salts, as is commonly seen for carbohydrates containing carboxylic groups.[19] Observed masses (*m/z*) in the Glc_5GlcA cluster are 1011.22 (Glc_5GlcLA+Na), 1013.25 (Glc_6+Na), 1029.24 (Glc_5GlcA+Na), 1045.21 (Glc_5GlcA+K), 1051.24 (Glc_5GlcA-H+2Na), and 1067.23 (Glc_5GlcA-H+K+Na). GlcLA indicates the lactone form of the oxidized cello-oligosaccharide.

nonreducing and an "oxidized reducing end," that is, an aldonic acid. It was also shown that the activity of CBM33 proteins could be boosted by adding external electron donors such as ascorbic acid.

These observations raise the question whether there exist proteins that act in a similar way on cellulose. Indeed, several bacteria that are able to degrade crystalline cellulose contain multiple CBM33 proteins[13,14] and some of these are known to be coregulated with cellulases.[15,16] Proteins classified as glycoside hydrolase family 61 (GH61) are structurally similar to CBM33[17,18] and are known to act synergistically with cellulases. However, the potentiating mechanism of these proteins remains enigmatic and activity has so far only been shown for rather complex substrates, that is, not for pure cellulose.[17] In this study, we demonstrate that CelS2 from *Streptomyces coelicolor* A3(2), a two-domain protein consisting of a 194-residue CBM33 supplemented with a well-known 99-residue cellulose-binding domain (CBM2), indeed is capable of cellulose cleavage, generating oxidized chain ends and boosting cellulase activity. This is the first time such an activity is experimentally shown.

Results and Discussion

We have cloned and analyzed the functionality of CelS2 (Uniprot ID: Q9RJY2) from *S. coelicolor* A3(2) consisting of a CBM33 domain and a C-terminal cellulose-binding domain classified as a CBM2. Figure 1 shows high pressure anion exchange chromatography (HPAEC) and matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) analyses of soluble oligosaccharides released from Avicel by CelS2 and demonstrates that this protein cleaves crystalline cellulose by a mechanism that leads to the formation of oxidized products. The observed masses [Fig. 1(B,C)] correspond to those for oxidized cello-oligosaccharides and the detection of the lactone form of the oxidized cello-oligosaccharides [two atomic mass units smaller than the native cello-oligosaccharide; Fig. 1(C)] confirms that the CelS2-derived products are aldonic acids (i.e., analogous to the products detected for CBM33 enzymes acting on chitin). The MS data and the chromatogram [Fig. 1(A)] of the oxidized CelS2-generated products also correlated with what was observed for in-house generated cello-oligosaccharide aldonic acids (Glc_nGlcA; Supporting Information Fig. S1). Figure 1(A and C) reveals the presence of both native and oxidized cello-oligosaccharides in the product mixtures. This is likely due to the low degree of polymerization (DP) of Avicel (~100; See Ref. 20), which implies that a soluble native cello-oligosaccharide is generated each time CelS2 cleaves a chain near the reducing end. As the CBM33 enzymatic mechanism includes both a hydrolytic and oxidative step,[12] it is also possible

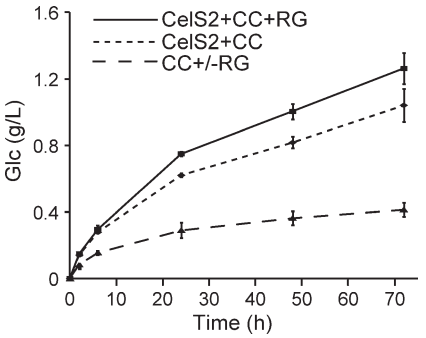


Figure 2. Degradation of high-molecular weight filter paper cellulose by Celluclast (CC), in the presence or absence of CelS2 and reduced glutathione (RG) as shown by the increase of soluble cello-oligosaccharides (Glc and Glc_2; converted to total Glc) over time. Under these conditions, reactions with only CelS2 did not yield detectable amounts of Glc or Glc_2 (not shown). The chitin-active CBM33, CBP21[12] did not affect CC efficiency (not shown). RG had no effect in reactions with only CC; only one of the two overlapping curves is shown. Data are mean ± SD ($N = 3$); error bars indicate SD. See the Materials and Methods section for experimental details.

that CelS2 occasionally generates normal reducing ends, by-passing the oxidative step. This would imply that the hydrolytic step must be the first in the order of events. Clarification of this issue awaits further studies on the catalytic mechanism of CBM33s.

As the solubility of cello-oligosaccharides is low and as these oligosaccharides tend to remain bound to the crystalline substrate, it is difficult to show soluble CelS2-generated products when using substrates with high DP. Indeed, we detected only very minor amounts of soluble products on incubation of filter paper (estimated DP ~2000; see Ref. 21) with CelS2 under the conditions used for producing Figure 1. However, by addition of a "catalytic amount" of cellulase activity, the action of CelS2 on filter paper could be visualized as a "boosting effect" on cellulose degradation, as shown in Figure 2. The data show that the hydrolytic enzymes become more active as they are presented to a more amenable substrate resulting from the action of CelS2. Figure 2 also shows that this boosting effect is increased in the presence of reductants, albeit less vigorously than previously observed for CBM33 proteins acting on chitin.

Interestingly, oxidized soluble products generated by CelS2 [Fig. 1(A and B)] show a dominance of products with an even number of sugars. This may indicate that CelS2 cleavage happens on a well-ordered chain, as would be the case when the chain is in a crystalline context. CBM33s acting on chitin show similar product patterns.[12]

These results reveal a new paradigm for enzymatic cellulose conversion and perhaps for enzymatic conversion of polysaccharides in general. Proteins classified as GH61 seem to be the fungal counterpart of the bacterial CBM33 proteins, as they are structurally similar and have a potentiating effect on cellulases,[17] although their mechanism remains unknown. The fact that such a "cellulase-boosting activity" not only occurs in fungi (GH61) but also in bacteria (CBM33) indicates the importance of this activity in nature. The occurrence of this activity in bacteria also has practical consequences as it is easier to produce the bacterial CBM33s recombinantly, with a cellulase-free background.

One important conclusion from our results is that CBM33 proteins vary with respect to their substrate specificities. In addition, we note that many microorganisms contain multiple CBM33 or GH61 encoding genes. Thus, it seems possible that additional substrate specificities may be discovered in the future. Another important issue for future studies is the role of the divalent metal ion which remains somewhat enigmatic for both CBM33 and GH61 enzymes.[11,17,18] We have used Mg^{2+} in this study because an initial screening showed high activity when this metal was added. However, several metals worked well, confirming the apparent promiscuity that has been observed in other studies of both CBM33s[11] and GH61s.[17] Purified CelS2 retained considerable activity without the addition of metals and it is likely that this activity is due to high affinity binding of another as yet unidentified metal. Addition of ethylenediaminetetraacetic acid (EDTA) inhibited the enzyme. Clearly, more work needs to be done on the role of metals in CelS2, as well as in other CBM33s and in GH61s.

Finally, and most importantly, the present findings have major implications for further development of an efficient cellulose-based biorefinery. CBM33 and GH61 enzymes may turn out to be important tools for achieving more efficient enzymatic conversion of recalcitrant lignocellulosic biomass. Further process optimization work is needed, as it is currently not immediately obvious what enzyme ratios (i.e., hydrolases versus CBM33/GH61) lead to optimal conversion rates and how these ratios depend on the substrate and the presence of reducing power (in the substrate or externally added; Forsberg *et al*., unpublished observations and Refs. 14,17). Still, there is no doubt that this novel class of enzymes will play a major role in the current quest for efficient enzymatic processes for biomass conversion.

Materials and Methods

Cloning, expression, and purification

The gene encoding the mature form of CelS2 (Uniprot ID: Q9RJY2; residues 35-364) from *S. coelicolor* A3(2) was cloned into the pET-32 LIC vector following the instructions provided by the supplier (Novagen). Successful constructs were sequenced for verification and transformed into *E. coli* Rosetta DE(3)

cells that were cultured at 37°C, and induced by 0.1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) at O.D. = 0.6, followed by 20 h culturing at 20°C and finally harvesting by centrifugation. Cell pellets were resuspended in 20 m*M* Tris-HCl pH 8.0, 100 μ*M* phenylmethylsulfonyl fluoride (PMSF), 0.1 mg/mL lysozyme (Sigma), and 1U/mL DNAse (Fluka) and lysed by sonication. Cell debris was removed by centrifugation and CelS2 was purified by standard immobilized metal affinity chromatography (IMAC) purification protocols using the Nickel-NTA IMAC resin (Qiagen). Purified protein was concentrated using Sartorius Vivaspin protein concentration devices with a 10 kDa cutoff. To obtain a native N-terminus, which is crucial for enzyme activity,[12] the pure protein was digested with Factor Xa according to the instructions supplied by the manufacturer (Novagen). The free His-Tag and Factor Xa were removed using IMAC chromatography and Xarrest agarose beads (Novagen), respectively. The buffer of the pure protein was finally changed to 20 m*M* Tris pH 8.0. Processing of the His-Tag and protein purity were verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis. Protein concentrations were quantified using the Bio-Rad Bradford micro assay (Bio-Rad).

Chemical oxidation of cello-oligosaccharides to aldonic acids/lactones

Cello-oligosaccharides were obtained by trifluoroacetic acid (TFA) hydrolysis according to the method described by Wing and Freer[22] and lyophilized. The material was oxidized using a mild oxidation method that has been shown to selectively oxidize the hemiacetal carbon of carbohydrates to generate aldonic acids.[23,24] The oligosaccharide (2.90 g which included an unspecified amount of salt) was suspended in a minimum amount of water (10 mL) and mixed with an iodine solution (7.3 mmol iodine in 15 mL methanol). While stirring, a 4% (w/w) solution of KOH in methanol (48 mL) was added dropwise for ≈15 min. The solution was heated to 40°C for 1 h until the color disappeared. Cooling in the refrigerator overnight yielded a precipitate of white crystals that was filtered and washed with cold methanol. Drying in a desiccator gave 0.96 g of off-white crystals. ^{13}C NMR spectra showed the carboxylic carbon resulting from the oxidation at about 178 ppm (internal), anomeric carbons at 100–105 ppm, and carbinol carbons at 60 – 80 ppm. Signals for the hemiacetal region (91–97 ppm) were low relative to the signal for the internal C1 carbons at 100–105 ppm (i.e., much lower than in the nonoxidized material), confirming that C1 had been oxidized. For assignment of NMR signals, see Ref. 25; for an example of a study showing that the hemiacetal C1 still would show chemical shifts in the 91–97 ppm region if another carbon, such as C6, had been oxidized, see Ref. 26.

Separation and fractionation of cello-oligosaccharide aldonic acids/lactones

Oligosaccharides and their lactones/aldonic acids were separated by porous graphitic carbon chromatography run in reverse phase mode on an Ultimate 3000RSLC (Dionex corp.) with a Hypercarb 10 × 150 mm^2 (5 μm) column (Thermo Scientific), operated at 70° with 5 mL/min flow rate, a 693 μL sample loop, and charged aerosol detection (Corona Ultra, ESA). The following eluents were used: 0.05% (v/v) trifluoroacetic acid (A), acetonitrile with 0.05% (v/v) trifluoroacetic acid (B). Oligosaccharides were eluted using the following gradient; 100% A for 1.8 min, then a linear gradient running for 25.6 min to reach 27.5% B, and finally running 24.4 min to reach 60% B. 60%B was kept for 13.4 min, followed by a rapid change back to initial conditions, which was kept for 16.5 min (column reconditioning). Fractions were collected using a 1:20 custom-made post-column split directing the flow to the detector and the collecting tubes, respectively. The identities of the solutes were verified by MALDI-TOF analysis (using an identical method as described in Ref. 12).

Qualitative analysis of native and oxidized cello-oligosaccharides

Soluble products generated by CelS2 activity on cellulosic substrates were identified by MALDI-TOF MS, using previously published methods,[12] and by HPAEC using a Dionex Bio-LC equipped with a CarboPack PA1 column operated with a flow rate of 0.25 mL/min 0.1*M* NaOH and column temperature of 30°C. Cello-oligosaccharides were eluted by applying a stepwise linear gradient with increasing amounts of NaOAc, going from 0.1*M* NaOH to 0.1*M* NaOH/0.1*M* NaOAc in 10 min, then to 0.1*M* NaOH/0.3*M* NaOAc in 25 min and then to 0.1*M* NaOH/1.0*M* NaOAc in 5 min. Column reconditioning was achieved running initial conditions for 9 min. Eluted oligosaccharides were monitored by PAD detection. Chromatograms were recorded and analyzed using Chromeleon 7.0.

Cellulose degradation experiments

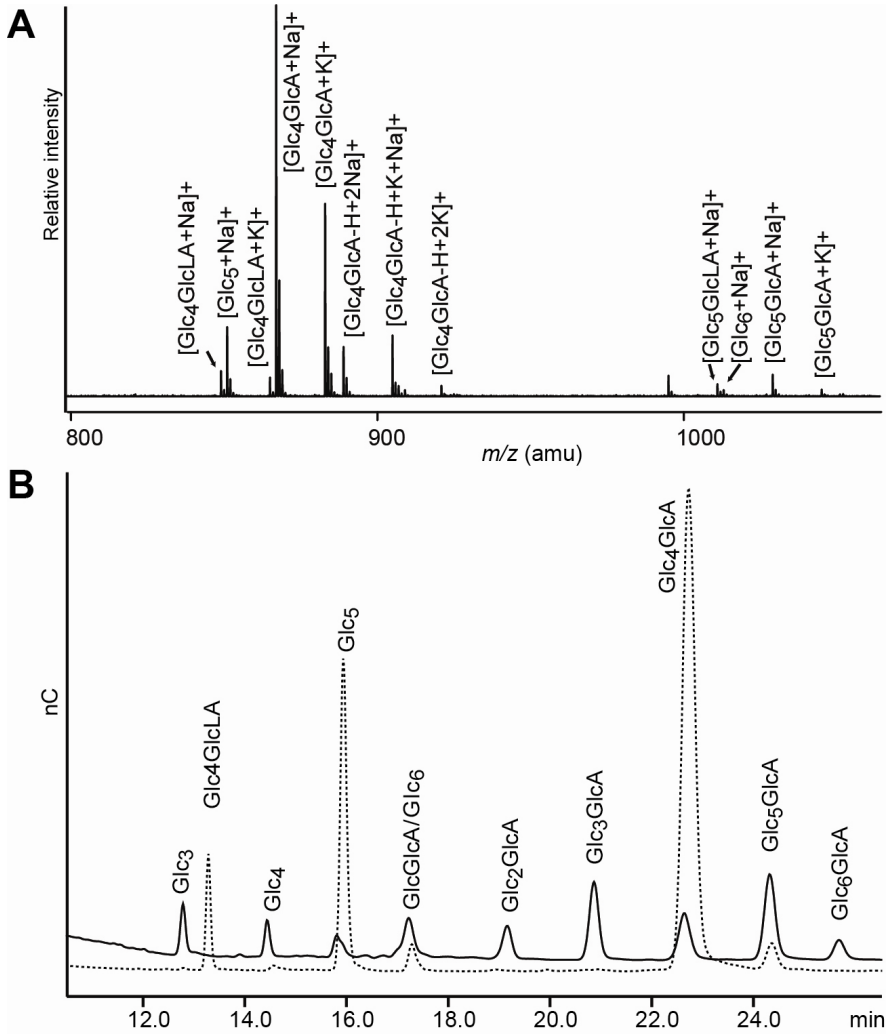
Quantitative assays were performed using 10 mg/mL filter paper (Whatman no.1) in 20 m*M* sodium acetate buffer pH 5.5 in the presence or absence of 40 μg/mL CelS2 and/or 0.8 μg/mL Celluclast (Novozymes). Celluclast is an enzyme cocktail produced by *T. reesei* (Rut C-30) where the dominant enzymes are Cel7A (40–60%), Cel6A (12–20%), Cel7B (5–10%), and Cel5A (1–10%); see Ref. 27. In reactions where no CelS2 was present, 40 μg/mL purified bovine serum albumin (BSA) (NEB) was added to maintain an identical protein load. Qualitative analysis of soluble products generated by CelS2 alone was performed by MALDI-TOF MS or HPAEC using

10 mg/mL Avicel PH-101 (Sigma) as substrate in 25 m*M* Bis-Tris pH 6.5. Reduced glutathione (0.5 m*M*) and ascorbic acid (1.0 m*M*) were used as external electron donors in the quantitative and qualitative experiments, respectively. All reactions contained 1.0 m*M* $MgCl_2$ and were incubated with shaking at 900 rpm at 50°C. Peak assignments in HPAEC were based on the use of purified external standards of native or chemically oxidized cello-oligosaccharides.

References

1. Merino ST, Cherry J (2007) Progress and challenges in enzyme development for biomass utilization. Biofuels 108:95–120.
2. Teeri TT (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. Trends Biotechnol 15:160–167.
3. Din N, Damude HG, Gilkes NR, Miller RC, Warren RA, Jr, Kilburn DG (1994) C1-Cx revisited: intramolecular synergism in a cellulase. Proc Natl Acad Sci USA 91:11383–11387.
4. Eijsink VGH, Vaaje-Kolstad G, Varum KM, Horn SJ (2008) Towards new enzymes for biofuels: lessons from chitinase research, Trends Biotechnol 26:228–235.
5. Reese ET, Siu RGH, Levinson HS (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J Bacteriol 59:485–497.
6. Beckham GT, Crowley MF (2011) Examination of the alpha-chitin structure and decrystallization thermodynamics at the nanoscale. J Phys Chem B 115: 4516–4522.
7. Zakariassen H, Eijsink VGH, Sorlie M (2010) Signatures of activation parameters reveal substrate-dependent rate determining steps in polysaccharide turnover by a family 18 chitinase. Carbohydr Polym 81:14–20.
8. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238.
9. Suzuki K, Suzuki M, Taiyoji M, Nikaidou N, Watanabe T (1998) Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. Biosci Biotechnol Biochem 62:128–135.
10. 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 280:28492–28497.
11. Vaaje-Kolstad G, Houston DR, Riemen AHK, Eijsink VGH, van Aalten DMF (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. J Biol Chem 280:11313–11319.
12. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu ZL, Zhai H, Sorlie M, Eijsink VGH (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science 330:219–222.
13. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). Nature 417:141–147.
14. Moser F,Irwin, Chen SL, Wilson DB (2008) Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. Biotechnol Bioeng 100:1066–1077.
15. Garda AL, Fernandez Abalos JM, Sanchez P, Ruiz Arribas A, Santamaria RI (1997) Two genes encoding an endoglucanase and a cellulose-binding protein are clustered and co-regulated by a TTA codon in Streptomyces halstedii JM8. Biochem J 324:403–411.
16. Ramachandran S, Magnuson TS, Crawford DL (2000) Cloning, sequencing and characterization of two clustered cellulase-encoding genes, celS1 and celS2, from *Streptomyces viridosporus* T7A and their expression in *Escherichia coli*. Actinomycetologica 14:11–16.
17. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, 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 49:3305–3316.
18. 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 angstrom resolution, J Mol Biol 383:144–154.
19. Coenen GJ, Bakx EJ, Verhoef RP, Schols HA, Voragen AGJ (2007) Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. Carbohydr Polym 70:224–235.
20. Mormann W, Michel U (2002) Hydrocelluloses with low degree of polymerisation from liquid ammonia treated cellulose. Carbohydr Polym 50:349–353.
21. Zhang YHP, Lynd LR (2005) Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. Biomacromolecules 6:1510–1515.
22. Wing RE, Freer SN (1984) Use of trifluoroacetic-acid to prepare cellodextrins. Carbohydr Polym 4:323–333.
23. Kobayashi K, Kamiya S, Enomoto N (1996) Amylose-carrying styrene macromonomer and its homo- and copolymers: synthesis via enzyme-catalyzed polymerization and complex formation with iodine. Macromolecules 29:8670–8676.
24. Kobayashi K, Sumitomo H, Ina Y (1985) Synthesis and functions of polystyrene derivatives having pendant oligosaccharides. Polym J 17:567–575.
25. Bubb WA (2003) NMR spectroscopy in the study of carbohydrates: characterizing the structural complexity. Concept Magn Reson A 19A:1–19.
26. Attolino E, Catelani G, D'Andrea F, Puccioni L (2002) Rare and complex saccharides from D-galactose and other milk derived carbohydrates. Part 15. Chemical transformation of lactose into 4-*O*-beta-D-galactopyranosyl-D-glucuronic acid (pseudolactobiouronic acid) and some derivatives thereof. Carbohydr Res 337:991–996.
27. Rosgaard L, Pedersen S, Cherry JR, Harris P, Meyer AS (2006) Efficiency of new fungal cellulase systems in boosting enzymatic degradation of barley straw lignocellulose. Biotechnol Progr 22:493–498.

Supplementary material



Supporting Information Figure 1. Analysis of an in-house generated aldonic acid standard (Glc4GlcA) generated by chemical oxidation of cellooligosaccharides. Note that the efficiency of the oxidation reaction is less than 100% and that some components partially co-elute in the subsequent purification step. Consequently, the Glc4GlcA sample also contains minor amounts of other compounds such as Glc5, Glc6 and Glc5GlnA. (A) Verification of the oxidized oligosaccharide by MALDI-TOF analysis. The mass spectrum shows the K and Na adducts of the lactone (Glc4GlcLA), aldonic acid (Glc4GlcA) and the native oligosaccharide (Glc5). Additionally, the K and Na adducts of the K and Na aldonic acid salts are observed. A distribution of ions representing the oxidized hexameric cellooligosaccharide (Glc5GlcLA/ Glc5GlcA/ Glc6) can also be observed. (B) HPAEC analysis. The picture shows an overlay of a chromatogram for the chemically oxidized oligosaccharide (dotted line) with a chromatogram for products generated by CelS2 action on Avicel (solid line). The chromatogram of the pentameric standard shows the same species as detected by MS analysis: the aldonic acid, the lactone and native form (Nb. Formally, the peak labeled Glc4GlcLA must be regarded as “unknown”; the peak was annotated as Glc4GlcLA by inference from the mass spectrum of panel A). The absence of the lactone form in the CelS2 product mix is due to the slow equilibrium between lactone and aldonic acid at the pH of the reaction (pH 6.5).

Paper II

Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases

Zarah Forsberg, Åsmund K. Røhr, Sophanit Mekasha, K. Kristoffer Andersson, Vincent G. H. Eijsink, Gustav Vaaje-Kolstad and Morten Sørlie, *Biochemistry*, 2014, 53, 1647-1656.

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Comparative Study of Two Chitin-Active and Two Cellulose-Active AA10-Type Lytic Polysaccharide Monooxygenases

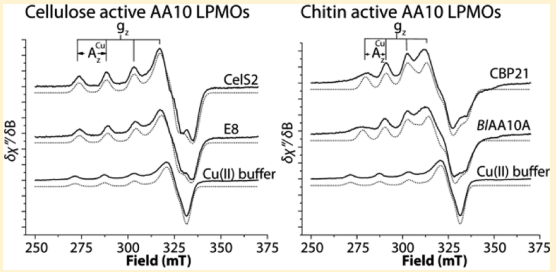
Zarah Forsberg,[†] Åsmund Kjendseth Røhr,[‡] Sophanit Mekasha,[†] K. Kristoffer Andersson,[‡] Vincent G. H. Eijsink,[†] Gustav Vaaje-Kolstad,[†] and Morten Sørlie*[,†]

[†]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway

[‡]Department of Biosciences, University of Oslo, P.O. Box 1066, Blindern, N-0316 Oslo, Norway

Supporting Information

ABSTRACT: Lytic polysaccharide monooxygenases (LPMOs), found in family 9 (previously GH61), family 10 (previously CBM33), and the newly discovered family 11 of auxiliary activities (AA) in the carbohydrate-active enzyme classification system, are copper-dependent enzymes that oxidize sp^3-carbons in recalcitrant polysaccharides such as chitin and cellulose in the presence of an external electron donor. In this study, we describe the activity of two AA10-type LPMOs whose activities have not been described before and we compare in total four different AA10-type LPMOs with the aim of finding possible correlations between their substrate specificities, sequences, and EPR signals. EPR spectra indicate that the electronic environment of the copper varies within the AA10 family even though amino acids directly interacting with the copper atom are identical in all four enzymes. This variation seems to be correlated to substrate specificity and is likely caused by sequence variation in areas that affect substrate binding geometry and/or by variation in a cluster of conserved aromatic residues likely involved in electron transfer. Interestingly, EPR signals for cellulose-active AA10 enzymes were similar to those previously observed for cellulose-active AA9 enzymes. Mutation of the conserved phenylalanine positioned in close proximity to the copper center in AA10-type LPMOs to Tyr (the corresponding residue in most AA9-type LPMOs) or Ala, led to complete or partial inactivation, respectively, while in both cases the ability to bind copper was maintained. Moreover, substrate binding affinity and degradation ability seemed hardly correlated, further emphasizing the crucial role of the active site configuration in determining LPMO functionality.



Lytic polysaccharide monooxygenases (LPMOs) make up a newly discovered group of carbohydrate-active enzymes, that as of 2013 are classified as auxiliary activities (AA) in the CAZy database.[1] Family AA9, previously classified as glycoside hydrolase family 61 (GH61), and family AA10, previously classified as carbohydrate binding module family 33 (CBM33), use an oxidative mechanism to cleave crystalline polysaccharides such as cellulose and chitin.[2−4] Recently, a new family (AA11) active on chitin that appears to have characteristics of both AA9 and AA10 has been discovered.[5] AA10-type proteins occur in all domains of life and are best known from work on bacterial and viral family members, which initially were believed to be chitin binding proteins lacking a catalytic function.[6,7] AA9-type proteins occur in fungi and were originally classified as glycoside hydrolases because of the observation of a low endo activity in cellulose degradation.[8,9] Most LPMO-containing organisms encode several LPMOs that are upregulated when different cellulose or chitin substrates are used as the sole carbon source, indicating that these enzymes are important for the organism's ability to degrade biomass.[10,11]

The first indications of AA10 enzyme activity were published in 2005 and 2008; chitin binding protein 21 (CBP21) from *Serratia marcescens* and E7 and E8 from *Thermobifida fusca*, respectively, were shown to enhance the enzymatic conversion of insoluble polysaccharides.[12,13] In 2010, Vaaje-Kolstad et al.[2] showed that CBP21 is an oxidative enzyme that introduces chain breaks into crystalline chitin in a reaction dependent on a divalent metal ion, an external electron donor, and molecular oxygen. Shortly afterward, it was shown that an AA10[3] as well as AA9-type LPMOs[4,14−16] could cleave and oxidize cellulosic substrates. After some initial confusion about the nature of the divalent metal ion, it has become clear that LPMOs are copper-dependent enzymes.[4,17] In addition to having different substrate specificities, LPMOs also differ in terms of their preference for the position of reaction. All AA10 enzymes described so far oxidize C1, whereas for AA9 enzymes, oxidation has been reported at C1 and C4 or C6.[4,15,16,18]

The CBP21 crystal structure was the first LPMO structure to be determined.[19] This AA10 enzyme has a small and globular structure with a flat binding surface that seems to be adapted to binding to crystalline parts of the chitin substrate primarily through polar interactions. The first AA9-type LPMO structure

Received: January 10, 2014
Revised: February 18, 2014
Published: February 21, 2014

   dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647−1656

to be determined was that of Cel61B from *Hypocrea jecorina* (*Hj*AA9B, CAZy abbreviation).[20] AA9- and AA10-type LPMOs have no significant overall sequence identity but are structurally similar and contain a highly conserved metal binding site in their catalytic centers.[21,22] X-ray crystallography and EPR studies of AA9-type LPMOs have shown that these enzymes contain a type 2 copper binding site where the copper has six ligands. The four ligands in the equatorial plane include nitrogens from two fully conserved histidine side chains and the N-terminal amino group, together forming a T-shaped histidine brace; the fourth equatorial ligand is a water molecule coordinated by a conserved and functionally important glutamine.[4,9,21,22] The axial ligands are the hydroxyl group of a conserved tyrosine (i.e., most likely a tyrosinate) and another water molecule.[4] Available structural information for AA10-type LPMOs also shows or suggests a T-shaped coordination of the metal, but only involving the equatorial sites. The tyrosine residue in AA9-type LPMOs is replaced by phenylalanine in most AA10-type LPMOs; the glutamine coordinating the axial water is absent, and access to this axial site seems to be restricted by a side chain (most commonly an alanine residue).[21] Another dissimilarity between the two LPMO families concerns the N-terminal histidine, which tends to be methylated (Nε-Me) in AA9-type LPMOs expressed in their fungal hosts,[4,23] whereas such methylation has not been observed for AA10-type LPMOs [which were all expressed heterologously (see below)].

Recently, EPR studies of an AA9-type LPMO active on cellulose (*Ta*AA9A, CAZy abbreviation) and an AA10-type LPMO with strong affinity for chitin (*Ba*AA10A, CAZy abbreviation) have revealed differences in the electronic structures of their copper binding sites.[4,21] The AA9-type LPMO displayed a typical type 2 copper EPR signal based on the Peisach−Blumberg classifications of type 1 and 2 copper enzymes. The EPR signal for *Ba*AA10A indicated an intermediate between type 1 and type 2 copper protein centers. On the basis of the overall axial envelope of the EPR spectrum, it was concluded that a type 2 classification is appropriate for members of the AA10 family; the deviation from standard values (and from members of the AA9 family) was ascribed to a distorted coordination geometry, which indeed is visible in available crystal structures.[21] Notably, *Ba*AA10A probably oxidizes C1 in chitin (there are in fact no activity data for this enzyme), whereas *Ta*AA9A oxidizes both C1 and C4 or C6 in cellulose (Quinlan et al.,[4] whether the enzyme oxidizes C4, C6, or both is a matter of debate).

These dissimilarities in the EPR signals could reflect inherent differences that separate LPMOs into two families, AA9 and AA10, but could also relate to differences in substrate specificity, catalytic efficiency, and/or oxidation mode. To address this, we have expressed and analyzed the activity of two AA10-type LPMOs whose oxidative activity had not been described previously, one active on chitin (from *Bacillus licheniformis*, *Bl*AA10A, CAZy abbreviation) and one active on cellulose (E8 from *T. fusca* YX, *Tf*AA10B, CAZy abbreviation). This then allowed us to compare the sequences (mapped on available structures) and EPR spectra of four AA10-type LPMOs, in total, two acting on cellulose and two acting on chitin. The importance of the Phe/Tyr variation close to the copper site mentioned above, which may codetermine differences between members of the AA9 and AA10 families, was addressed by site-directed mutagenesis of Phe219 in one of the cellulose-active AA10 LPMOs, CelS2 (*Sc*AA10C, CAZy abbreviation).

MATERIALS AND METHODS

Cloning, Site-Directed Mutagenesis, and Protein Expression. A DNA fragment encoding CBP21 (residues 1−197) from *S. marcescens* BJL200 (*Sm*AA10A, UniProt entry O83009) was cloned into the pRSET B expression vector (Invitrogen) with its native signal peptide (residues 1−27), as previously described.[19] Full length CelS2 (residues 35−364) from *Streptomyces coelicolor* A3(2) (*Sc*AA10C, UniProt entry Q9RJY2) was cloned into the pET-32 LIC vector (Novagen) as previously described.[3] The N-terminal LPMO domain (residues 35−230) of CelS2, hereafter termed CelS2-N, was amplified from the pET-32 LIC_*cels2* vector and fused into linearized pRSET B using the In-Fusion HD cloning kit (Clontech). The pRSET B vector was precut with restriction endonucleases (BsmI and HindIII) to remove the nucleotides encoding the mature CBP21 protein but preserving its signal sequence for periplasmic expression. The N-terminal LPMO domain (residues 32−225) of E8 from *T. fusca* YX (*Tf*AA10B, UniProt entry Q47PB9), hereafter termed E8-N, was amplified from genomic DNA (ATCC catalog no. BAA-629D-5), as was the LPMO from *B. licheniformis* (*Bl*AA10A, UniProt entry Q62YN7) (residues 32−203) (genomic DNA, ATCC catalog no. 14580D-5). Both of these latter LPMOs were cloned without their natural signal peptide and inserted into the pRSET B expression vector containing the signal sequence of CBP21, as described for CelS2-N above. A general drawing of the constructs and their domains is presented in Figure S1 of the Supporting Information.

The pRSET B_*cels2-n* plasmid was used as a template for site-directed mutagenesis of the conserved phenylalanine residues near the catalytic site. CelS2-N mutants F219A and F219Y were generated using the QuikChange II site-directed mutagenesis kit (Agilent Technologies). After the mutated expression vectors had been verified by DNA sequencing, they were transformed by heat shock into chemically competent One Shot BL21 Star (DE3) cells (Invitrogen).

For protein expression, fresh colonies were inoculated into LB-Amp (Luria broth medium containing 50 μg/mL ampicillin) medium and grown at 30 °C for 20 h at 200 rpm, with the exception of the CBP21-producing strain, which was grown at 37 °C for 16 h. Cells were harvested by centrifugation, and periplasmic fractions were prepared by osmotic shocking.[24] The periplasmic extracts were sterilized by filtration (0.2 μm) and stored at 4 °C prior to protein purification.

Purification and Generation of Apoenzymes. CelS2-N wild type and mutants, E8-N, and *Bl*AA10A were purified using a two-step protocol, starting with an anion exchange chromatography step followed by gel filtration. The periplasmic fraction was adjusted to buffer A [50 mM Tris-HCl (pH 7.5 or 8.5 for the *Bl*AA10A enzyme)] and loaded onto a 5 mL HiTrap DEAE FF column (GE Healthcare) connected to an ÄKTA purifier fast protein liquid chromatography system (GE Healthcare). LPMOs were eluted by applying a linear salt gradient from 100% buffer A [50 mM Tris-HCl (pH 7.5 or 8.5)] to 50% buffer B (buffer A with 1 M NaCl) over 100 min at a flow rate of 3 mL/min (200 min at 4.5 mL/min for *Bl*AA10A). CelS2-N and its mutants and E8-N were eluted at approximately 12% buffer B, whereas *Bl*AA10A eluted at 8% buffer B. The LPMO-containing fractions were pooled and concentrated to <1 mL, using Amicon Ultra centrifugal filters

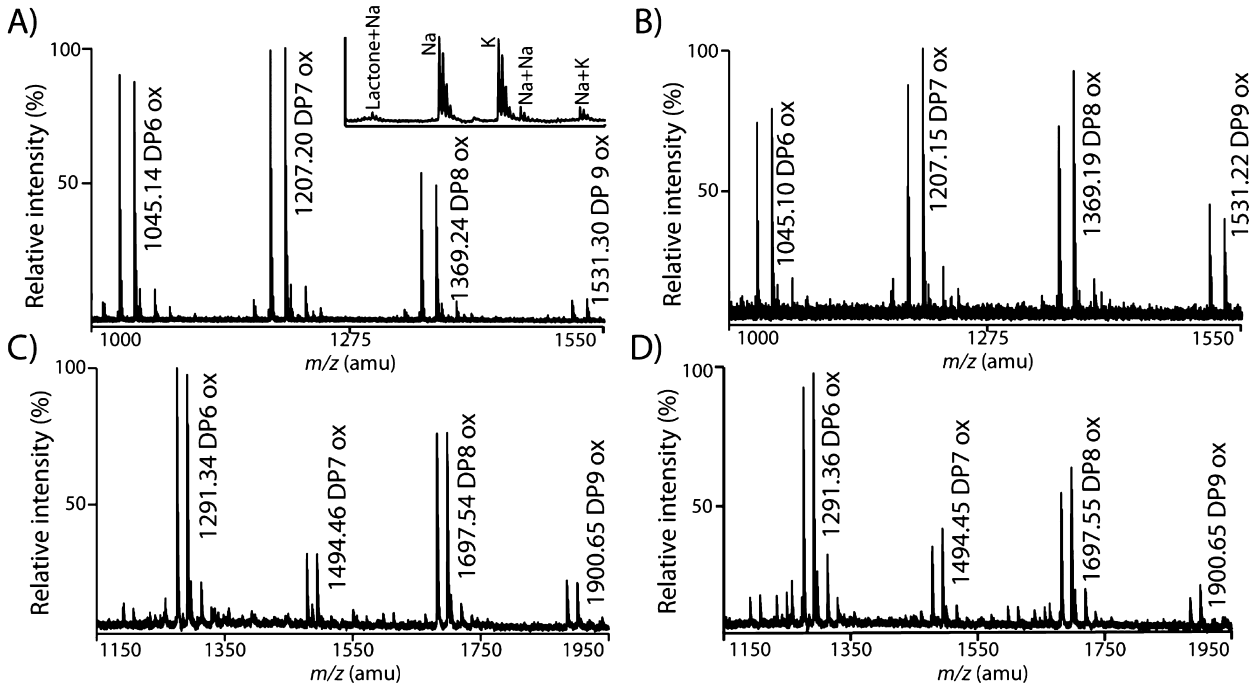


Figure 1. MALDI-TOF MS analysis of LPMO-generated soluble oxidized products. The top two panels show products generated from PASC degradation by CelS2-N (A) and E8-N (B). The bottom two panels show products generated from β-chitin by CBP21 (C) and *Bl*AA10A (D). The annotated masses correspond to the potassium adducts of products in the range from DP6 to DP9 of oxidized cello- or chito-oligosaccharides (Glc_{5-8}Glc1A or $GlcNAc_{5-8}$GlcNAc1A). The sodium adducts, at values m/z 16 lower, yield approximately equally intense signals. The inset in panel A shows the different adducts found in an ionic cluster starting with the sodium adduct of the 1,5 δ-lactone, followed by the sodium and potassium adducts of the aldonic acid, followed by the sodium adducts of the sodium and potassium salts; 100% relative intensity represents 2.2×10^4, 1.5×10^4, 5.6×10^4, and 5.2×10^4 arbitrary units (a.u.) in panels A–D, respectively.

(Millipore) with a molecular mass cutoff of 10000 Da. Subsequently, samples were loaded onto a HiLoad 16/60 Superdex 75 size exclusion column (GE Healthcare), with a running buffer consisting of 50 mM Tris-HCl (pH 7.5) and 200 mM NaCl, using a flow rate of 1 mL/min. All LPMOs eluted approximately 75 min after injection.

CBP21 was purified by chitin affinity chromatography using chitin beads (NEB) equilibrated in 50 mM Tris-HCl (pH 8.0) and 1 M $(NH_4)_2SO_4$ and eluted with 20 mM acetic acid (pH 3.6).[19]

Full length CelS2 from intracellular expression was purified as described previously,[3] using standard immobilized metal affinity chromatography (IMAC) and the nickel-NTA IMAC resin (Qiagen) followed by removal of the His tag using Factor Xa (Novagen). The free His tag and Factor Xa were removed using IMAC chromatography and Xarrest agarose beads (Novagen), respectively.

Protein purity was analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), and fractions containing pure protein were pooled and concentrated using Amicon filters, followed by a change of buffer to 20 mM Tris-HCl (pH 8.0). Protein concentrations were determined using the Bradford assay (Bio-Rad).

Apo-LPMOs were generated by incubating 100 μM enzyme with 400 μM EDTA in 20 mM Pipes buffer (pH 6.5) for 30 min at room temperature. Subsequently, the buffer was changed to 20 mM Chelex 100 resin (Bio-Rad)-treated Pipes buffer (pH 6.5), until the EDTA concentration was diluted and estimated to be <0.1 nM. The protein concentrations were measured using the Bradford assay.

Product Analysis by MALDI-TOF MS and HPAEC. Products formed after incubation of the Cu(II)-saturated LPMOs with various substrates were analyzed using MALDI-TOF MS. Purified LPMOs at a concentration of 1 μM were incubated with 1 mg/mL substrate (phosphoric acid-swollen cellulose, PASC, prepared from Avicel),[25] squid pen β-chitin (France Chitin, Marseille, France), or shrimp shell α-chitin (Hov Bio, Tromsø, Norway) in 20 mM Bis-Tris buffer (pH 6.0) at 37 °C in the presence of 1 mM ascorbic acid. After incubation for 16 h, the reaction mixtures were centrifuged at 16100*g* for 5 min. Samples from the supernatants were mixed with a 9% solution of 2,5-dihydroxybenzoic acid (DHB) matrix in a 1:2 ratio, after which the samples were air-dried and MALDI-TOF MS analysis was performed as described previously.[2]

For semiquantitative analysis of the activity of CelS2-N wild type and mutants, reactions were set up for each enzyme variant using the conditions described above for PASC degradation. After incubation for 16 h, soluble oxidized products were analyzed by HPAEC using a CarboPac PA1 column and a PAD detector as described by Westereng et al.[26]

Analysis of Substrate Binding by CelS2. One microgram of full length CelS2 or CelS2-N was incubated with an excess of substrate [5–15 mg; filter paper (Whatman #1, 50 μm), dried shrimp shell α-chitin or squid pen β-chitin] in 50 μL of 50 mM Bis-Tris (pH 6.2) at room temperature with slow inversion of the tubes. After 3.5 h, the supernatant was removed by centrifugation and the substrate was washed three times with 500 μL of buffer to remove unspecifically bound enzyme. The bound enzyme was released by denaturing with SDS when boiling the substrates in 50 μL of SDS–PAGE sample buffer for

dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647–1656

Table 1. Spin Hamiltonian Parameters[a]

	Cu(II) buffer	CelS2-N[b,c]	E8-N[b]	CBP21	*Bl*AA10A	CelS2-N F219A	CelS2-N F219Y
g_x	2.059	2.015	2.018	2.039	2.038	2.056	2.048
g_y	2.059	2.102	2.103	2.116	2.108	2.058	2.062
g_z	2.270	2.267	2.262	2.260	2.262	2.258	2.248
A_x^{Cu} [d]	12.3	11.7	6.6	42.3	41.3	1.8	23.0
A_y^{Cu} [d]	12.3	17.0	15.3	50.3	48.3	14.0	3.6
A_z^{Cu} [d]	165	153	156	116	125	170	175

[a]Assuming collinear **g** and $\mathbf{A}^{Cu}$ tensors in all simulations. [b]The spectra for CelS2-N and E8-N indicate the presence of free copper, which was corrected by subtracting 30 and 28% of the Cu(II) signal before simulation of the spectra for CelS2-N and E8-N, respectively. [c]The full length CelS2 has EPR parameters identical to those of truncated CelS2-N. [d]Units of $\times 10^{-4}$ cm^{-1}.

10 min. Released proteins, as well as the unbound enzyme fraction and purified enzyme, were then analyzed on an SDS−PAGE gel, using Coomassie Brilliant Blue for staining (Bio-Rad).

Structural Sequence Alignment. PyMod[27] was used to make a structure-based sequence alignment[28] of five AA10-type LPMOs with known structures: *Ef*AA10A (PDB entry 4A02), *Bp*AA10A (PDB entry 3UAM), *Vc*AA10B (PDB entry 2XWX), *Ba*AA10A (PDB entry 2YOX), and CBP21 (PDB entry 2BEM). Chitin activity and binding have been described for all these LPMOs, except for *Bp*AA10A.[19,21,29,30] MUSCLE[31] was then used to add the sequences of the AA10-type LPMOs *Bl*AA10A, *Tf*AA10B, and *Sc*AA10C, which are the subjects of this study.

Electron Paramagnetic Resonance. The EPR spectra were recorded using a BRUKER EleXsyS 560 SuperX instrument equipped with an ER 4122 SHQE SuperX high-sensitivity cavity. Typical settings when recording spectra at 77 K were a microwave power of 0.1−5.0 mW and a modulation amplitude of 10 G (copper quantification purposes) or 30 K a microwave power of 0.5 mW and a modulation amplitude of 5 G when using a liquid helium-cooled Oxford ESR900 cryostat. To estimate the Cu(II) content in the samples, double integrals of baseline-corrected EPR spectra, recorded for the samples and a 100 μM Cu(II) standard in 1 M perchloric acid, were compared.

Apoenzymes were generated as described above and were in Chelex-treated 20 mM Pipes buffer (pH 6.5). EPR studies using the standard conditions described below showed that the apo forms of CelS2-N (enzyme concentration of 108 μM), CBP21 (108 μM), *Bl*AA10A (88 μM), CelS2-N F219A (108 μM), or CelS2-N F219Y (108 μM) did not contain detectable levels of Cu(II). For E8-N (31 μM), approximately 22 μM Cu(II) was detected in the "apoenzyme" (EDTA-treated). Examples of apo-LPMO EPR spectra are shown in Figure S2 of the Supporting Information. Then, Cu(II) was added to the apoenzyme samples to achieve an ∼0.9:1 Cu(II):enzyme ratio. Subsequently, samples were frozen in liquid nitrogen. All Cu(II) added, or initially present and added in the case of E8-N, in the samples could be identified within a margin of error of ∼20%, which is within the expected precision of such analyses. For some of the samples, it was necessary to subtract a fraction of the Cu(II) from the buffer EPR spectrum prior to EPR spectrum simulation (see the simulation parameters listed in Table 1), indicating that not all Cu(II) remained bound to the protein upon the making of the frozen EPR samples. The EasySpin toolbox developed for Matlab was used to simulate and fit EPR spectra.[32]

RESULTS AND DISCUSSION

LPMO Activity Analysis. The four wild-type LPMOs were successfully expressed in *Escherichia coli*. After being purified, stripped of metal ions, and reconstituted with Cu(II), all enzymes showed activity on insoluble polysaccharides [activity measurements on apo-LPMO are not straightforward as it appears that all LPMOs free of metal ions are able to scavenge Cu(II) from the substrate and a large excess of EDTA is required to inhibit oxidation].[4,17,29] The N-terminal LPMO domains of CelS2[3] and E8,[13] called CelS2-N and E8-N, respectively, produced C1-oxidized cello-oligosaccharides from cellulose only (panels A and B of Figure 1, respectively), whereas CBP21[2] and *Bl*AA10A produced C1-oxidized chito-oligosaccharides from α- and β-chitin [panels C and D of Figure 1, respectively (data for α-chitin not shown)], but not from cellulose. To the best of our knowledge, this is the first time that oxidative cleavage of chitin and cellulose is shown for *Bl*AA10A and E8, respectively.

For comparative purposes, full length CelS2 was also produced and purified.[3] Binding studies of CelS2 showed that full length CelS2 binds to filter paper whereas the truncated version, CelS2-N, showed barely detectable binding. For CelS2-N, SDS−PAGE analyses repetitively indicated a minor loss of protein material (compare the filter paper unbound fraction with the control in Figure 2, right panel), which could indicate weak binding, as one would expect considering the detected activity. Remarkably, both full length and truncated CelS2 showed strong binding to both α- and β-chitin. Whereas it is well-known that CBM2s, as present in full length CelS2, bind

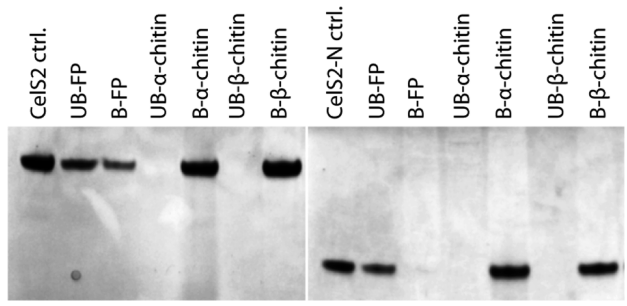


Figure 2. Binding of CelS2 to various substrates. One microgram of full length CelS2 (left) or truncated CelS2-N (right) was incubated with filter paper (FP), α-chitin, or β-chitin for 3.5 h in 20 mM Bis-Tris buffer (pH 6.2) at room temperature. The SDS−PAGE gels document the protein content in the supernatant (unbound, UB) or bound to the substrate (bound, B). The bound protein fraction was released from the substrate by being denatured via boiling in 50 μL of SDS-PAGE sample buffer for 10 min. Purified enzymes were used as a control (ctrl.), not incubated with any substrate, to evaluate protein binding.

dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647−1656

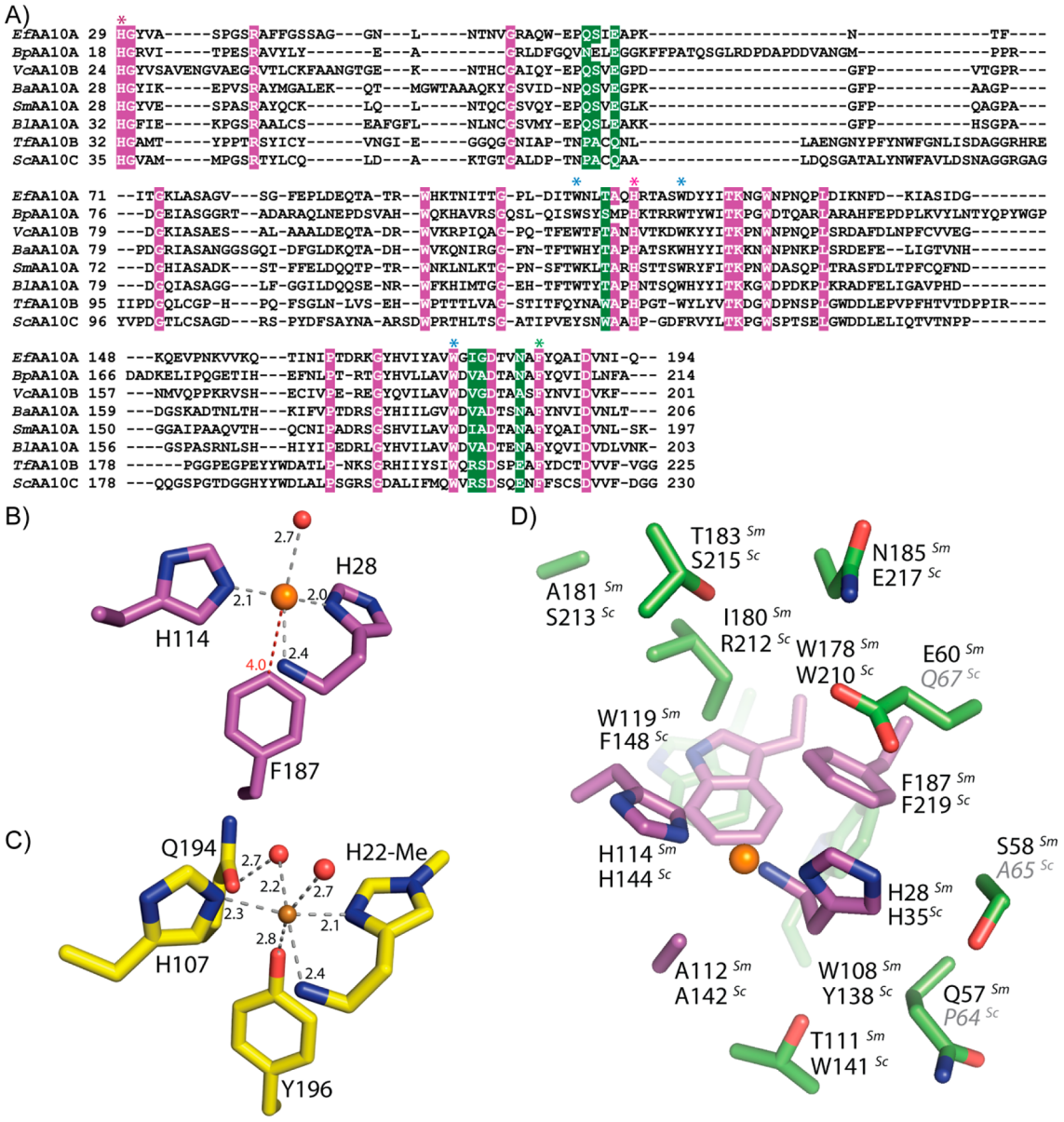


Figure 3. Sequence and structural comparison of LPMO active sites. (A) Structure-based sequence alignment of AA10-type LPMOs. Fully conserved residues are shown with white letters on a pink background; the Cu-coordinating histidines are marked with pink asterisks, and the phenylalanine mutated in CelS2-N (F219A and F219Y) is marked with a green asterisk. The three conserved tryptophans in CBP21 are marked by blue asterisks. Residues highlighted in green are located close to the active site and show variation, potentially determining differences between chitin- and cellulose-active members of the AA10 family (D). (B) Metal coordination in CBP21 representing the AA10-type LPMOs (PDB entry 2BEM).[19] (C) Copper coordination in *Ta*AA9A representing the AA9-type LPMOs (PDB entry 2YET).[4] (D) Expanded view of the active site of CBP21, showing side chains close to the catalytic center that differ between the chitin-active and cellulose-active members of the AA10 family shown in panel A. Pink residues are conserved between all members of the AA10 family in this study, and green residues tend to vary between chitin- and cellulose-active members of the AA10 family. The putative corresponding residues in CelS2 in the unreliable region of the sequence alignment are shown in gray italics. The naming used in the sequence alignment is that from the recommendations of the CAZy database: *Sm*AA10A, CBP21; *Tf*AA10B, E8; *Sc*AA10C, CelS2. Abbreviations in panel D: *Sm*, *Sm*AA10A = CBP21; *Sc*, *Sc*AA10C = CelS2. Figures were made using PyMOL.[38]

to several insoluble polysaccharides, including chitin and cellulose,[33] the strong binding of CelS2-N to α- and β-chitin was unexpected, considering the lack of catalytic activity on these substrates. Upon long incubations (>70 h), we observed barely detectable amounts of soluble products generated from β-chitin by both CelS2-N and E8-N (results not shown). Also, Moser et al.[13] have described experiments indicating that E8 could act synergistically with a chitinase, suggesting some activity on β-chitin. Still, taken together, these results lead to the important conclusion that the ability to bind the substrate and substrate activity and/or specificity are not connected. In other words, the chitin binding demonstrated in Figure 2 is nonproductive. This suggests that other features, such as the geometry of the catalytic center or the character of the redox-active species formed during turnover,[22] could determine which polysaccharides an LPMO can oxidize.

LPMO Sequence Analysis. To provide a structure- and sequence-based basis for interpreting the divergent functions of AA10-type LPMOs, a structure-based alignment of all members of the AA10 family with known structures was made and used

dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647−1656

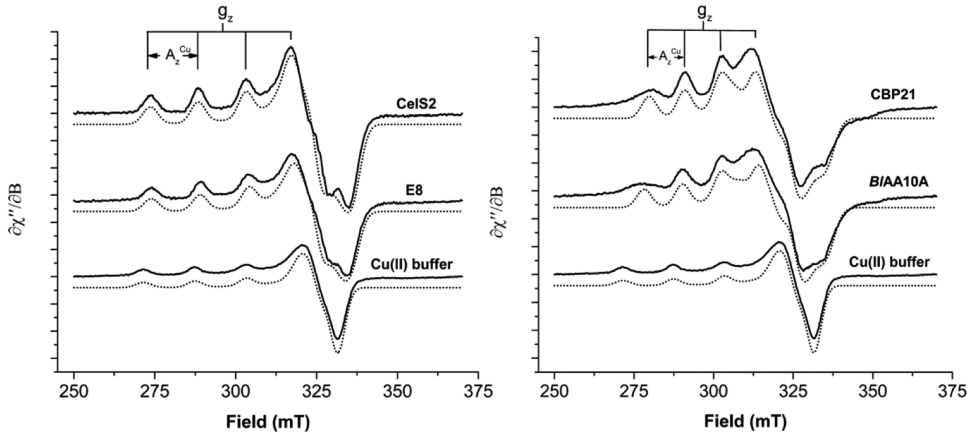


Figure 4. X-Band EPR spectra (—) with simulations (···) for cellulose-oxidizing (left) and chitin-oxidizing (right) AA10-type LPMOs. The superhyperfine splitting occurs at approximately 325 mT. An inset of this for CelS2-N is shown in Figure 6. The spin Hamiltonian parameters in the parallel region are different for chitin- and cellulose-active LPMOs (Table 1). The EPR spectra were recorded at 30 K using a microwave power of 0.5 mW.

to guide alignment of the proteins with unknown structures included in this study (Figure 3A). The level of sequence identity between the cellulose-oxidizing and chitin-oxidizing LPMOs is low (21 and 25% for CelS2-N and CBP21 or *Bl*AA10A, respectively, and 29 and 31% for E8-N and CBP21 or *Bl*AA10A, respectively), whereas the level of sequence identity is ∼50% among LPMOs with the same substrate specificity (46% between the two cellulose-active and 51% between the two chitin-active members of the AA10 family). The metal binding histidines (His28 and His114 in CBP21) are conserved, as are the phenylalanine (Phe187 in CBP21) and the alanine (Ala 112 in CBP21) that are believed to help in shaping the copper binding site in AA10-type LPMOs (Figure 3A,B,D).[21,22]

Notably, several residues close to the copper binding site seem to differ between the cellulose-active and chitin-active LPMOs (Figure 3A,D). A residue that could be important for substrate specificity is Ile180 in CBP21, which is conserved or replaced by Val in other chitin-active members of the AA10 family. This residue lines a pocket on the enzyme surface in the direct proximity of the active site. The two cellulose-active members of the AA10 family contain an arginine at this position.

Another difference between cellulose-active and chitin-active LPMOs that possibly relates to substrate specificity stands out. There are three residues in the N-terminal part of the chitin-active LPMOs, close to the active site (CBP21 residues Gln57, Ser58, and Glu60) that all are solvent-exposed and highly conserved. Because this sequence region is highly diverse (Figure 3A), the sequence alignment is unreliable, and it is thus difficult to predict the corresponding residues in the cellulose-active members of the AA10 family; it is clear though that the sequences vary in this region.

Additional features appearing from this comparison concern the catalytic center itself and the flow of electrons to the copper ion. LPMOs can recruit electrons from a variety of sources, including cellobiose dehydrogenase.[14,15,34] Nothing is known about how electrons are transferred to and through LPMOs. The chitin-active LPMOs contain a fully conserved cluster of three tryptophans that is located directly below the copper binding site (Figure 3D) and that, by analogy to other redox enzymes,[35,36] could play a role in electron transfer.[19,21,22] While the tryptophan residue closest to the metal site (Trp178 in CBP21) is conserved in the two cellulose-active LPMOs, the other two are not (Figure 3A,D). CelS2 has Tyr, Phe, and Trp, whereas E8 has Tyr, Trp, and Trp.

A potentially important difference between the copper binding sites of AA9-type and AA10-type LPMOs concerns the methylation of the N-terminal histidine in the former, a post-transcriptional modification with unknown function. This modification has not been observed in heterologously expressed (and active) LPMOs, including *Phanerochaete chrysosporium* AA9D[37] produced by *Pichia pastoris* and the AA10 LPMOs discussed here, which were all expressed in *E. coli*. To verify the state of the N-terminal histidines in the proteins used for the EPR studies, we analyzed tryptic peptides from recombinantly produced CelS2-N, E8-N, *Bl*AA10A, and CBP21 and found, as expected, that none of the N-terminal tryptic fragments were methylated. Importantly, a similar analysis of CBP21 separated by SDS−PAGE from a culture supernatant of *S. marcescens* grown on chitin did not exhibit methylation of the N-terminal histidine (results not shown).

EPR Spectroscopy. The resting EPR spectra of the Cu(II)-charged AA10 LPMOs showed differences between the two cellulose-degrading enzymes, CelS2-N and E8-N, and the two chitin-degrading enzymes, CBP21 and *Bl*AA10A (Figure 4). The EPR spectra of these enzymes have been simulated, and the estimated spin Hamiltonian parameters are summarized in Table 1. The $\mathbf{g}$ and $\mathbf{A}^{Cu}$ tensors reflect the active site copper coordination environment. The simulations indicated slightly rhombic $\mathbf{g}$ tensors for all enzymes, with CelS2-N and E8-N displaying a higher degree of rhombicity (i.e., a larger difference between g_x and g_y) than CBP21 and *Bl*AA10A. The $g_{x,y}$ and $A_{x,y}^{Cu}$ parameters could not be simulated with the same accuracy as g_z and A_z^{Cu} because the EPR lines in the high-field region are broad and overlap. This is in accordance with what was observed in the work of Hemsworth et al.[21] for *Ba*AA10A, a chitin binding AA10 from *Bacillus amyloliquefaciens*.

By plotting A_z values versus g_z values for several types of copper-containing proteins and model compounds, Peisach and Blumberg observed clear grouping of type 1 and 2 copper proteins.[39] The cellulose-oxidizing *Ta*AA9A (g_z = 2.27, and $A_z^{Cu} = 162 \times 10^{-4}$ cm^{-1})[4] falls nicely into the group of typical type 2 proteins, as do CelS2-N and E8-N (Table 1; g_z and A_z^{Cu} values are very similar to those found for *Ta*AA9A). Interestingly, the A_z^{Cu} hyperfine splitting for the chitin-

dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647−1656

oxidizing LPMOs is substantially smaller than for the cellulose-oxidizing LPMOs (Table 1; values of 116 and 125 × 10^{-4} cm^{-1} vs 153 and 156 × 10^{-4} cm^{-1} for the cellulose-active LPMOs). This places the two chitin-oxidizing LPMOs between the usual Peisach−Blumberg type 1 and type 2 classifications. The same was observed for *Ba*AA10A (g_z = 2.25, and A_z^{Cu} = 135 × 10^{-4} cm^{-1}),[21] which has a sequence 49% identical to that of CBP21 (Figure 3A). Hemsworth et al.[21] noted that although *Ba*AA10A falls between the usual Peisach−Blumberg classifications of type 1 and type 2 copper enzymes, the overall axial envelope of the EPR signal would suggest that a type 2 classification is appropriate.[21,39]

Another difference between the cellulose and chitin-active enzymes is that the CelS2-N and E8-N EPR spectra show superhyperfine splitting on the order of 11−16 × 10^{-4} cm^{-1} that is likely to originate from at least two nitrogen atoms coordinating the copper atom (an example is shown for CelS2-N in Figure 4, left). Such superhyperfine splitting is virtually absent from the spectra for the chitin-active LPMOs (Figure 4, right), and this can indicate differences in active site *g* strain and *A* strain. The physical origin of *g* and *A* strain lies in distributions of spin Hamiltonian parameters (i.e., the orientation of the paramagnetic centers) caused by the structural flexibility of the active site. Because cellulose-active LPMOs have less *g* and *A* strain than the chitin-active enzymes, they are more likely to have more structurally defined Cu(II) environments.

It has previously been suggested that the observed differences in the electronic structures of *Ta*AA9A[4] and *Ba*AA10A[21] are due to inherent differences between AA9-type and AA10-type LPMOs.[21] Our data suggest that these previously observed differences are not family-dependent. Instead, they may relate to substrate specificity and/or other aspects of the catalytic features of the enzymes (see below for further discussion).

Site-Directed Mutagenesis of CelS2-N. All four AA10-type LPMOs investigated have a phenylalanine reaching out from the enzyme core toward the copper ion, independent of the substrate preference. While this is a conserved residue in AA10-type LPMOs, it tends to be a tyrosine in AA9-type LPMOs, including *Ta*AA9A [Tyr196 (Figure 3C)], though it should be mentioned that a minor number of members of the AA10 family have a tyrosine and those of the AA9 family have a phenylalanine. To gain insight into the importance of this residue, CelS2-N mutants F219A and F219Y were generated [Phe219 corresponds to Phe187 in CBP21 (shown in Figure 3B)]. The mutants were expressed at levels comparable to that of the wild type and purified in the same way. The activity of the mutants was analyzed by MALDI-TOF MS and HPAEC. The F219Y mutant showed no activity, whereas the F219A mutant showed reduced activity (Figure 5). The amount of soluble products obtained upon PASC degradation suggested that the F219A mutant has a relative activity of approximately 15% compared to that of the wild-type enzyme. These data show that the aromatic ring of Phe219 is not essential for LPMO activity, although it does contribute. It is interesting that the mutation of Phe219 to Tyr, which is present in almost all members of the AA9 family at this position (Figure 3C), was detrimental to activity.

We also investigated copper coordination for the CelS2-N mutants using EPR. The overall EPR envelopes (Figure 6) for the F219 mutants are similar to that of the wild type, with lower g_z and higher A_z^{Cu} values. The data clearly show that the copper site remains intact in both mutants, although with minor

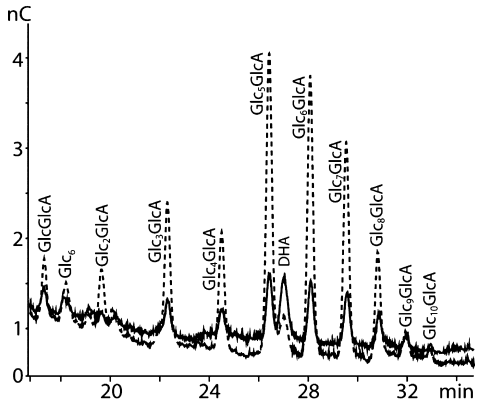


Figure 5. HPAEC chromatograms of oxidized products generated by variants of CelS2-N. Products were analyzed after incubation of CelS2-N wild type or CelS2-N F219A with PASC in 20 mM Bis-Tris (pH 6.0) at 37 °C for 16 h in the presence of 1 mM ascorbate: (···) CelS2-N wild type and (—) CelS2-N F219A mutant. The peak labeled Glc_6 is the result of a cleavage near the reducing end of the polymer. Such sugar moieties will not contain any oxidized functional groups when a C1-oxidizing LPMO interacts with cellulose. Shorter oligomers were also observed prior to the 16 min elution time but were omitted from the figure for the sake of clarity. The peak labeled DHA eluting right after Glc_5GlcA corresponds to dehydroascorbate.

modifications (see values in Table 1). For inactive CelS2-N F219Y, the superhyperfine splitting is even more prominent than for wild-type CelS2-N (Figure 6).

CONCLUDING REMARKS

In this study, we demonstrate the activity of two members of the AA10 family, whose catalytic action has not previously been described, *Bl*AA10A, which oxidizes chitin substrates, and the catalytic LPMO domain of E8 (*Tf*AA10B), which oxidizes cellulose. This allowed us to conduct a comparative study of four AA10-type LPMOs, two acting on chitin and two acting on cellulose.

Substrate binding studies with CelS2 and CelS2-N yielded remarkable results in which both showed strong binding to chitin, on which the enzymes are not active, and weaker and almost no binding to cellulose, on which the enzymes are active. Notably, when LPMO functionality is being assessed, it is important to realize that functional studies normally are conducted with non-natural (i.e., heavily processed) substrates that may be both heterogeneous and structurally dissimilar from "real life". In any case, the data for CelS2 clearly show that differences in the ability to bind certain substrates cannot explain differences in the substrate specificity of catalytic action. All in all, our data suggest that variation in the substrate specificity of LPMOs is determined by the copper active site configuration (because we did observe differences by EPR) and/or the geometry of the enzyme−substrate interaction. Both are conceivably affected by sequence variations directly adjacent to the copper binding site, as discussed above. There are major variations in residues possibly involved in determining substrate specificity (Figure 3D), and there are variations in a cluster of conserved aromatic residues, including Trp178/210 and possibly extending to Phe187/219 (Figure 3D), that may be involved in electron transfer.[21]

The EPR data show that there is considerable variation in the copper-coordinating environments of AA10-type LPMOs. Importantly, the EPR spectra of cellulose-active AA10-type

dx.doi.org/10.1021/bi5000433 | *Biochemistry* 2014, 53, 1647−1656

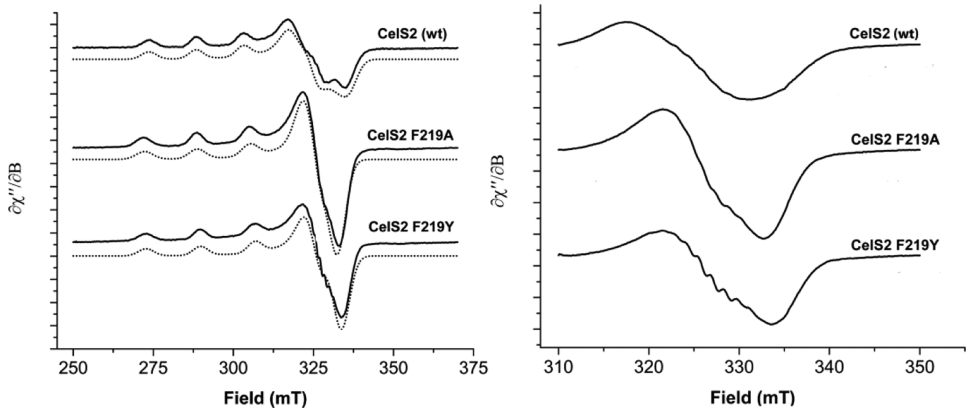


Figure 6. X-Band EPR spectra (—) with simulations (···) for CelS2-N WT, F219A, and F219Y (left). The right panel shows close-ups of the superhyperfine splitting. The EPR spectra were recorded at 30 K using a microwave power of 0.5 mW.

LPMOs are very similar to the EPR spectrum of the cellulose-active AA9-type LPMO *Ta*AA9A,[4] despite the fact that this AA9 enzyme shows some structural differences in or close to the copper binding site. First, the N-terminal copper-coordinating histidine is methylated, which is not the case in the AA10 LPMOs. Second, a conserved Phe in AA10 LPMOs, with its Cζ atom located 4.0 Å from the metal ion in CBP21 (Figure 3B), is replaced by a Tyr in most AA9-type enzymes, with its hydroxyl group 2.8 Å from the metal ion (Figure 3C). The tyrosine residue [Tyr196 in *Ta*AA9A (Figure 3C)] may not be as important as anticipated, because the distance between the tyrosine/tyrosinate in *Ta*AA9A is longer (2.8 Å) than the average 2.38 Å Cu−OTyr bond.[40] Moreover, while mutation of Phe219 to Tyr in CelS2 inactivated the enzyme, the EPR spectrum showed that the copper site remained intact, although with minor modifications. There is no crystal structure for CelS2 or any other cellulose-active AA10 to conclude if there is sufficient space for the added hydroxyl group, which may be the reason why CelS2-N F219Y is inactive. Other mutations (Phe219Ala in this study and Tyr153Phe in *Tt*AA9E described by Harris et al.[9]) indicate that neither the Phe in AA10-type LPMOs nor the Tyr in AA9-type LPMOs is absolutely necessary for activity.

While our data show that there is a correlation between EPR observable variation in the copper binding sites and variation in LPMO functionality, more work needs to be done before the structural basis of LPMO substrate specificity and oxidative power can be unraveled in detail.

As stated by Hemsworth et al.,[21,22] there are clear differences between the active sites of AA10 and AA9. One such difference concerns the steric congestion caused by conserved hydrophobic amino acid sides (Ala112 and Phe187 in CBP21), possibly forcing the substrate to bind solely in the equatorial positions. Our results suggest that the presence of these residues does not determine substrate specificity, which rather seems to be determined by variation at positions that are slightly more remote from the catalytic center. One would certainly expect the structural differences in the copper binding sites themselves to have effects, but these apparently cannot be detected by EPR (or in terms of variation in substrate specificity). Several monomeric copper−oxygen species are known and proposed to be capable of sp^3 carbon oxidation,[22,41] and it is possible that the AA9 and AA10 enzymes differ in terms of how they are generated. Recently, Kim et al. showed, using quantum mechanical calculations, that oxygen binds end-on (η^1) to copper and that a copper-oxyl-mediated oxygen rebound mechanism is energetically preferred on an AA9-type LPMO.[42] It would be interesting to undertake same analyses of the oxygen activation process for members of the AA10 family with varying substrate specificities and compare C1-oxidizing members of the AA9 and AA10 families with similar substrate specificities.

■ ASSOCIATED CONTENT

Supporting Information

A general drawing of the constructs and the domain architecture of LPMOs in this study (Figure S1) and X-band EPR spectra for apo and Cu(II)-charged CelS2 and CBP21 (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: morten.sorlie@nmbu.no. Telephone: +47 64965902. Fax: +47 64965901.

Funding

This work was primarily funded by the VISTA program of the Norwegian Academy of Science and Letters, Grant 6505, and by contributions from Norwegian Research Council Projects 196885, 214138, 209335, 218412, and 214239.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Morten Skaugen for help with mass spectrometry and Alasdair Mackenzie for help with protein purification.

■ ABBREVIATIONS

AA, auxiliary activities; *Ba*, *B. amyloliquefaciens*; *Bl*, *B. licheniformis*; CBM, carbohydrate binding module; EDTA, ethylenediaminetetraacetic acid; EPR, electron paramagnetic resonance; Glc1A, gluconic acid; GlcNAc1A, gluconic acid of *N*-acetylglucosamine; HPAEC, high-performance anion exchange chromatography; LPMO, lytic polysaccharide monooxygenase; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; PASC, phosphoric acid-swollen cellulose; PDB, Protein Data Bank; *Sc*, *S. coelicolor*; *Sm*, *S. marcescens*; *Ta*, *Thermoascus aurantiacus*; *Tf*, *T. fusca*.

REFERENCES

(1) Levasseur, A., Drula, E., Lombard, V., Coutinho, P., and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels 6*, 41.

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

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

(4) Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jorgensen, 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.

(5) Hemsworth, G. R., Henrissat, B., Davies, G. J., and Walton, P. H. (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat. Chem. Biol. 10*, 122−126.

(6) Suzuki, K., Suzuki, M., Taiyoji, M., Nikaidou, N., and Watanabe, T. (1998) Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosci., Biotechnol., Biochem. 62*, 128−135.

(7) Zeltins, A., and Schrempf, H. (1997) Specific interaction of the *Streptomyces* chitin-binding protein CHB1 with α-chitin: The role of individual tryptophan residues. *Eur. J. Biochem. 246*, 557−564.

(8) Karlsson, J., Saloheimo, M., Siika-aho, M., Tenkanen, M., Penttila, M., and Tjerneld, F. (2001) Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur. J. Biochem. 268*, 6498−6507.

(9) 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.

(10) Adav, S. S., Ng, C. S., Arulmani, M., and Sze, S. K. (2010) Quantitative iTRAQ Secretome Analysis of Cellulolytic *Thermobifida fusca*. *J. Proteome Res. 9*, 3016−3024.

(11) Takasuka, T. E., Book, A. J., Lewin, G. R., Currie, C. R., and Fox, B. G. (2013) Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Sci. Rep. 3*, 1030.

(12) Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M. F., Synstad, B., and Eijsink, V. G. H. (2005) The Non-catalytic Chitin-binding Protein CBP21 from *Serratia marcescens* Is Essential for Chitin Degradation. *J. Biol. Chem. 280*, 28492−28497.

(13) Moser, F., Irwin, D., Chen, S. L., and Wilson, D. B. (2008) Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. *Biotechnol. Bioeng. 100*, 1066−1077.

(14) 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.

(15) 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.

(16) Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G. H., 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.

(17) Aachmann, F. L., Sørlie, M., Skjåk-Bræk, G., Eijsink, V. G. H., 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.

(18) 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.

(19) Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H., Eijsink, V. G. H., and van Aalten, D. M. (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J. Biol. Chem. 280*, 11313−11319.

(20) Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C., and Sandgren, M. (2008) The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J. Mol. Biol. 383*, 144−154.

(21) 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.

(22) 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.

(23) Li, X., Beeson, W. T., IV, 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.

(24) Manoil, C., and Beckwith, J. (1986) A Genetic Approach to Analyzing Membrane-Protein Topology. *Science 233*, 1403−1408.

(25) Wood, T. M. (1988) Preparation of Crystalline, Amorphous, and Dyed Cellulase Substrates. *Methods Enzymol. 160*, 19−25.

(26) Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrøm, Y. H., and Eijsink, V. G. H. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J. Chromatogr., A 1271*, 144−152.

(27) Bramucci, E., Paiardini, A., Bossa, F., and Pascarella, S. (2012) PyMod: Sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinf. 13* (Suppl. 4), S2.

(28) Shindyalov, I. N., and Bourne, P. E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng. 11*, 739−747.

(29) Vaaje-Kolstad, G., Bøhle, L. A., Gåseidnes, S., Dalhus, B., Bjørås, M., Mathiesen, G., and Eijsink, V. G. H. (2012) Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J. Mol. Biol. 416*, 239−254.

(30) 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., and van Aalten, D. M. F. (2012) The *Vibrio cholerae* Colonization Factor GbpA Possesses a Modular Structure that Governs Binding to Different Host Surfaces. *PLoS Pathog. 8*, e1002373.

(31) Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32*, 1792−1797.

(32) Stoll, S., and Schweiger, A. (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson. 178*, 42−55.

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

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

(35) Gray, H. B., and Winkler, J. R. (2005) Long-range electron transfer. *Proc. Natl. Acad. Sci. U.S.A. 102*, 3534−3539.

(36) Gray, H. B., and Winkler, J. R. (2010) Electron flow through metalloproteins. *Biochim. Biophys. Acta 1797*, 1563−1572.

(37) 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., 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.
(38) DeLano, W. L., and Lam, J. W. (2005) PyMOL: A communications tool for computational models. *Abstracts of Papers of the American Chemical Society 230*, U1371−U1372.
(39) Peisach, J., and Blumberg, W. E. (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins. *Arch. Biochem. Biophys. 165*, 691−708.
(40) Abriata, L. A. (2012) Analysis of copper-ligand bond lengths in X-ray structures of different types of copper sites in proteins. *Acta Crystallogr. D68*, 1223−1231.
(41) Himes, R. A., and Karlin, K. D. (2009) Copper-dioxygen complex mediated C-H bond oxygenation: Relevance for particulate methane monooxygenase (pMMO). *Curr. Opin. Chem. Biol. 13*, 119−131.
(42) 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.

Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases

Zarah Forsberg,[†] Åsmund Kjendseth Røhr,[‡] Sophanit Mekasha, [†] K. Kristoffer Andersson,[‡] Vincent G.H. Eijsink, [†] Gustav Vaaje-Kolstad,[†] and Morten Sørlie*[†]

Supporting Information

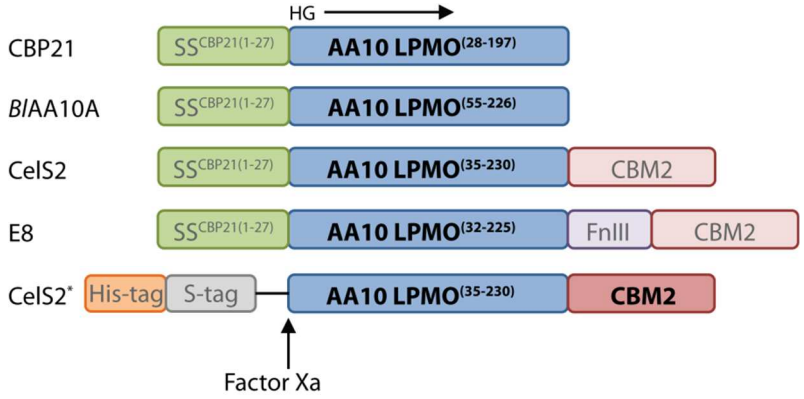


Figure S1. General drawing of the constructs and the domain architecture of the LPMOs in this study. CBP21 and *Bl*AA10A are single domain proteins whereas CelS2 and E8 have a C-terminal family 2 carbohydrate-binding module (CBM2). The catalytic LPMO domain (blue) of all enzymes were cloned into the pRSET B expression vector using the native signal sequence from CBP21 (green) for periplasmic expression.[1] The expressed proteins are marked with bold letters. Full length CelS2* was cloned into the pET-32 LIC system with an N-terminal his-tag that was removed using Factor Xa after IMAC purification as described previously.[2]

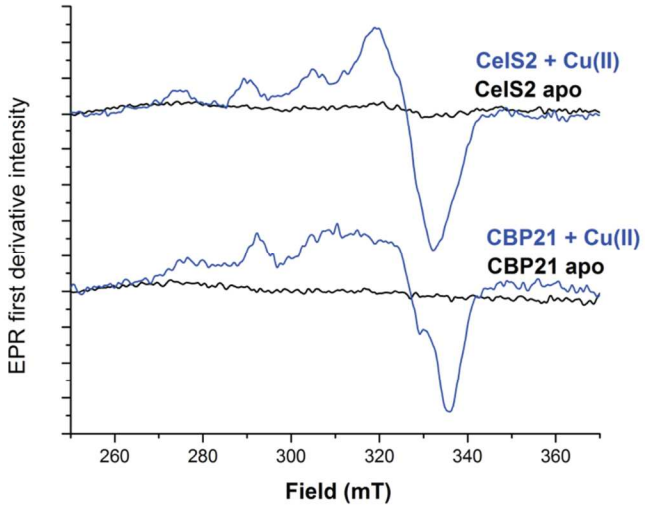


Figure S2. X-band EPR spectra for cellulose oxidizing apo (black) and Cu(II)-charged (blue) CelS2 (top) and chitin oxidizing apo (black) and Cu(II)-charged blue CBP21 (bottom). The EPR spectra were recorded at 77 K using a microwave power of 1 mW for apo-samples and 0.5 mW.

REFERENCES

(1) Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H., Eijsink, V. G. H., and van Aalten, D. M. (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J. Biol. Chem.* 280, 11313-11319.

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

Paper III

Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases

Zarah Forsberg, Alasdair K. Mackenzie, Morten Sørlie, Åsmund K. Røhr, Ronny Helland, Andrew S. Arvai, Gustav Vaaje-Kolstad and Vincent G. H. Eijsink, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, published ahead of print May 27, 2014, doi:10.1073/pnas.1402771111.

Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases

Zarah Forsberg[a], Alasdair K. Mackenzie[a], Morten Sørlie[a], Åsmund K. Røhr[b], Ronny Helland[c], Andrew S. Arvai[d], Gustav Vaaje-Kolstad[a], and Vincent G. H. Eijsink[a,1]

[a]Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, N-1432 Aas, Norway; [b]Department of Biosciences, University of Oslo, N-0316 Oslo, Norway; [c]Department of Chemistry, Norwegian Structural Biology Centre, Faculty of Science and Technology, University of Tromsø, 9019 Tromsø, Norway; and [d]Department of Molecular Biology, Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037

Edited* by Arnold L. Demain, Drew University, Madison, NJ, and approved April 28, 2014 (received for review February 14, 2014)

For decades, the enzymatic conversion of cellulose was thought to rely on the synergistic action of hydrolytic enzymes, but recent work has shown that lytic polysaccharide monooxygenases (LPMOs) are important contributors to this process. We describe the structural and functional characterization of two functionally coupled cellulose-active LPMOs belonging to auxiliary activity family 10 (AA10) that commonly occur in cellulolytic bacteria. One of these LPMOs cleaves glycosidic bonds by oxidation of the C1 carbon, whereas the other can oxidize both C1 and C4. We thus demonstrate that C4 oxidation is not confined to fungal AA9-type LPMOs. X-ray crystallographic structures were obtained for the enzyme pair from *Streptomyces coelicolor*, solved at 1.3 Å (*Sc*LPMO10B) and 1.5 Å (CelS2 or *Sc*LPMO10C) resolution. Structural comparisons revealed differences in active site architecture that could relate to the ability to oxidize C4 (and that also seem to apply to AA9-type LPMOs). Despite variation in active site architecture, the two enzymes exhibited similar affinities for Cu^{2+} (12–31 nM), redox potentials (242 and 251 mV), and electron paramagnetic resonance spectra, with only the latter clearly different from those of chitin-active AA10-type LPMOs. We conclude that substrate specificity depends not on copper site architecture, but rather on variation in substrate binding and orientation. During cellulose degradation, the members of this LPMO pair act in synergy, indicating different functional roles and providing a rationale for the abundance of these enzymes in biomass-degrading organisms.

GH61 | CBM33

The enzymatic conversion of plant biomass is an issue of major scientific and commercial interest. Although this process originally was thought to involve only hydrolytic enzymes, such as cellulases, we now know that oxidative enzymes called lytic polysaccharide monooxygenases (LPMOs) play an important role (1). Using powerful oxidative chemistry (2, 3), LPMOs cleave glycosidic bonds in polysaccharides that are inaccessible to cleavage by hydrolytic enzymes, such as endoglucanases and cellobiohydrolases. By increasing substrate accessibility, LPMOs boost the overall efficiency of enzymatic degradation of insoluble polysaccharides (4–6). Indeed, the latest generation of commercial cellulase mixtures for processing of lignocellulosic biomass benefits from the presence of LPMOs (7).

In the CAZy database (8), LPMOs are classified into the auxiliary activity (AA) families 9 (AA9; previously known as GH61), 10 (AA10; previously known as CBM33), and 11 (AA11) (9). Families AA9 and AA11 comprise fungal enzymes, whereas family AA10 comprises enzymes from all domains of life. Hereinafter, these three families are referred to as LPMO9, LPMO10, and LPMO11, respectively. Members of these families share low sequence identity, but have similar Ig-like folds with a flat substrate-binding surface. The solvent-exposed active site contains two histidines that coordinate a copper ion in a histidine brace (2, 10, 11). The role of the copper ion is to reduce dioxygen, which requires electrons from an external electron donor. The reduced dioxygen likely abstracts a hydrogen from the substrate, which eventually leads to cleavage of the β-1,4 glycosidic linkage (12–14). As a result of this reaction, a carbon in the scissile glycosidic bond is oxidized. Some LPMOs exclusively oxidize C1, others exclusively oxidize C4, and a third group can oxidize either C1 or C4 (13, 15). The latter LPMO group can generate "double-oxidized" products that are formed when a polysaccharide chain is cleaved twice, once with C1 oxidation and once with C4 oxidation (13). The reaction products are either δ-1,5 lactones, which may be hydrated to form aldonic acids (C1 oxidation) (5, 16), or 4-ketoaldoses, which may be hydrated to form gemdiols (C4 oxidation) (17). Notably, the existence of other types of non-C1 oxidation has been suggested (10, 12). Based on previous studies (15, 17) and experimental evidence presented here, and for the sake of simplicity, only C4 oxidation is considered hereinafter.

The abundance of LPMOs in the genomes of biomass-degrading organisms (18), expression data (19), and functional studies (5, 6) demonstrate the great importance of LPMOs in biomass processing. Thus, studying the structure and function of these enzymes is

Significance

The discovery of lytic polysaccharide monooxygenases (LPMOs) has profoundly changed our understanding of the enzymatic conversion of recalcitrant polysaccharides, such as cellulose. Although in-depth studies of fungal cellulolytic LPMOs have been reported, the structures and functions of their bacterial counterparts with no detectable sequence similarity remain largely elusive. We present the structures of a conserved pair of bacterial cellulose-active LPMOs supplemented with extensive functional characterization. The structural data allow a thorough comparative assessment of fungal and bacterial LPMOs, providing insight into the structural basis of substrate specificity and the oxidative mechanism (C1/C4 oxidation). Importantly, we show that this LPMO pair acts synergistically when degrading cellulose, a finding that may help explain the occurrence of multiple LPMOs in a single microbe.

Author contributions: Z.F., M.S., G.V.-K., and V.G.H.E. designed research; Z.F., A.K.M., M.S., Å.K.R., R.H., and A.S.A. performed research; Å.K.R., R.H., and A.S.A. contributed new reagents/analytic tools; Z.F., A.K.M., M.S., Å.K.R., G.V.-K., and V.G.H.E. analyzed data; and Z.F., A.K.M., M.S., G.V.-K., and V.G.H.E. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: Atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4OY6, 4OY7, and 4OY8).

[1]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.1402771111/-/DCSupplemental.

of interest. Fungal LPMO9s have been relatively well studied and are known to act on cellulose and oxidize C1, C4, or both (11, 13). In contrast, little is known about bacterial cellulose-degrading LPMO10s. Oxidative cleavage of cellulose has been described for only two enzymes, both oxidizing C1 (16, 20), and no structural information is available. All LPMO10s with known structures act on chitin and oxidize C1. This lack of information limits our understanding of bacterial cellulose degradation as well as LPMO functionality and diversity.

The dominant bacterial genus responsible for aerobic biomass decomposition in soil is the Gram-positive Streptomycetes (21). A recent secretome/transcriptome study showed that *Streptomyces* sp. SirexAA-E (ActE) secretes a plethora of enzymes targeting carbohydrates, including abundantly expressed LPMO10s (19). Two of its six LPMOs (SACTE_3159 and SACTE_6428) were up-regulated and secreted during growth on pure cellulosic substrates or plant biomass, and three other LPMOs were up-regulated when chitin served as the substrate (19). Homologs of SACTE_3159 and SACTE_6428 in the well-studied cellulolytic bacterium *Thermobifida fusca* YX, known as E7 and E8, are also up-regulated during growth on cellulose (22). Homologs of this pair of putative LPMOs also exist among the seven LPMOs encoded in the *Streptomyces coelicolor* A3 (2) genome (*Sc*LPMO10B and CelS2, respectively; the formal name of CelS2 is *Sc*LPMO10C). Previous studies have indicated that CelS2 is coexpressed with a cellulase (23), and we previously showed that CelS2 is a C1-oxidizing LPMO acting in synergy with cellulases (16). The enzymatic properties of other bacterial LPMOs putatively acting on cellulose, and the functional significance of the coexpression of CelS2-*Sc*LPMO10B-like pairs of LPMOs, have not yet been described.

Here we show that *Sc*LPMO10B (and E7) are C1- and C4-oxidizing LPMOs that complement the activity of C1-oxidizing CelS2 (and E8), yielding synergy when combined in a reaction. We also present the X-ray crystallographic structures of both *S. coelicolor* enzymes, that is, the first structures of bacterial cellulose-degrading LPMOs. The analysis and comparison of these structures is supported by studies of copper binding by isothermal titration calorimetry (ITC), determination of the redox potentials ($E°$), and analysis of copper coordination by electron paramagnetic resonance (EPR) spectroscopy.

Results

Enzyme Activity. Five LPMOs were produced and characterized: CelS2 from *S. coelicolor* (comprising an LPMO domain and a family 2 carbohydrate-binding module termed CBM2), the N-terminal LPMO domain of CelS2 (CelS2-N), the N-terminal domain of E8 (E8-N), and the single-domain proteins *Sc*LPMO10B and E7. Studies with phosphoric acid swollen cellulose (PASC) and Avicel found that CelS2, CelS2-N, and E8-N generated C1-oxidized products only, whereas *Sc*LPMO10B and E7 showed a different product profile (Fig. 1 and *SI Appendix*, Fig. S1). Chromatographic peak assignments from a recent in-depth study of a C4-oxidizing LPMO9 (17) allowed identification of the additional products generated by these latter two enzymes as C4-oxidized (4-ketoaldoses) and double (C4/C1)-oxidized cello-oligosaccharides (*SI Appendix*, Fig. S2). MS data support this conclusion (Fig. 1).

To verify the presence of double-oxidized products, we degraded soluble products formed by *Sc*LPMO10B, E7 or CelS2-N (negative control) with a cellobiohydrolase, followed by C1 oxidation by cellobiose dehydrogenase from *Myriococcum thermophilum*

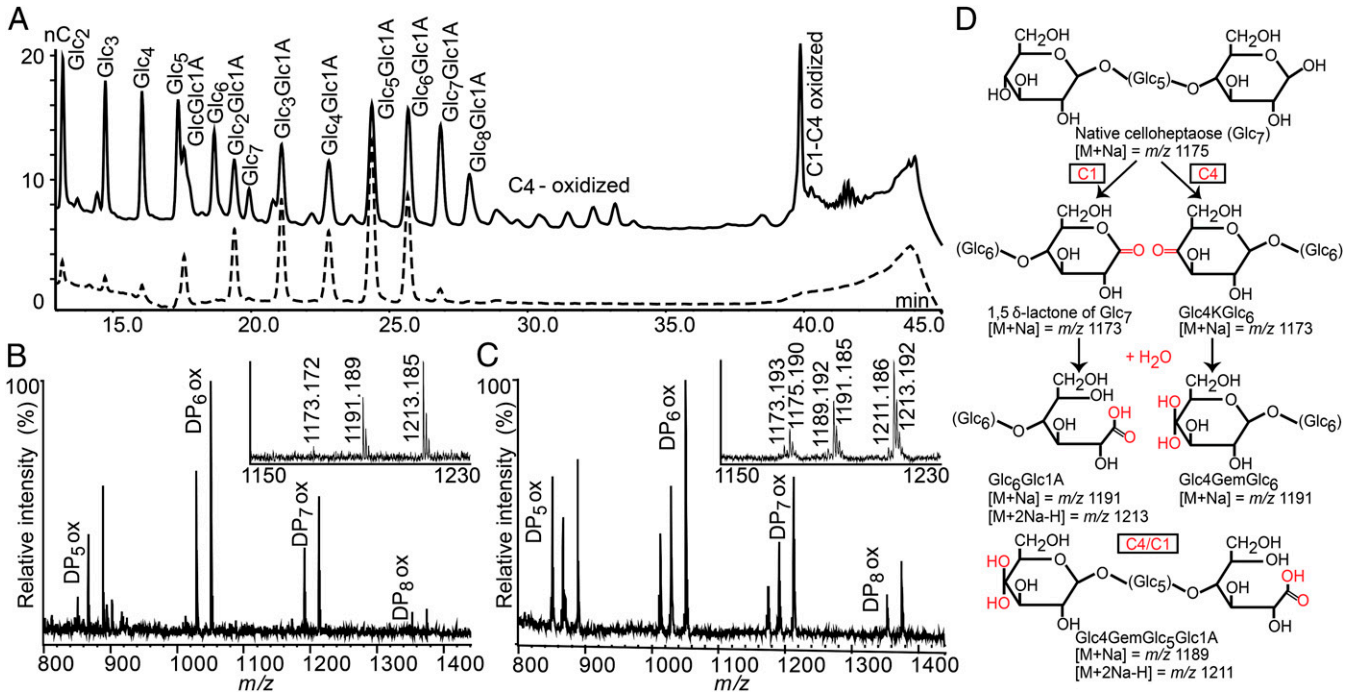


Fig. 1. Comparison of the two cellulose-active LPMOs from *S. coelicolor* in PASC degradation. (*A*) High-performance anion-exchange chromatography product profile for CelS2-N (dotted chromatogram) and *Sc*LPMO10B (solid chromatogram). Peaks were annotated using native and C1-oxidized standard cello-oligosaccharides (16), and products were compared with those obtained previously with C4-oxidizing *Nc*LPMO9C (17) (*SI Appendix*, Fig. S2). (*B* and *C*) MALDI-TOF MS analysis of products generated by CelS2 (*B*) and *Sc*LPMO10B (*C*), with sodium saturation. (*Insets*) Details of the heptamer ion clusters. (*D*) Possible products in these clusters. Both *B* and *C* show the lactone or ketoaldose (1173), the aldonic acid or gemdiol (4-ketoaldose + water:1191), and the sodium adduct of the aldonic acid sodium salt (1213). *C* also shows native Glc_7 (1175), the double-oxidized heptamer (1189), and the sodium adduct of the sodium salt of the double-oxidized heptamer (1211). Here 100% relative intensity represents 4.2×10^4 arbitrary units (a.u.) for the full spectra and 2.8×10^4 a.u. for the spectra in the *Insets*. Similar results for the E7/E8 enzyme pair are shown in *SI Appendix*, Fig. S1.

(*Mt*CDH). As expected, peaks assigned as double-oxidized disappeared after cellobiohydrolase treatment, whereas double-oxidized products, resulting from C1 oxidation of C4-oxidized fragments, were detected again after subsequent treatment with *Mt*CDH (*SI Appendix*, Fig. S3). Taken together, these data clearly show that *Sc*LPMO10B and E7 have mixed activity, yielding C1- and C4-oxidized products. Double-oxidized products and native oligomers alike are the result of two oxidative cleavages in the same polysaccharide chain (one C1 and one C4) with different outcomes, whereas for all LPMOs, native oligomers also result from oxidative cleavage near an original chain end (24).

*Sc*LPMO10B and E7, but not CelS2 and E8-N, also showed activity on squid pen β-chitin (*SI Appendix*, Fig. S4), whereas no products were obtained from crab shell α-chitin. The product profiles differed from those previously described for chitin-active LPMOs (5, 9), in that considerable amounts of partially deacetylated oligomers were produced. This indicates that the enzymes also (or preferably) act on deacetylated regions of the substrate, which notably resemble cellulose rather than fully acetylated parts.

Product formation over time was assessed by incubating CelS2, CelS2-N, and *Sc*LPMO10B with PASC, individually or in combination. Clear synergistic effects were observed after CelS2 and *Sc*LPMO10B were combined in the same reaction (Fig. 2 and *SI Appendix*, Fig. S5). A similar effect was seen when E7 and E8 were combined (*SI Appendix*, Fig. S6).

Three-Dimensional Structures of CelS2 and *Sc*LPMO10B. The structure of *Sc*LPMO10B was determined to 2.1 Å using zinc single-wavelength anomalous diffraction. The partially refined structure was then used as a search model for molecular replacement to obtain the structures of *Sc*LPMO10B with zinc (1.4 Å) and copper (1.3 Å), as well as the structure of CelS2-N with copper (1.5 Å). Statistics for diffraction data and structure refinement are summarized in *SI Appendix*, Table S1. Both *Sc*LPMO10B and CelS2 have the Ig-like β-sandwich fold observed in other LPMO structures, which also includes the sequence-disparate family AA9 LPMOs (Fig. 3 and *SI Appendix*, Fig. S7). The family AA10 LPMOs consist of a distorted β-sandwich comprising two β-sheets, one containing three antiparallel strands (S1, S4, and S7) and the other containing four antiparallel strands (S5, S6, S8, and S9). Additional strands adorn the four-stranded sheet in either an antiparallel (S2) or parallel (S3) fashion. The β-sandwich contains several conserved aromatic residues, potentially contributing to the electron transfer pathway necessary for reduction of the copper ion.

Structural diversity in both AA10 and AA9 families is confined mainly to the region between strands S1 and S3, known as loop 2 (L2 loop) in LPMO9s. This region forms a large protuberance, which in LPMO10s contains several helices and contributes to at least 50% of the putative substrate-binding surface. As shown in Fig. 3, the L2 loops of the two cellulose-active *S. coelicolor* LPMOs are larger than the L2 loop of the archetypal chitin-active *Sm*LPMO10A (CBP21). A similarly small L2 loop is present in chitin-active *Ef*LPMO10A (4AO2) (25), whereas the L2 loop in *Ba*LPMO10A (2YOY, which probably is chitin-active) (3, 26) is larger. Both *S. coelicolor* LPMOs have two disulfide bridges in the L2 region (*SI Appendix*, Fig. S7). One of these disulfide bridges links helix H1 to helix H1.1, which is conserved in CBP21, and the other tethers the L2 loop to the four-stranded β-sheet (via strand S9). The *Ba*LPMO10A and *Ef*LPMO10A structures do not contain any disulfide bridges. Relative to CBP21, CelS2 and (by inference from sequence alignment) E8 (*SI Appendix*, Fig. S7) have an insertion between strand S6 and S7 that contributes to the putative binding surface and corresponds to an area designated the LS loop in LPMO9s (27).

Aromatic residues are often involved in enzyme–carbohydrate interactions and are indeed found on the surface of LPMO9s in conformations (i.e., rings parallel to the binding surface) that suggest a role in substrate binding (13, 27). The two *S. coelicolor* LPMO10s have only one aromatic residue with a ring parallel to the binding surface, in structurally equivalent positions, Tyr79 in CelS2 and Trp88 in *Sc*LPMO10B (Fig. 3). This single aromatic residue is conserved and positioned similarly in essentially all LPMO10s with known structures, including Tyr54 in CBP21 (2BEM) (28), Trp58 in *Ef*LPMO10A (4A02) (25), and Trp50 in *Ba*LPMO10A (2YOW) (3). The binding surfaces of chitin-active LPMO10s have a cavity close to the catalytic center that has been postulated to accommodate dioxygen (3); this cavity is absent in the cellulose-active LPMO10s (Fig. 3 and *SI Appendix*, Fig. S8).

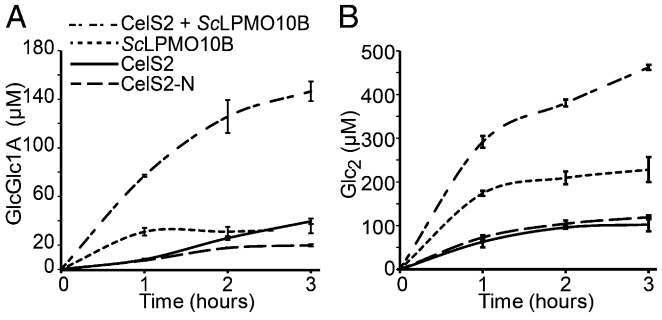


Fig. 2. Time course of released oxidized (*A*) and native (*B*) cellobiose after incubation of 1 μM LPMO (CelS2, CelS2-N, or *Sc*LPMO10B) or a mixture of 0.5 μM CelS2 and 0.5 μM *Sc*LPMO10B with 2 g/L PASC and 2 mM ascorbic acid in 20 mM ammonium acetate buffer (pH 6.0). Before analysis, oligomeric products generated by the LPMOs (Fig. 1*A*) were converted to shorter fragments by treatment with a cellobiohydrolase to facilitate product quantification. In the sample with the highest sugar concentration (i.e., CelS2 + *Sc*LPMO10B; 3 h), the amounts of cellobionic acid (GlcGlc1A) and cellobiose (Glc_2) represent 2.4% ± 0.2% and 7.5% ± 0.1% conversion, respectively, of the substrate. Glc_2Glc1A was an additional minor end product, and the Glc_2Glc1A: GlcGlc1A ratio was approximately constant. Control experiments (dose–response curves for the individual LPMOs) are shown in *SI Appendix*, Fig. S5. SDs were calculated based on three independent reactions. Similar results for the *T. fusca* LPMO pair are shown in *SI Appendix*, Fig. S6.

Copper Site. The copper site of *Sc*LPMO10B is highly similar to that of the C1/C4 cellulose-oxidizing LPMO9A from *Thermoascus aurantiacus* (*Ta*LPMO9A) (10) and exhibits an octahedral coordination geometry with Jahn–Teller distortion (Fig. 3). Three equatorial ligands form a "histidine brace" (10) composed of His43, His150, and the N-terminal amino group (His43). An acetate ion occupies the solvent-facing axial position, while its other oxygen extends partially into the fourth equatorial position (*SI Appendix*, Fig. S9). The axial position on the protein-facing side is occupied by the hydroxyl group of Tyr219 at a distance of 3.3 Å. A third histidine, His214, stabilizes His150 through π–π interaction, as seen in *Ta*LPMO9A (His164 stabilizes His86) (10). Ala148, which is highly conserved in LPMO10s, has been suggested to restrict access to the solvent-facing axial position (3) and is present in *Sc*LPMO10B as well.

In contrast, the copper site of CelS2 is similar to that of chitin-active LPMO10s (3, 25, 28) (Fig. 3), clearly showing that the active site features found in *Sc*LPMO10B, and first described for *Ta*LPMO9A (10), are not essential for activity on cellulose. Minor differences between CBP21 and CelS2 occur directly outside the copper coordination shell; Glu60 (CBP21, strand 2) and Glu217 (CelS2, strand 9) are at structurally equivalent positions but located at different positions in the protein sequences. In CelS2, this glutamate has a hydrogen bond with Arg212, which fills in the cavity typically seen in the chitin-active LPMOs (Fig. 3 and *SI Appendix*, Fig. S8). Interestingly, the eight molecules in the asymmetric unit show variation in copper

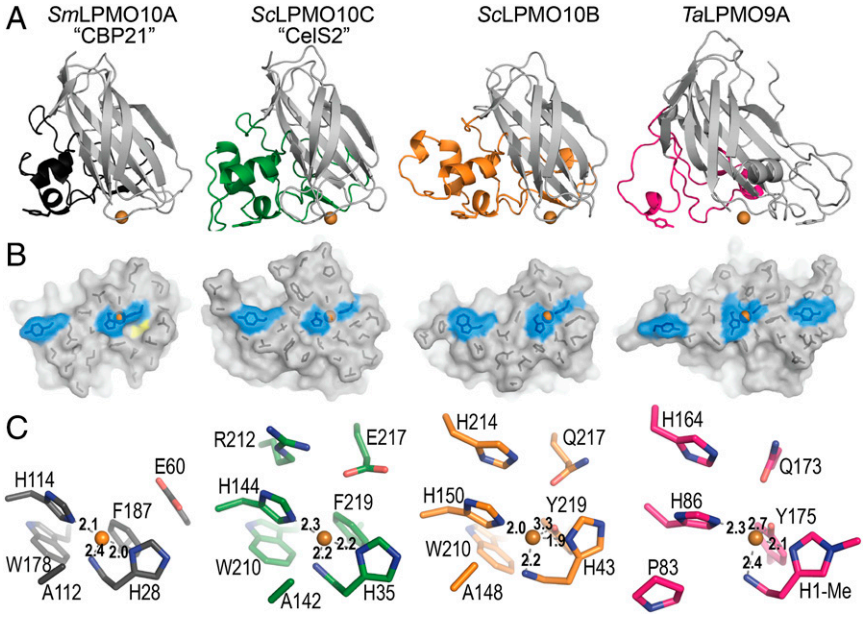


Fig. 3. Structural comparison of four LPMOs. (*A*) Cartoon representation of CBP21 (black; chitin-active, C1-oxidizing, PDB ID code 2BEM), CelS2 (green; cellulose-active, C1-oxidizing), *Sc*LPMO10B (orange; a C1/C4-oxidizer of cellulosic substrates, also capable of C1 oxidation of β-chitin), and *Ta*LPMO9A (pink; a C1/C4-oxidizer of cellulosic substrates, PDB ID code 2YET). The L2 loop is colored, and the rest of the enzyme is in gray. Residues with aromatic rings lying parallel to the putative substrate-binding surface are shown as sticks. Metal ions are shown as orange spheres. (*B*) Surface projection of the proposed substrate-binding surface, related to Fig. 3*A* by a 90° rotation along the horizontal axis. The side chains of residues forming the relatively flat surface are shown as black sticks. Metal ions are shown as orange spheres; coordinating histidine residues and residues with aromatic rings parallel to the surface are in blue. The yellow area in CBP21 indicates a characteristic cavity on the surface of chitin-active LPMO10s (*SI Appendix*, Fig. S8). (*C*) LPMO active sites, showing residues within 6 Å of the respective metal ions. Figures were created in PyMOL (32).

coordination and ligand bond lengths that may reflect variation in the oxidation state of the copper (*SI Appendix*, Fig. S10); see Hemsworth et al. (11) for a discussion. All chains (A–H) show density (refined as a water molecule) in the equatorial position located 1.6–2.0 Å from the copper and in close contact with Glu217 (2.7 Å). Chains E and G have additional density (refined as water) positioned 2.4–2.9 Å from the copper, askew from the axial position. Electron density has been observed in the same position for *Ef*LPMO10A containing oxidized copper (Cu^{2+}) (PDB ID code 4ALC) (11). We note that the average distance between the density in the equatorial position (refined as water) and copper is 1.8 Å (*SI Appendix*, Fig. S10), and that this distance corresponds to that of the Cu(II)-oxyl species described in a recently suggested catalytic mechanism for LPMOs (1.8 Å) (14) and in other studies (1.8–1.9 Å) (29–31).

Metal Binding, Redox Potential, and EPR Spectroscopy. ITC-measured dissociation constants were 31 nM for CelS2-Cu^{2+} and 12 nM for *Sc*LPMO10B-Cu^{2+} (*SI Appendix*, Fig. S11 and Table 1). Redox potentials of 242 ± 7 mV for CelS2 and 251 ± 15 mV for *Sc*LPMO10B were determined as described previously (2) and shown in *SI Appendix*, Fig. S12. Combining the redox potentials and dissociation constants for Cu^{2+} in three thermodynamic relationships (*SI Appendix*, Fig. S12) allowed estimation of the dissociation constants for reduced copper (Cu^{+}), resulting in values of 1.1 nM for CelS2 and 0.3 nM for *Sc*LPMO10B.

Copper coordination was studied by EPR spectroscopy (*SI Appendix*, Fig. S13). The EPR spectra were simulated; the estimated spin Hamiltonian parameters are summarized in Table 2. The g and A^{Cu} tensors reflect the active site copper coordination environment, and of these the g_z and A_z^{Cu} tensors could be modeled with greatest accuracy (3). The g_z and A_z^{Cu} tensors of CelS2 (g_z = 2.267; A_z^{Cu}= 153 • 10^{-4} cm^{-1}) and *Sc*LPMO10B (g_z = 2.270; A_z^{Cu}= 158 • 10^{-4} cm^{-1}) were similar and resemble those reported previously for cellulose-active *Ta*LPMO9A (g_z = 2.27; A_z^{Cu}= 162 • 10^{-4} cm^{-1}) (10). Interestingly, chitin-active LPMO10s showed a lower A_z^{Cu} tensor (*SI Appendix*, Fig. S14).

Table 1. Thermodynamic parameters obtained from ITC experiments performed in triplicates for binding of Cu^{2+} to apo-CelS2 or apo-*Sc*LPMO10B at pH 5.5 and t = 10 °C

LPMO	K_d, nM	$\Delta G°$, kcal/mol	$\Delta H°$, kcal/mol	$-T\Delta S°$, kcal/mol
CelS2	31 ± 6	−9.7 ± 0.1	−6.4 ± 0.6	−3.3 ± 0.6
*Sc*LPMO10B	12 ± 6	−10.3 ± 0.3	−9.6 ± 1.3	−0.7 ± 1.3

Discussion

All of the LPMO10s that were characterized before the present study are C1 oxidizers (5, 16, 20, 25), whereas the family of cellulose-active fungal LPMO9s contains strict C1 oxidizers, strict C4 oxidizers, and enzymes that can oxidize both C1 and C4 (15). Here we show that the cellulolytic enzyme systems of *S. coelicolor* and *T. fusca* also are equipped with more than one LPMO type, a strict C1 oxidizer (CelS2 and E8) and a C1/C4 oxidizer (*Sc*LPMO10B and E7). Importantly, the two enzymes display synergy when acting on cellulose (Fig. 2 and *SI Appendix*, Fig. S6), providing a possible explanation for the occurrence of a multitude of LPMOs in biomass-degrading microorganisms. Interestingly, sequence analysis shows that this pair of enzymes is common in other cellulolytic actinomycetes, including *Cellulomonas*, *Micromonospora*, *Streptomyces*, *Thermobifida*, and *Xylanimonas* (Fig. 4).

The two crystal structures presented here and the data on redox properties and metal binding allow, for the first time to our knowledge, structural and functional comparison of LPMOs belonging to the same family but having different substrate specificities. Whereas phylogenetic analysis of LPMO10s with known or inferred activities on chitin and cellulose shows clear separation between the two substrate specificities (Fig. 4), the structural and functional data reveal few pronounced differences. The chitin- and

Table 2. Spin Hamiltonian parameters*

Parameter	Cu(II) buffer	*Sc*LPMO10B	CelS2
g_x	2.059	2.020	2.015
g_y	2.059	2.090	2.102
g_z	2.270	2.270	2.267
A_x^{Cu}†	12.3	5.0	11.7
A_y^{Cu}†	12.3	10.0	17.0
A_z^{Cu}†	165	158	153

*Assuming collinear *g* and A^{Cu} tensors in all simulations.
†(10^{-4} cm^{-1}).

cellulose-active LPMO10s have similar redox potentials and copper affinities [data for chitin from Aachmann et al. (2) and Hemsworth et al. (3)], and the cores of their copper-binding sites, the histidine braces, are essentially identical. The copper-binding sites of CelS2 and chitin-active LPMO10s, such as CBP21, are remarkably similar (Fig. 3), whereas the copper-binding site of *Sc*LPMO10B is clearly different, resembling that of LPMO9s. On the other hand, the EPR spectrum of *Sc*LPMO10B is very similar to that of CelS2.

Two clear findings stand out. First, chitin-active and cellulose-active LPMO10s seem to be separated by their A_z^{Cu} (*SI Appendix*, Fig. S14), and the values suggest a more distorted axial geometry in the former than in the latter (3, 20). It should be noted, however, that the recently discovered chitin-active LPMO11 from *Aspergillus oryzae* is grouped with the cellulose-active LPMOs according to the Peisach–Blumberg plot (*SI Appendix*, Fig. S14) (9). Second, the substrate-binding surface of chitin-active LPMO10s consistently contains a cavity adjacent to the active site (Fig. 3 and *SI Appendix*, Fig. S8) that is lacking from the cellulose-active enzymes, potentially accommodating an *N*-acetyl group.

The conserved active site alanine in LPMO10s (Fig. 3) has been postulated to prevent ligands from binding in the solvent-facing axial position, making dioxygen activation more likely in the equatorial position (3, 11). Conversely, binding of dioxygen in the solvent-facing axial position has been proposed for a C1/C4-oxidizing LPMO9 (14), which lacks this alanine (10). Indeed, C1-oxidizing CelS2 seems incapable of ligand coordination in the solvent-facing axial position (*SI Appendix*, Fig. S10). In contrast, *Sc*LPMO10B does coordinate a ligand (acetate) in this position, despite the presence of the alanine (*SI Appendix*, Fig. S9). Importantly, comparison of the structures presented here shows that the loop hosting the conserved alanine adopts different conformations in the two enzymes (Fig. 5). In *Sc*LPMO10B, Asp146 forces the main chain of the loop into a conformation that positions Ala148 2.5 Å away from the position of the corresponding alanine (Ala142) in CelS2, allowing sufficient space for a ligand in the solvent-facing axial position. Asp146 is conserved in E7 and in other LPMO10s that phylogenetically cluster with E7 and *Sc*LPMO10B (Fig. 4 and *SI Appendix*, Fig. S7). Importantly, closer inspection of available LPMO9 structures revealed a similar scenario (*SI Appendix*, Fig. S15); LPMO9s with C4- and C1/C4-oxidizing activity have an open solvent-facing axial coordination site, whereas LPMO9s known to be strict C1 oxidizers have a tyrosine, preventing optimal axial access to the copper ion (*SI Appendix*, Fig. S15). Thus, the ability to bind a ligand in the axial position could be a determinant of C4-oxidizing activity.

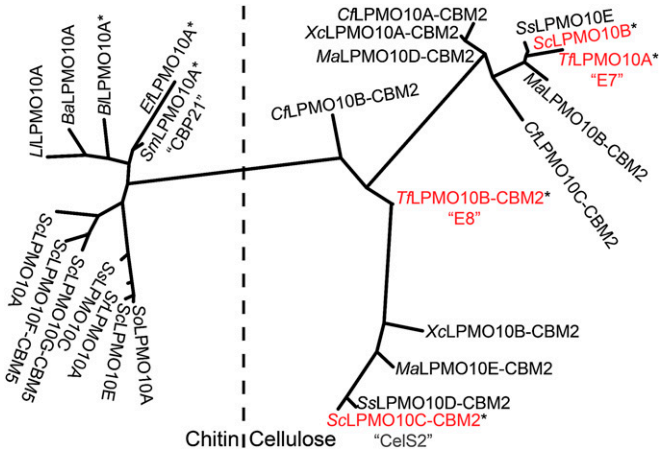


Fig. 4. Phylogeny of six *S. coelicolor* LPMO10s and a selection of other LPMO10s, selected on the basis of literature data documenting up-regulation during growth on biomass (19), binding to chitin (26, 33, 34), synergism with glycoside hydrolases (35), and/or substrate cleavage and oxidation (5, 16, 20, 25). Proteins for which substrate degradation and oxidation has been demonstrated are labeled with an asterisk.

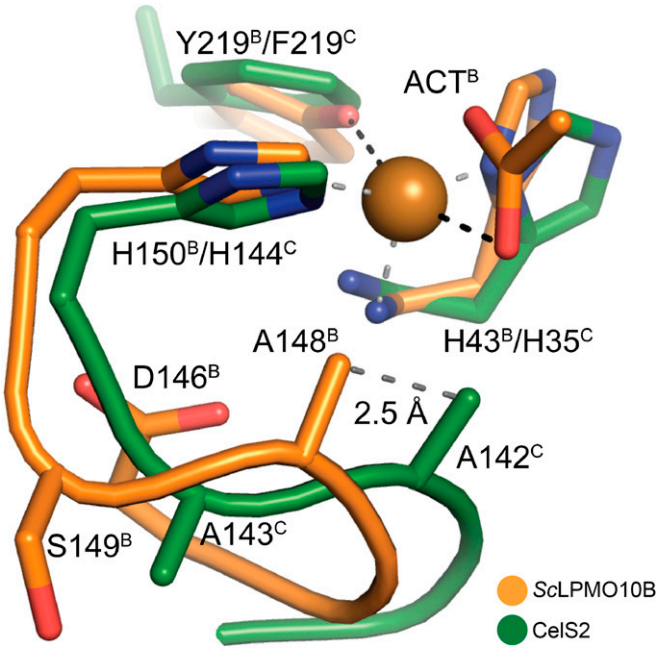


Fig. 5. Position of the conserved active site alanine in CelS2 (green) and *Sc*LPMO10B (orange). Black and gray dashed lines indicate the axially and equatorially coordinated ligands, respectively, in *Sc*LPMO10B. ACT indicates the acetate ion; the skewed equatorial contact between the acetate and the copper is not drawn (*SI Appendix*, Fig. S9). Note the 2.5-Å relative shift in position of Ala142/148 between CelS2 and *Sc*LPMO10B; in the former, the alanine (Ala142) is much closer to the acetate oxygen.

When the first combined functional and structural data on cellulose-active LPMO9s and chitin-active LPMO10s became available, it was suggested that differences in the copper-containing catalytic centers of these enzymes (i.e., between these families) could yield variable oxidative power, which in turn could affect the ability to cleave, for example, cellulose (11). Our data show no obvious correlation between the geometry of the copper center and substrate specificity. In fact, including recent data on a LPMO11, it would seem that LPMOs use a continuum of active site configurations, with the histidine brace being the key conserved element. All in all, it seems that substrate specificity depends on variations that are more remote from the active site and that affect substrate binding and positioning, as well as possibly electron transfer in the enzyme–substrate complex. The cavity observed in the binding surface of chitin-active LPMO10s supports this notion. On this note, one might expect some substrate promiscuity, as we indeed observed for *Sc*LPMO10B. Clearly, additional data on enzyme–substrate interactions, supplementing data from modeling and NMR studies (2, 13, 27), are needed to gain more insight into the issue of substrate specificity.

Materials and Methods

Detailed information for all experimental procedures is provided in *SI Appendix, Materials and Methods*.

Enzyme Activity. LPMOs were expressed heterologously in *Escherichia coli* and purified by chitin-affinity, ion-exchange, and size-exclusion chromatography. Substrate degradation was analyzed by MALDI-TOF MS and high-performance anion-exchange chromatography. Standard reactions were set up

BIOCHEMISTRY

with 2 g/L substrate (PASC or β-chitin), 1 μM LPMO, and 2 mM ascorbate in 20 mM ammonium acetate buffer (pH 6.0) in a shaking incubator at 50 °C.

Metal Binding by ITC. Dissociation constants and thermodynamic data for binding of Cu^{2+} to *Sc*LPMO10B and CelS2 were obtained by measuring the heat produced by titrating 4-μL aliquots of metal ion solution (120–150 μM) to 5 μM apo-LPMO in 20 mM Mes (pH 5.5) in a 1.42-mL reaction cell of a VP-ITC system (MicroCal) at 10 °C.

Determination of Redox Potential and Dissociation Constant. The cell potential for the LPMO-Cu^{2+}/LPMO-Cu^{+} redox couple was determined by monitoring a reaction between reduced *N,N,N′,N′*-tetramethyl-1,4-phenylenediamine (TMP_{red}) and LPMO-Cu^{2+}. The calculated $E°$ value was then combined with ITC data to estimate the dissociation constant for Cu^{+} as described previously (2).

EPR Spectroscopy. Resting-state EPR spectra were recorded for Cu^{2+}-charged *Sc*LPMO10B and CelS2 in Pipes buffer (pH 6.0) using a Bruker EleXsyS 560 SuperX instrument equipped with an ER 4122 SHQE SuperX high-sensitivity cavity and a liquid nitrogen cooled cold finger. The instrument settings were 0.5 mW microwave power, 5 G modulation amplitude, and a temperature of 77 K.

Protein Crystallization and Data Collection. Details and statistics for protein crystallization, data collection, structure determination, and refinement are provided in *SI Appendix, Materials and Methods* and Table S1.

ACKNOWLEDGMENTS. We thank Roland Ludwig for the purified *Mt*CDH, Gabriele Cordara for assistance with data collection, the European Synchrotron Radiation Facility and Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung staff for help and beamtime, and Gregg Beckham and Seonah Kim (National Renewable Energy Laboratory, Golden, CO) for helpful discussions. This work was funded by Norwegian Research Council Grants 190965, 214138, and 214613 and by the Norwegian Academy of Science and Letters Vista Program Grant 6505. A.S.A.'s work on structural methods is supported by the US Department of Energy's Integrated Diffraction Analysis Technologies program.

1. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH (2012) Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels* 5(1):45.
2. Aachmann FL, Sørlie M, Skjåk-Bræk G, Eijsink VGH, 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.
3. Hemsworth GR, et al. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc* 135(16):6069–6077.
4. Vaaje-Kolstad G, Horn SJ, van Aalten DM, 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.
5. Vaaje-Kolstad G, et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330(6001):219–222.
6. 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.
7. Cannella D, Jørgensen H (2014) Do new cellulolytic enzyme preparations affect the industrial strategies for high solids lignocellulosic ethanol production? *Biotechnol Bioeng* 111(1):59–68.
8. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels* 6(1):41.
9. Hemsworth GR, Henrissat B, Davies GJ, Walton PH (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* 10(2):122–126.
10. 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.
11. Hemsworth GR, Davies GJ, Walton PH (2013) Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol* 23(5):660–668.
12. Beeson WT, Phillips CM, Cate JH, Marletta MA (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc* 134(2): 890–892.
13. 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.
14. Kim S, Ståhlberg J, Sandgren M, Paton RS, Beckham GT (2014) Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc Natl Acad Sci USA* 111(1):149–154.
15. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc* 136(2):562–565.
16. Forsberg Z, et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20(9): 1479–1483.
17. Isaksen T, et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem* 289(5):2632–2642.
18. Berka RM, et al. (2011) Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol* 29(10):922–927.
19. Takasuka TE, Book AJ, Lewin GR, Currie CR, Fox BG (2013) Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Sci Rep* 3:1030.
20. Forsberg Z, et al. (2014) Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry* 53(10): 1647–1656.
21. Hodgson DA (2000) Primary metabolism and its control in streptomycetes: A most unusual group of bacteria. *Adv Microb Physiol* 42:47–238.
22. Adav SS, Ng CS, Arulmani M, Sze SK (2010) Quantitative iTRAQ secretome analysis of cellulolytic *Thermobifida fusca*. *J Proteome Res* 9(6):3016–3024.
23. Garda AL, Fernández-Abalos JM, Sánchez P, Ruiz-Arribas A, Santamaría RI (1997) Two genes encoding an endoglucanase and a cellulose-binding protein are clustered and co-regulated by a TTA codon in *Streptomyces halstedii* JM8. *Biochem J* 324(Pt 2): 403–411.
24. Westereng B, et al. (2011) The putative endoglucanase *Pc*GH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE* 6(11):e27807.
25. Vaaje-Kolstad G, et al. (2012) Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J Mol Biol* 416(2):239–254.
26. Chu HH, Hoang V, Hofemeister J, Schrempf H (2001) A *Bacillus amyloliquefaciens* ChbB protein binds beta- and alpha-chitin and has homologues in related strains. *Microbiology* 147(Pt 7):1793–1803.
27. 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.
28. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VGH, van Aalten DM (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* 280(12):11313–11319.
29. Yoshizawa K, Kihara N, Kamachi T, Shiota Y (2006) Catalytic mechanism of dopamine beta-monooxygenase mediated by Cu(III)-oxo. *Inorg Chem* 45(7):3034–3041.
30. Huber SM, et al. (2009) Generating Cu(II)-oxyl/Cu(III)-oxo species from Cu(I)-alpha-ketocarboxylate complexes and O2: In silico studies on ligand effects and C-H-activation reactivity. *Chemistry* 15(19):4886–4895.
31. Crespo A, Martí MA, Roitberg AE, Amzel LM, Estrin DA (2006) The catalytic mechanism of peptidylglycine alpha-hydroxylating monooxygenase investigated by computer simulation. *J Am Chem Soc* 128(39):12817–12828.
32. DeLano WL, Lam JW (2005) PyMOL: A communications tool for computational models. *Abstr Pap Am Chem S* 230:U1371–U1372.
33. Kolbe S, Fischer S, Becirevic A, Hinz P, Schrempf H (1998) The *Streptomyces reticuli* alpha-chitin–binding protein CHB2 and its gene. *Microbiology* 144(Pt 5):1291–1297.
34. 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. *Mol Microbiol* 13(5):807–819.
35. Vaaje-Kolstad G, Bunaes AC, Mathiesen G, Eijsink VGH (2009) The chitinolytic system of *Lactococcus lactis* ssp. *lactis* comprises a nonprocessive chitinase and a chitin-binding protein that promotes the degradation of alpha- and beta-chitin. *FEBS J* 276(8):2402–2415.

Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases

Zarah Forsberg[a], Alasdair K. Mackenzie[a], Morten Sørlie[a], Åsmund K. Røhr[b], Ronny Helland[c], Andrew S. Arvai[d], Gustav Vaaje-Kolstad[a], and Vincent G.H. Eijsink[a,1]

a. Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Aas, Norway
b. Department of Biosciences, University of Oslo, P.O. Box 1066, Blindern, N-0316 Oslo, Norway
c. Department of Chemistry, NorStruct, Forskningsparken 3, Faculty of Science and Technology, University of Tromsø, Norway
d. Department of Molecular Biology, Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037

[1] Corresponding author:
Vincent G.H. Eijsink Tel.: +47 6496 5892 Address: Department of Chemistry, Biotechnology and Food Science Norwegian University of Life Sciences P.O. Box 5003, NO-1432 Aas, Norway, vincent.eijsink@nmbu.no

Supporting information

Table of Contents:

Materials and Methods

Cloning, Expression and Purification. Genes encoding *Sc*LPMO10B (residues 43-228, UniProt ID; Q9RJC1) and CelS2 (*Sc*LPMO10C residues 35-364, Uniprot ID; Q9RJY2) from *S. coelicolor* A3(2) were codon optimized for *E. coli* expression (GenScript) and cloned into the pRSET B expression vector (Invitrogen) containing the native signal sequence of a *S. marcescens* chitin-active LPMO known as CBP21 (1). The pRSET B_*cbp21* vector was pre-cut with restriction endonucleases (*Bsm*I and *Hind*III), to remove the nucleotides encoding the mature CBP21 protein but preserving its signal sequence, prior to fusing the amplified LPMO genes using the In-Fusion HD cloning kit (Clontech). An expression plasmid encoding a truncated form of CelS2 (residues 35-234) lacking the CBM2 cellulose binding domain, CelS2-N, was also produced. The *T. fusca* YX genes encoding E7 (*Tf*LPMO10A residues 37-222; Uniprot ID Q47QG3) and E8 (truncated form with LPMO domain only; *Tf*LPMO10B residue 32-225; Uniprot ID Q47PB9) were cloned into the same expression vector from genomic DNA (ATCC no. BAA-629D-5™). Sequenced vectors were transformed by heat shock into chemically competent One Shot® BL21 Star™ (DE3) cells (Invitrogen). Fresh colonies were inoculated in LB-Amp (50 µg/mL) media and grown at 30 °C (*Sc*LPMO10B, CelS2-N and E8-N) or 37 °C (CelS2 and E7) for 20 hours at 200 rpm. After harvesting cells by centrifugation, periplasmic fractions were prepared using an osmotic shock method (2), which were sterilized by filtration (0.2 µm) prior to enzyme purification.

CelS2, CelS2-N and E8-N were purified by loading periplasmic extracts adjusted to Buffer A (50 mM Tris/HCl pH7.5) onto a 5 mL HiTrap DEAE FF anion exchanger (GE Healthcare) connected to an ÄKTA purifier FPLC system (GE Healthcare). LPMOs were eluted by using a linear salt gradient (0-500 mM NaCl) over 100 min at a flow rate of 3.5 mL/min. The LPMO containing fractions were pooled and concentrated using Amicon Ultra centrifugal filters (Millipore) with a molecular weight cut-off of 10 kDa. For further clean-up, samples were loaded onto a HiLoad 16/60 Superdex 75 size exclusion column (GE Healthcare), with a running buffer consisting of 50 mM Tris/HCl pH 7.5 and 200 mM NaCl, using a flow rate of 1 mL/min.

*Sc*LPMO10B was purified by a one-step protocol using a 1 mL cation HiTrap CM FF ion exchange column (GE Healthcare) with 20 mM sodium citrate pH 3.5 as running buffer. Protein was eluted using a linear NaCl gradient (0-500 mM) over 60 minutes at a flow rate of 1.5 mL/min. E7 was purified by chitin affinity chromatography as described previously for CBP21(1), using chitin beads (NEB) with 50 mM Tris/HCl pH 8.0 + 1 M $(NH_4)_2SO_4$ as starting buffer and using 20 mM acetic acid pH 3.6 for elution. Protein purity was analyzed by SDS-PAGE. Fractions containing pure protein were pooled and concentrated and the buffer was changed to 20 mM MES pH 5.5 prior to storage at 4 °C. Protein concentrations were determined using the Bradford assay (Bio-Rad).

Apo-enzymes (for ITC and EPR) and Cu^{2+}-saturated LPMOs (for degradation experiments) were generated by 30 min incubation with EDTA or Cu^{2+}, respectively, in a 1:4 molar ratio (LPMO:EDTA/Cu^{2+}), at room temperature. Excess EDTA/Cu^{2+} was removed by separation through a EDTA pre-striped PD MidiTrap G-25 column (GE Healthcare) equilibrated with Chelex (Bio-Rad) treated 20 mM MES buffer pH 5.5 (ITC experiments and copper saturated enzyme) or 20 mM Pipes buffer pH 6.0 (EPR experiments).

Cellulose Degradation Experiments. For all cellulose degradation experiments, 1 µM LPMO (CelS2, CelS2-N, *Sc*LPMO10B, E7 or E8-N, charged with copper) was incubated with 2 g/L phosphoric acid swollen cellulose (PASC), prepared from Avicel (3), Avicel® PH-101, or squid

pen β-chitin (France Chitin; Marseille, France), in 20 mM ammonium acetate pH 6.0, in the presence of 2 mM ascorbic acid. The enzyme reactions were incubated in an Eppendorf Thermomixer set to 50 °C and 900 rpm. To analyze double oxidized products generated by *Sc*LPMO10B or E7 after 16 h incubation, 0.25 μM of partially pure *T. reesei* cellobiohydrolase, *Tr*Cel7A, purified from Celluclast (Novozymes), was added to the soluble fraction in order to hydrolyze double-oxidized products into two single (shorter) products, oxidized at either the reducing or the non-reducing end. After 8 h incubation, the sample was split into two and 0.3 μM of *M. thermophilum* cellobiose dehydrogenase (*Mt*CDH), purified according to (4), was added, followed by another 16 h of incubation. For synergy experiments, reactions were run for 4 hours with an enzyme load of 1 μM LPMO (CelS2, CelS2-N or *Sc*LPMO10B) or 0.5 μM CelS2 + 0.5 μM *Sc*LPMO10B, using the conditions described above. Samples were taken every hour, and the LPMO reaction was stopped by separating the soluble fraction from the remaining insoluble fraction by centrifugation (16,000 × g for 5 min). To facilitate product quantification, the soluble products generated by the LPMOs were treated with 0.25 μM cellobiohydrolase at 50 °C for 16 h followed by chromatographic quantification of the products for which standards were available; cellobiose (Glc_2) and cellobionic acid (GlcGlc1A; obtained by oxidizing cellobiose with *Mt*CDH). The synergy experiment was performed in the same manner for the *T. fusca* LPMO-pair.

Product Analysis. Native and oxidized products in soluble fractions generated from PASC were analyzed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS). The samples were mixed with a 9 % solution of 2,5-dihydroxybenzoic acid (DHB) matrix in a 1:2 ratio, and the MS analysis was performed as described previously (5). Sodium saturation was accomplished by adjusting the sample to 5 mM NaCl (using a 10 mM NaCl stock solution), followed by 10 min incubation at room temperature, addition of DHB and MS analysis. In addition, samples were analyzed by high performance anion exchange chromatography (HPAEC) using a Dionex Bio-LC connected to a CarboPac PA1 column operated with a flow rate of 0.25 mL/min 0.1 M NaOH and a column temperature of 30 °C. Native and oxidized products were eluted and largely separated using a stepwise gradient with increasing amount of eluent B (0.1 M NaOH + 1 M NaOAc), as follows: 0-10 % B over 10 min, 10-30 % B over 25 min, 30-100 % B over 5 min, 100-0 % B over 1 min, 0 % B over 9 min. For samples from the synergy experiments, only containing short products, a steeper gradient was used, as follows; 0-10 % B over 10 min, 10-18 % B over 10 min, 18-30 % B over 9 min, 30-100 % B over 1 min, 100-0 % B over 0.1 min and 0 % B over 13.9 min. Eluted oligosaccharides were monitored by a pulsed amperometric detector (PAD) and chromatograms recorded using Chromeleon 7.0 software (6).

Protein Crystallization. Crystals of *Sc*LPMO10B and CelS2-N were obtained by hanging drop vapor diffusion at 20 °C, by mixing an equal volume of the reservoir solution with the concentrated protein in storage buffer (20 mM Tris/HCl pH 8.0). *Sc*LPMO10B crystallized in 0.2 M zinc acetate, 0.1 M sodium cacodylate pH 6.5 and 9 % v/v 2-propanol at a protein concentration of 10.3 mg/mL. CelS2-N crystals were obtained using 9 % w/v PEG 10K, 0.1 M sodium citrate pH 4.5, 0.1 M calcium acetate and 5 % v/v glycerol, at a protein concentration of 9.2 mg/mL. Prior to freezing, the crystals were soaked in cryo-protectant consisting of reservoir solution with the addition of either 25 % PEG400 (*Sc*LPMO10B- Zn complex) or 15 % glycerol (CelS2). To obtain a *Sc*LPMO10B-Cu complex, crystals were transferred to a reservoir solution

with a reduced zinc acetate concentration (0.1 M) and containing different $CuCl_2$ concentrations (1 mM-20 mM). The crystals were soaked for various time periods, ranging from 2 min to overnight. Crystals were flash frozen in liquid nitrogen in a cryo-solution consisting of the copper soaking solution and 25% v/v PEG 400.

Data Collection and Structure Determination. A single-wavelength anomalous diffraction (SAD) data set was collected on a *Sc*LPMO10B zinc derivative diffracting to 2.14 Å (dataset I) at the zinc edge under cryogenic conditions at ID14-4 (ESRF, Grenoble, France). After data collection the crystal was translated to the opposite end of the cryo-loop, where the same crystal was found to diffract to 1.4 Å, resulting in a second high resolution data-set being collected (data set II). The data were processed using XDS (7) and SCALA (8), where auto-processing software determined the space group to be P 3_1 2 1 with unit cell dimensions a = 67.2 Å, b = 67.2 Å, c = 107.2 Å (with α, β, γ = 90.0°, 90.0°, and 120.0° respectively). The presence of one molecule in the asymmetric unit gave a V_M of 3.35 $Å^3$ Da^{-1} (9, 10) with a solvent content of 63.3 %. The structure was solved using the SHELX suite (11). The zinc sites were located using SHELXD. Solvent flattening and density modification in SHELXE resulted in sufficient tracing of main-chain atoms revealing the classically conserved AA10 fold. Visual inspection of the electron density in *Coot* (12) allowed partial docking of the sequence to the incomplete model. ARP/wARP (13, 14) succeeded in tracing the entire protein molecule, and the resulting model was manually inspected using *Coot*, where essentially all residues possessed well defined electron density. Two rounds of manual rebuilding and refinement ensued using *Coot* and REFMAC5 (14) gave a partially refined model of the entire *Sc*LPMO10B enzyme. The model included eight zinc ions, and had R_{cryst}/R_{free} values of 23.4% and 27.6% respectively.

The partial SAD model was used as a molecular replacement model for the high resolution data set (data set II), which was processed using iMOSFLM (15) and SCALA (8). The structure was solved using MOLREP (16), and subjected to rigid body refinement, and an initial refinement step using REFMAC5. Following manual rebuilding in *Coot* the refinement of the structure was continued using the PHENIX suit (17). Solvent atoms were added using PHENIX refine (18) and ARP-wARP (19), with the final stages of refinement performed with mixed anisotropic refinement (excluding solvent) in REFMAC5, to yield a model with R_{cryst}/R_{free} values of 12.5 % and 14.3 % respectively.

Three data sets were collected under cryogenic conditions at ID29 (ESRF, France) on *Sc*LPMO10B crystals that had been soaked in a 0.1 M sodium acetate reservoir solution at different $CuCl_2$ concentrations for various time periods. The data were processed using XDS and SCALA, with cell dimensions isomorphous to the previous datasets (data set II). The model of the Zinc-complex (data set II), excluding solvent and ions, was used as a search model for molecular replacement using MOLREP. After rigid body refinement in REFMAC5 the models were subjected to simulated annealing and refinement using PHENIX refine. The electron density in the metal binding site of the enzyme, formed by the histidine brace of His43 and His150, was examined in all three data sets. Crystals soaked overnight in 1 mM $CuCl_2$ or for 2 minutes at 20 mM $CuCl_2$ contained a metal ion occupying the exact same position as the zinc ion of previously collected zinc-complex. Crystals soaked in 20 mM CuCl2 for 30 minutes contained a bound metal ion shifted 0.5 Å towards Tyr219 (Cu-complex, data set III, Table S1). The positions of the bound metal ions in data set III were confirmed using SHELXD. The calculated

phased anomalous difference map revealed electron density for metal ions, ranging from ca. 5 – 20 e/Å^3 (map contouring 5-37 σ) with two stronger peaks at 34 and 41 e/Å^3 (map contouring 61 and 72 σ). The strongest anomalous peak corresponded to the metal ion bound at the histidine brace formed by His43 and His150. The data were collected at a wavelength of 1.3Å, which is above the zinc edge, but below the copper edge, and thus any copper ions will have a far greater anomalous contribution compared to any zinc ions binding from the cryo-protectant solution. Taken together this suggests that the significant anomalous signal seen at the known AA10 copper-binding site in dataset III is a copper ion. Refinement was continued using PHENIX, with the final stages of anisotropic refinement performed using REFMAC5 to yield a model with and R_{cryst}/R_{free} value of 12.4 % and 14.0 % respectively.

A native data set was collected on a CelS2 (*Sc*LPMO10C) crystal diffracting to 1.5 Å (dataset IV, Table S1) under cryogenic conditions on ID14-4 (ESRF, Grenoble, France). The data were processed using XDS (7) and SCALA (8), and the space group determined to be P 2_1 2_1 2_1 with unit cell dimensions a = 83.9 Å, b = 122.9 Å, c = 156.0 Å (α, β, γ = 90.0°). The presence of eight molecules in the asymmetric unit gave a V_M of 2.39 Å^3 Da^{-1} (9) with a solvent content of 48.6 %. The structure of CelS2 was then solved by PHASER (20) using a model of *Sc*LPMO10B (dataset II), excluding solvent and ions. The structure was built using the auto-build option in PHENIX, with subsequent manual building and refinement cycles, maintaining NCS restraints, performed using *Coot* and PHENIX. The electron density maps showed clear density in the metal-binding site at the histidine brace (His35 and His144, molecules A-H). The data were collected at a wavelength of 0.98 Å, which is below the copper edge. A phased anomalous difference map was calculated using PHENIX, and revealed weak anomalous signals for the metal ions located at the histidine brace, ranging from 0.11-0.25 e/Å3 (5-10 σ), whereas metal ions at Asp60, presumed to be Ca^{2+} from the reservoir solution, did not show any anomalous signal distinguishable from background noise (3.6 σ). Given the known very strong preference for copper over calcium in LPMO10 (21), as well the anomalous signal from the metal in the histidine brace, we modeled a copper ion in all eight molecules. Solvent atoms were added using PHENIX refine and ARP-wARP, which included water molecules in close proximity to the copper ion modeled in the copper-binding site. The final rounds of TLS-refinement, where each molecule was modeled as a TLS group, were performed using REFMAC5, maintaining relaxed NCS restraints, to yield a model with and R_{cryst}/R_{free} value of 19.4 % and 22.2 % respectively.

Determination of the Cell Potential (*E*°). The cell potential for the redox couple LPMO-Cu^{2+}/LPMO-Cu^{+} was determined as described by Aachmann *et al* (21). Oxygen free solutions (50 µL) of 305 or 153 µM *N,N,N',N'*-tetramethyl-1,4-phenylenediamine in its reduced form (TMP_{red}) were mixed with an equal volume of 70 µM Cu^{2+}-charged LPMO in UVettes (Eppendorf) in a Chelex-treated 20 mM MES buffer pH 5.5, and incubated at room temperature (298 K). The absorbance at λ = 610 nm was measured using a Hitachi U-1900 spectrophotometer until the signal became stable (8 minutes). The concentration of TMP_{ox}, which is equal to the concentration of LPMO-Cu^{+} (Eq. **S1**), was calculated from the extinction coefficient of 14.0 mM^{-1} cm^{-1} for TMP_{ox} (22). From the determined concentrations (TMP_{ox} and LPMO-Cu^{+}), the equilibrium constant (*K*) was calculated (Eq. **S2**). In Eq. **S3** the relationship between the free energy change (ΔG_r°), the equilibrium constant (*K*) and the cell potential (*E*°) is shown, where *R* is the gas constant, *T* is the temperature in Kelvin, *n* is the electrons transferred in the reaction, and *F* is the Faraday constant. The cell potential for the LPMO-Cu^{2+}/LPMO-Cu^{+} redox couple

was determined by adding the known cell potential of 273 mV for TMP_{red}/TMP_{ox} (23) to the cell potential of the equilibrium reaction of TMP_{red} and $LPMO_{ox}$ as outlined in Fig. S12 and (21).

$$TMP_{red} + LPMO^{Cu^{2+}} \rightleftarrows TMP_{ox} + LPMO^{Cu^{+}} \quad \textbf{[S1]}$$

$$K = \frac{[TMP_{ox}][LPMO^{Cu^{+}}]}{[TMP_{red}][LPMO^{Cu^{2+}}]} \quad \textbf{[S2]}$$

$$\Delta G_r^\circ = -RT \ln K = -nFE^\circ \quad \textbf{[S3]}$$

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) was used to determine the dissociation constants (K_d) for CelS2 and *Sc*LPMO10B with Cu^{2+} as ligand. The ITC experiments were performed with a VP-ITC system from Microcal (24) by measuring the heat produced by titrating copper to apo-LPMO. 4 µL aliquots (50 in total, 180 s intervals) of Cu^{2+} ($CuSO_4$) at a concentration of 150 µM (CelS2) or 120 µM (*Sc*LPMO10B) were titrated into the reaction cell containing 1.42 mL of 5 µM apo-LPMO at a temperature of 10 °C (283 K) and a stirring speed of 260 rpm. Prior to the experiment, solutions were degassed for 20 min to avoid air bubbles. Enzyme and ligand were prepared in identical 20 mM Chelex-treated MES buffer pH 5.5 and measurements were performed in triplicates.

ITC data were monitored and recorded using the Microcal Origin v.7.0 software accompanying the VP-ITC system (24). All data were corrected for the heat of dilution by subtracting the heat produced by injection of ligand into the reaction cell after completion of the binding reaction. These heats had the same magnitudes as when titrating ligand into buffer alone. Theoretical fits to experimental data were obtained utilizing a non-linear least-squares algorithm for a single-site binding model used by the Origin software that accompanies the VP-ITC system. For all binding reactions, data fitted well to a single-site binding model, yielding the stoichiometry (n), the equilibrium binding association constant (K_a), and the enthalpy change ($\Delta H_r°$) of the reaction. The value of n was found to be between 1.0 and 1.1 per enzyme molecule for all reactions. Thermodynamic parameters, i.e. changes in reaction free energy ($\Delta G_r°$) and entropy ($\Delta S_r°$) as well as the dissociation constant (K_d), were calculated using the following relationships: $\Delta G_r° = -RT\ln K_a = RT\ln K_d = \Delta H_r° - T\Delta S_r°$. Errors in $\Delta H_r°$, K_d, and $\Delta G_r°$ were obtained as SDs of three experiments. Errors in $\Delta S_r°$ and $-T\Delta S_r°$ were obtained as propagation of errors.

Electron Paramagnetic Resonance Spectroscopy. Metal-free full length CelS2 or *Sc*LPMO10B in a concentration of 110 µM were mixed with 100 µM Cu^{2+} in 20 mM Chelex-treated Pipes buffer pH 6.0 for resting state EPR spectroscopy analysis. After addition of Cu^{2+} the samples were frozen in liquid nitrogen and EPR spectra were recorded using a BRUKER EleXsyS 560 SuperX instrument equipped with an ER 4122 SHQE SuperX High-sensitivity cavity and a liquid nitrogen cooled cold-finger. The instrument settings were 0.5 mW microwave power, 5 G modulation amplitude, and a temperature of 77 K. The EasySpin toolbox developed for Matlab was used to simulate and fit EPR spectra (25). To estimate Cu^{2+} content in the samples, double integrals of base line corrected EPR spectra, recorded for the samples and a 100 µM Cu^{2+} standard in 1 M perchloric acid, were compared. The spectra for *Sc*LPMO10B and CelS2 indicated the presence of some free copper. Considering that the ITC measurements clearly show a 1:1 stoichiometry of copper binding and that the quantification of the EPR signals

could account for all added copper, for both LPMOs, the detection of a free copper signal indicates that not all copper remained bound to the protein upon freezing the EPR samples. To compensate for this, the EPR signal of free copper, at 20 % signal intensity, needed to be subtracted before simulation.

Sequence Alignment and Phylogeny.
PyMod (26) was used to make a structure-based sequence alignment (27) of the two *S. coelicolor* LPMOs (*Sc*LPMO10B and CelS2, *Sc*LPMO10C) and the chitin active CBP21 (*Sm*LPMO10A, PDB: 2BEM). MUSCLE (28) was then used to add the sequences of the *T. fusca* LPMOs, E7 (*Tf*LPMO10A) and the catalytic domain of E8 (*Tf*LPMO10B). Phylogenetic analysis was carried out for six *S. coelicolor* LPMOs and a selection of other LPMOs for which literature data on substrate specificity is available, using Phylogeny.fr (29).

Table S1. Structural data processing and refinement statistics.

Dataset	*Sc*LPMO10B Zn-SAD I	*Sc*LPMO10B Zn-complex II	*Sc*LPMO10B Cu-complex III	*Sc*LPMO10C (CelS2) IV
PDB code		4OY8	4OY6	4OY7
Data Collection				
beamline	ID14-4 (ESRF)	ID14-4 (ESRF)	ID29 (ESRF)	ID14.1 (BESSY)
Wavelength (Å)	1.282	1.286	1.300	0.918
Space group	P 3_1 2 1	P 3_1 2 1	P 3_1 2 1	P 2_1 2_1 2_1
Cell dimensions				
a, *b*, *c* (Å)	67.2 67.2 107.2	67.6 67.6 107.3	67.5 67.5 107.2	83.9 122.9 156.0
α, β, γ	90.0 90.0 120.0	90.0 90.0 120.0	90.0 90.0 120.0	90.0 90.0 90.0
Resolution (Å)	50 – 2.14	35 – 1.40	50 – 1.29	47.9 – 1.5
Unique reflections	15,162 (994) [a]	54,273 (6205)	67598 (6449)	254063 (34822)
Multiplicity	15.9 (4.4)	4.5 (2.0)	7.9 (2.4)	3.9 (2.7)
Completeness (%)	93.9 (64.9)	96.0 (76.7)	94.2 (63.2)	99.1 (94.2)
Mean I/ σI	28.1 (4.7)	13.6 (3.6)	13.9 (2.1)	5.9 (1.7)
R_{meas} [b]	0.108 (0.36)	0.08 (0.23)	0.088 (0.509)	0.103 (0.560)
Anomalous completeness	90.3 (50.8)	87.6 (46.5)	88.4 (36.4)	96.7 (80.1)
Anomalous multiplicity	8.7 (2.6)	2.4 (1.3)	4.2 (1.4)	2.0 (1.5)
Refinement Statistics				
R_{cryst} /R_{free} (%) [c]	23.4 (27.6)	12.5 (14.3)	12.4 (14.0)	19.4 (22.2)
R.ms.d bond lengths (Å)	0.0158	0.01	0.0105	0.0102
R.m.s.d angles (°)	1.477	1.400	1.425	1.385
Number of atoms				
Protein	1462	1470	1487	11,962
Solvent	–	261	246	1922
Acetate ions	–	12	8	–
Metal	8	8	11	12
Ramachandran plot [d] (%)	97.28	97.3	97.34	97.83
Most favoured	2.72	2.70	2.66	2.17
Additionally allowed	0.0	0.0	0.0	0.0
Outliers				

[a] Values in parentheses are for the highest resolution shells

[b] R_{meas} as defined by Diederichs & Karplus (30).

[c] $R_{cryst} = \Sigma_{hkl} \parallel F_o \parallel F_c \parallel / \Sigma_{hkl} \parallel F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is calculated from a randomly chosen 5% sample of all unique reflections.

[d] Defined using MolProbity (31).

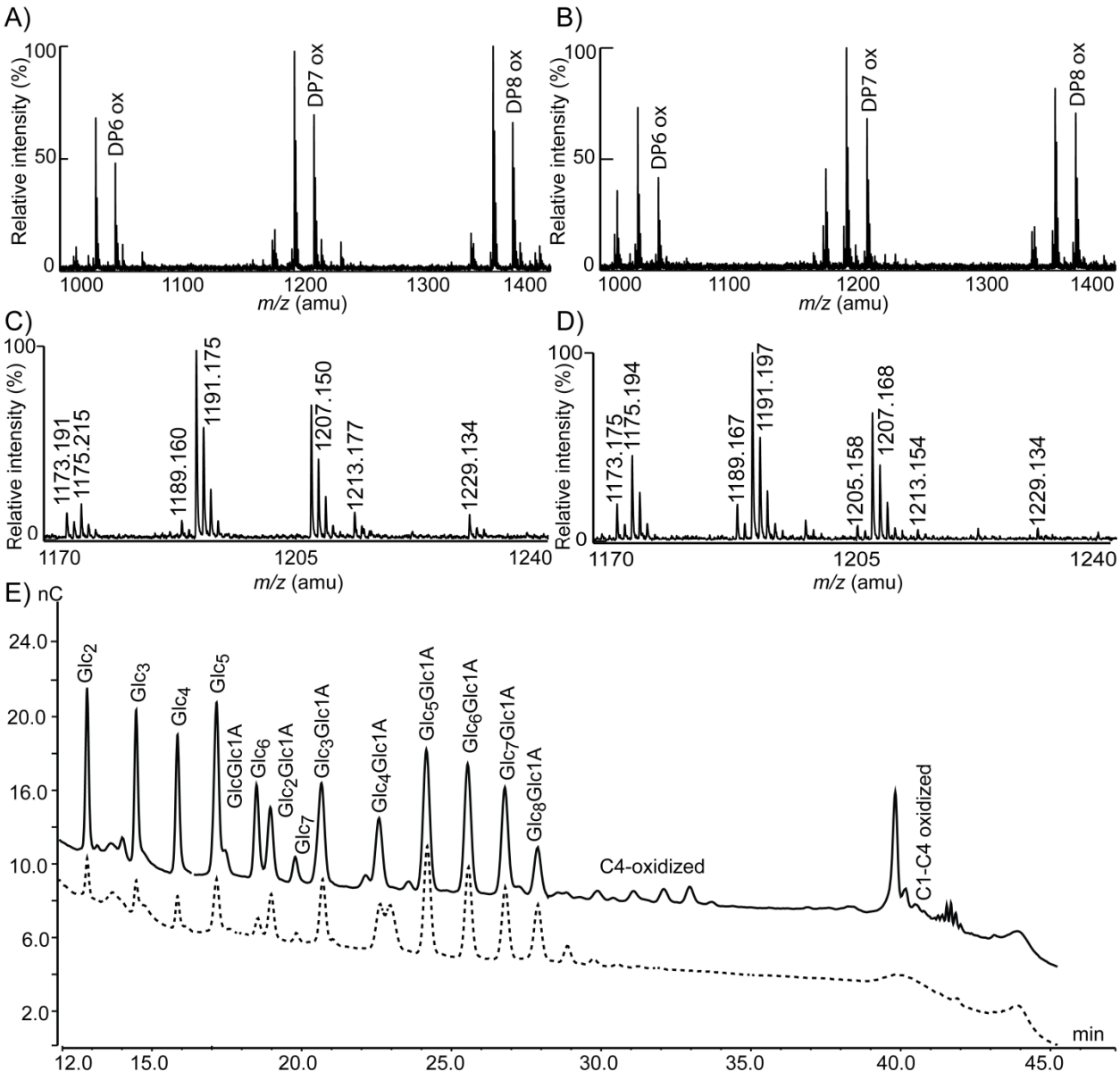


Figure S1. Product profiles from PASC degradation experiments with the *Thermobifida fusca* LPMOs E7 and E8-N, analyzed by MALDI-TOF MS (A-D) and HPAEC (E). MALDI TOF MS of E8-N (A and C) and E7 (B and D) shows that E8-N is a strict C1-oxidizing LPMO as CelS2, whereas E7 produces a mixture of C1-, C4- and double oxidized cello-oligosaccharides as *Sc*LPMO10B. Panels C) and D) are close ups of the heptamer ion cluster with *m/z* values corresponding to: 1173, sodium adduct of lactone or ketoaldose; 1175, sodium adduct of native Glc_7; 1189, potassium adduct of lactone or ketoaldose or sodium adduct of double oxidized heptamer; 1191, sodium adduct of aldonic acid or potassium adduct of native Glc_7 or sodium adduct of gemdiol [4-ketoaldose + water (32)]; 1205, potassium adduct of double oxidized sugar; 1207, potassium adduct of aldonic acid or gemdiol form of the 1189 species; 1213, sodium adduct of the aldonic acid sodium salt; 1229, sodium adduct of the aldonic acid potassium salt. The large signals in panel D at 1173 (relatively stable 4-ketoaldose compared to a relatively unstable lactone, and at 1189 (relatively stable 4-keto form) and at 1205 (double oxidized

product) confirm the C4-oxidizing activity of E7. 100 % relative intensity represents 2.4×10^4 (A, C) and 1.4×10^4 (B, D) arbitrary units (a.u.), respectively. HPAEC chromatograms (E) show results for E7 (solid line) and E8-N (dotted line). Peak assignments are based on native standards and on inferences from previous studies (32, 33); see main text for further explanation.

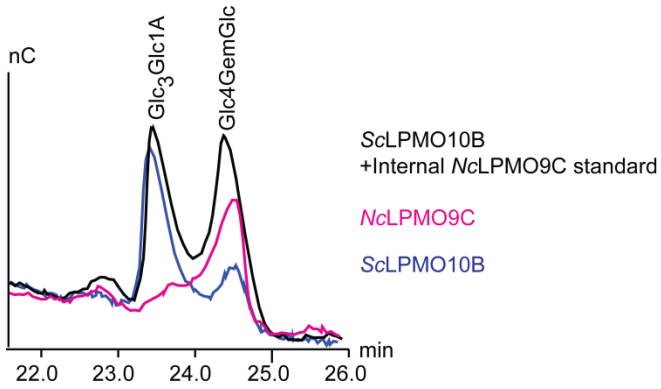


Figure S2. Evidence for C4 oxidation by *Sc*LPMO10B. The chromatogram shows product analysis of the following samples: Magenta, degradation of Glc_5 with *Nc*LPMO9C (sample size 0.2 µL); the product shown has been identified (by MS and, importantly, NMR) as Glc4GemGlc [C4 oxidized dimer; (32)]. Blue: degradation of PASC with *Sc*LPMO10B followed by hydrolysis of soluble products with cellobiohydrolase (sample size 2 µL). Black: mixture of 0.2 µL of the *Nc*LPMO9C sample and 1.8 µL of the *Sc*LPMO10B/cellobiohydrolase sample; as expected the black peak assigned as Glc4GemGlc corresponds approximately to the sum of the blue and the magenta peak underneath. The left peak, produced in the *Sc*LPMO10B reaction only is Glc_3Glc1A.

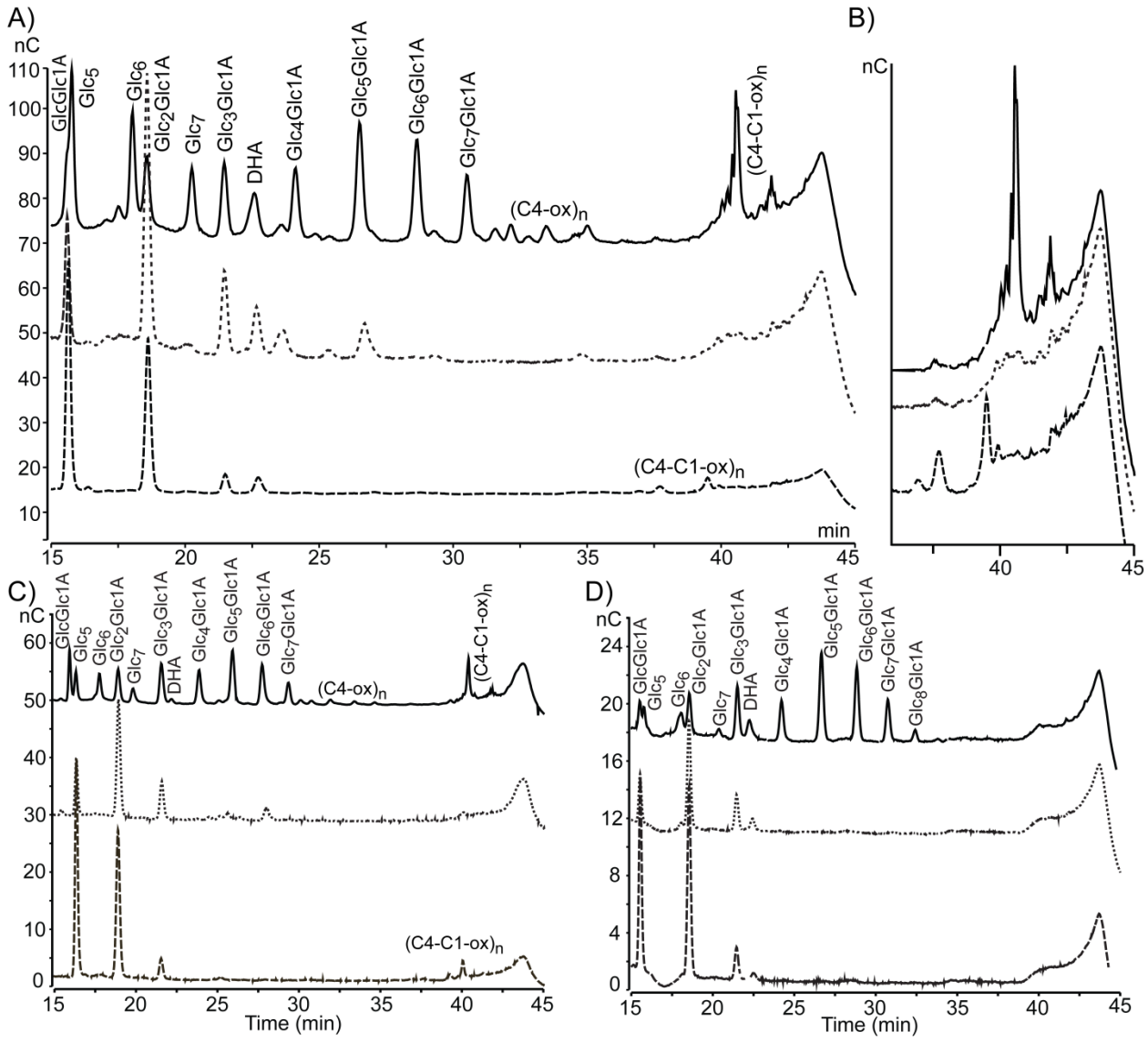


Figure S3. HPAEC chromatograms to verify the occurrence of double oxidized products formed by *Sc*LPMO10B (A-B) and E7 (C). First, *Sc*LPMO10B or E7 was incubated with PASC over night in the presence of ascorbate; subsequently the soluble fraction was taken and analyzed by HPAEC (panel A and C, top chromatogram). The remaining soluble fraction was digested by a cellobiohydrolase for 8 h and analyzed (panel A and C, dotted chromatogram). Finally, *Mt*CDH was added to the resulting mixture of short oligomers followed by incubation for another 16 h to oxidize the reducing ends of the C4-oxidized (i.e. non reducing end-oxidized), regenerating double oxidized species (panel A and C, lower chromatogram). Panel (B) shows a zoom in on the double oxidized products that elute late in the gradient and that are not well resolved. Panel D shows the same experiment for CelS2 (negative control), showing that this enzyme only produces C1-oxidized products (hence no difference is observed between the two lower chromatograms). DHA, dehydroascorbate.

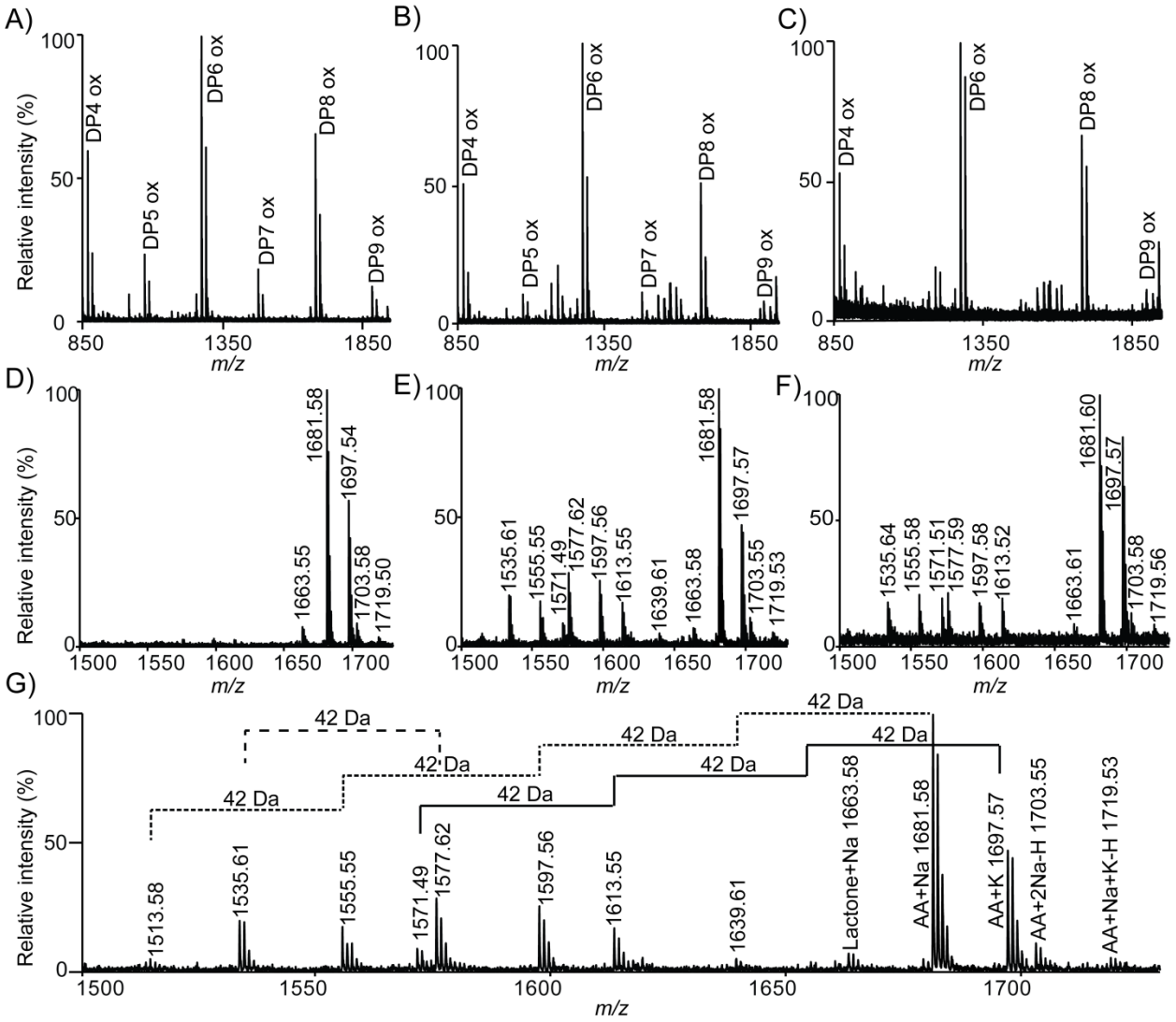


Figure S4. MALDI-TOF MS spectra showing LPMO activity on squid pen β-chitin for *S. marcescens* CBP21 (A, D), *Sc*LPMO10B (B, E), and E7 (C, F). Panels A-C show oxidized products ranging from DP 4-9, whereas panels D-F show a zoom in on the octamer ion cluster. Panel G shows more details of the *Sc*LPMO10B octamer ion cluster revealing the presence of deacetylated products (which are observed for *Sc*LPMO10B and E7 but hardly for CBP21). AA stands for the aldonic acid of the chito-octamer ($GlcNAc_7GlcNAc1A$). 100 % relative intensity represents 1.7×10^4 (A, B), 0.7×10^4 (C), 1.1×10^4 (D), 0.8×10^4 (E; G), and 0.5×10^4 (F) arbitrary units (a.u.), respectively.

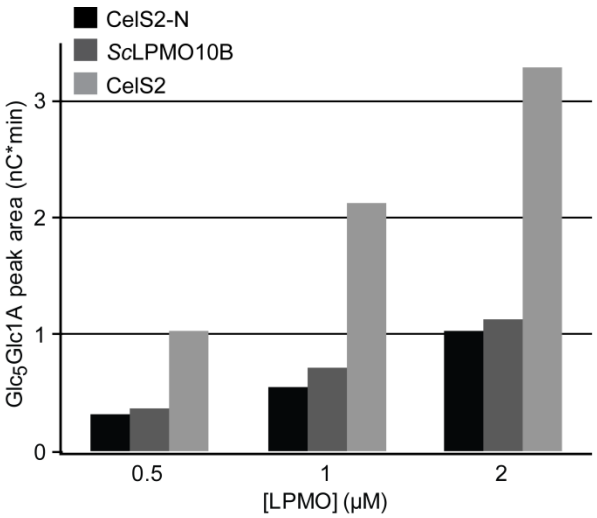


Figure S5. Dose-response effects for CelS2-N (black), *Sc*LPMO10B (dark grey) and full length CelS2 (light grey). Products were generated by incubation of PASC (2 g/L) with the indicated LPMO concentrations, in a 20 mM ammonium acetate buffer pH 6.0, at 50 °C, followed by cellobiohydrolase digestion of the soluble products. The reductant concentration was set to a ratio of 1:2000 enzyme/reductant in all samples.

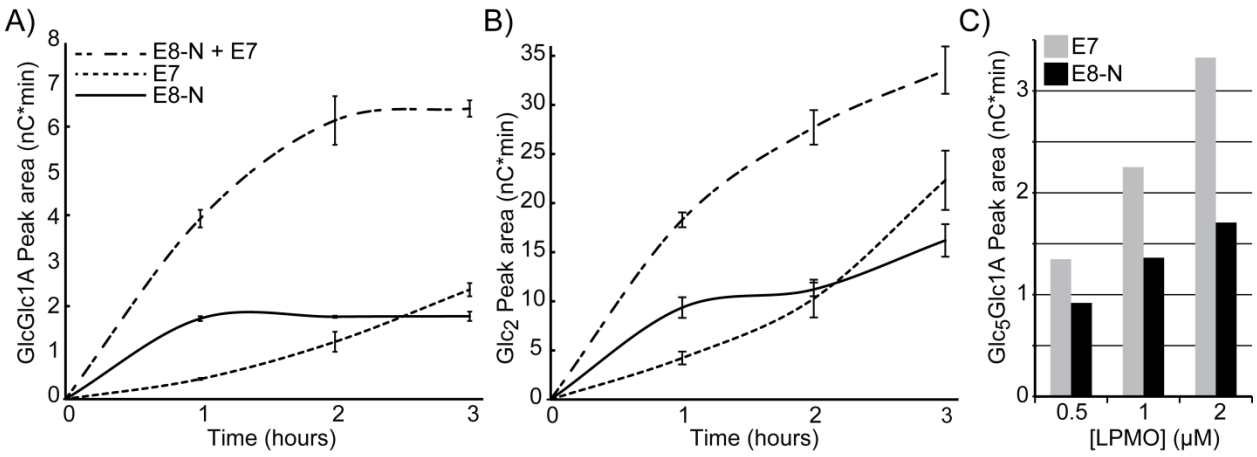


Figure S6. Time course of products released from PASC by the *T. fusca* LPMO pair. The figures show (A) cellobionic acid and (B) cellobiose obtained after incubation of 1 μM LPMO (E8-N or E7) or a mixture of 0.5 μM E8-N and 0.5 μM E7 with 2 g/L PASC and 2 mM ascorbic acid in 20 mM ammonium acetate buffer pH 6.0, at 50°C for three hours, followed by cellobiohydrolase hydrolysis of the soluble products. Standard deviations were calculated based on three independent reactions. Panel C) provides dose-response data for E7 (grey) alone and E8-N (black) alone at the indicated concentrations, using the same conditions. The reductant concentration was set to a ratio of 1:2000 enzyme/reductant in all samples.

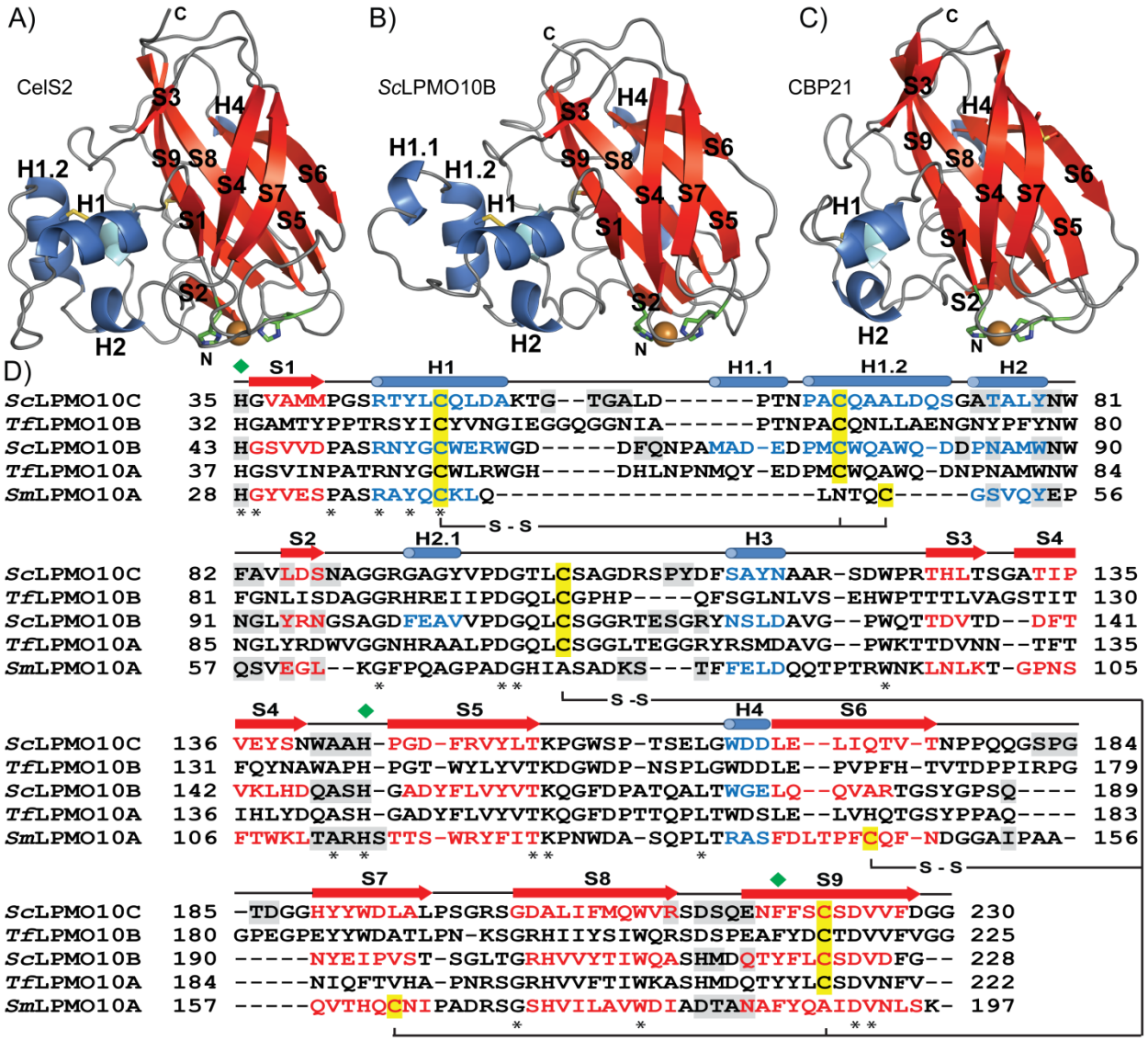


Figure S7. Structural overview of cellulose oxidizing LPMOs from *S. coelicolor* and comparison with the chitin active LPMO from *S. marcescens*. (A-C) Cartoon representations; secondary structure elements are shown in red, for β-strands, and blue, for α-helices. Metal ions are shown as orange spheres, with coordinating histidine side chains shown as green sticks. (A) C1-oxidizing CelS2 (*Sc*LPMO10C) (B) C1/C4-oxidizing *Sc*LPMO10B; (C) chitin-active CBP21 (*Sm*LPMO10A; 2BEM). Note that helix H3 (cyan) is hidden behind helix H1 in all three structures, and that helix H2.1 in *Sc*LPMO10B is hidden behind the β-sheets. (D) Structure-based sequence alignment made with PyMOD, including the sequences of the related *T. fusca* cellulose active LPMOs E7 (*Tf*LPMO10A; C1/C4-oxidizing) and the catalytic domain of E8 (*Tf*LPMO10B; C1-oxidizing). Cysteines involved in disulfide bridges are highlighted in yellow. Known secondary structures are indicated as colored text, β-strands (red) or α-helices (blue) and residues at the putative binding surface are highlighted in grey (for the proteins with known structures only). Asterisks indicate fully conserved residues, and green diamonds indicate amino acids coordinating the copper ion.

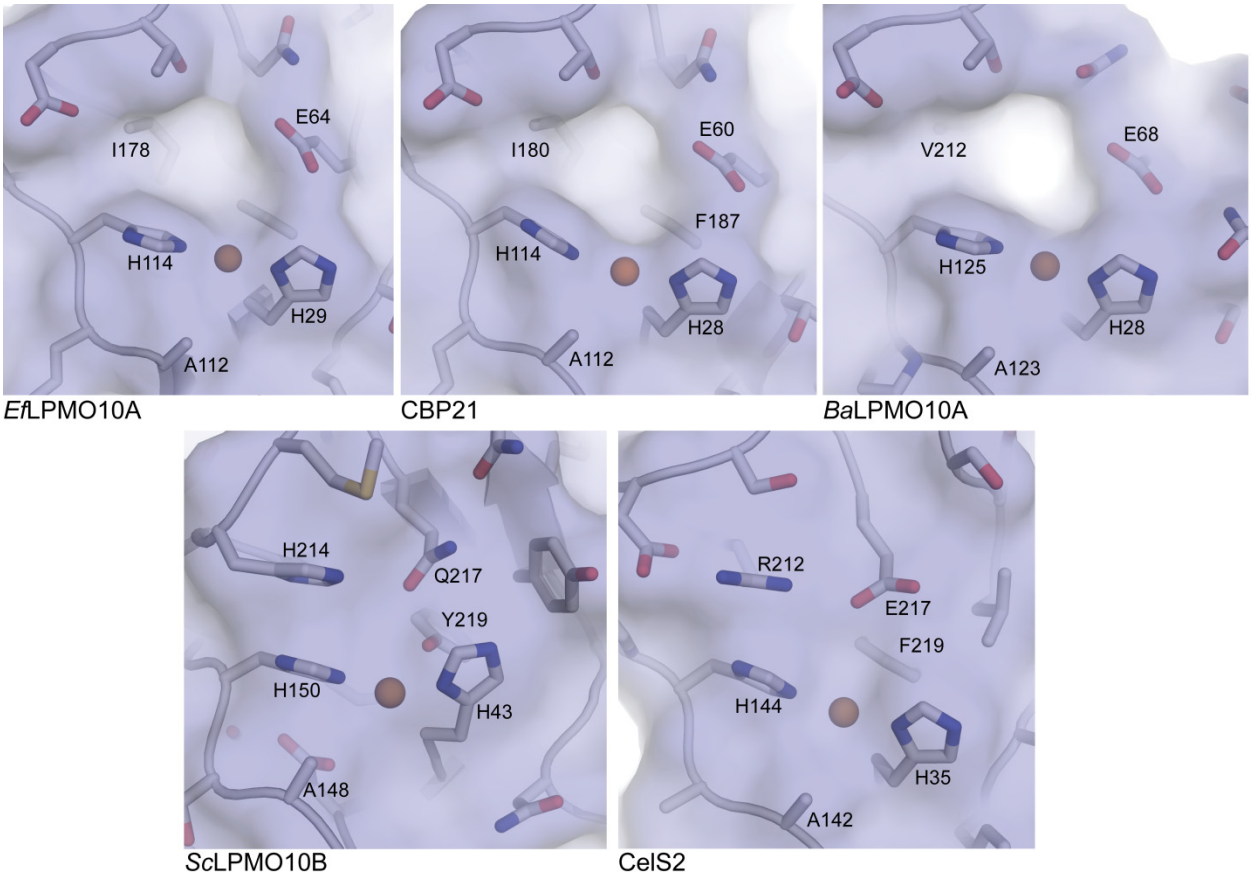


Figure S8. Conserved cavities on the substrate binding surface of chitin active LPMO10s. The picture shows the surfaces of three chitin-active LPMOs (*Ef*LPMO10A, CBP21 (*Sm*LPMO10A), and *Ba*LPMO10A; no activity data have been published for the latter but is known that this protein has high affinity for chitin substrates (34), and two cellulose-active LPMOs (*Sc*LPMO10B and CelS2). The cellulose-active LPMOs have a positively charged side chain (Arg212 in CelS2 and His214 in *Sc*LPMO10B) that fills the cavity present in the chitin-active LPMO10s. The latter have a hydrophobic side chain in the bottom of the cavity (Ile178 in *Ef*LPMO10A, Ile180 in CBP21 and Val189 in *Ba*LPMO10A). Protein surfaces are shown in a transparent representation of the molecular surface. Amino acid side chains and the protein main chain are shown in stick and cartoon representation, respectively. The copper ions are shown as golden colored spheres.

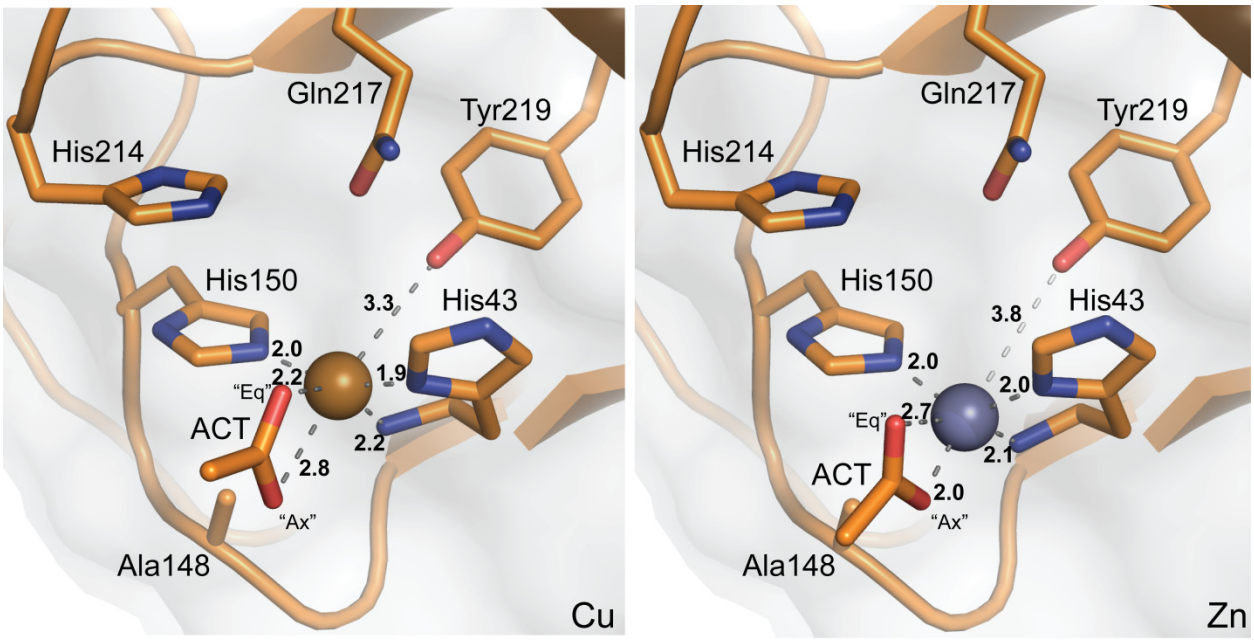


Figure S9. Active site of *Sc*LPMO10B with bound copper (left panel) or zinc (right panel). The two oxygen moieties of the acetate ion (ACT) are bound in the solvent-facing axial position (labeled "Ax") and in a slightly skewed equatorial position (labeled "Eq"). The structural data demonstrate the capability of *Sc*LPMO10B to bind a ligand in the solvent-facing axial position. Amino acid side chains and the protein main chain are shown in orange colored stick and cartoon representation, respectively. The copper and zinc ions are shown as golden and grey colored spheres, respectively. The acetate ion is shown in stick representation with red colored oxygen atoms and orange colored carbon atoms. Distances between atoms are indicated by dashed lines.

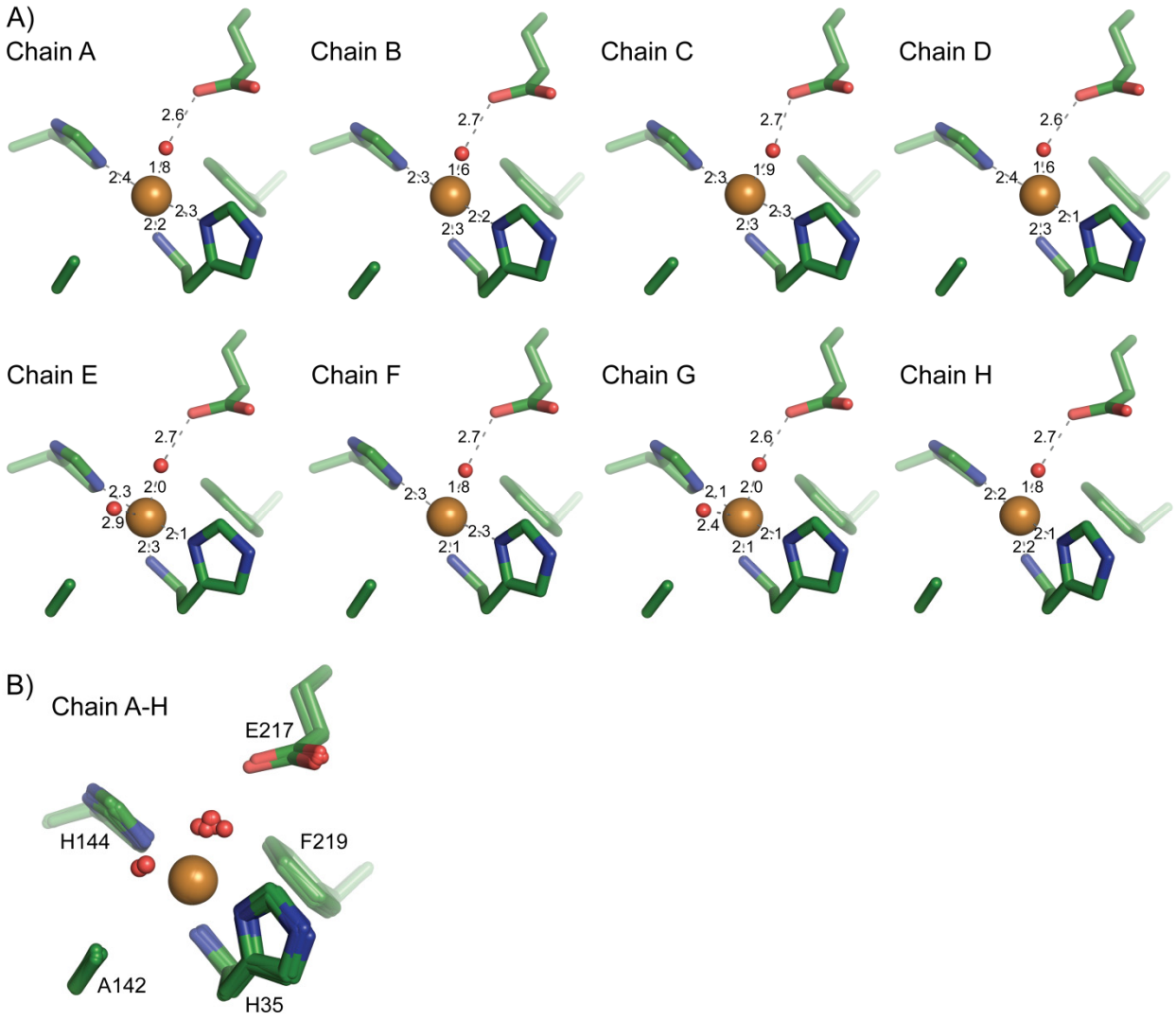


Figure S10. Active sites in all eight CelS2 molecules in the asymmetric unit. Panel (A) shows the individual sites, whereas panel (B) shows a superposition. Amino acid side chains are shown in stick representation. The water molecules and copper ions are shown as red and golden spheres, respectively.

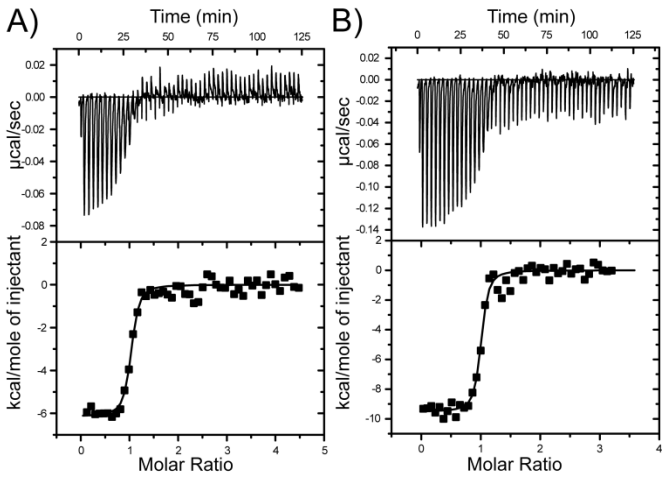


Figure S11. Thermograms (upper) and binding isotherms with theoretical fits (lower) for A) titration of 150 µM Cu^{2+} to 5 µM apo-CelS2 and B) titration of 120 µM Cu^{2+} to 5 µM apo-*Sc*LPMO10B in MES buffer pH 5.5 at 10 °C.

I) $TMP_{red} + LPMO\text{-}Cu^{2+} \rightleftharpoons TMP_{ox} + LPMO\text{-}Cu^{+}$ $TMP_{ox} + e^{-} \longrightarrow TMP_{red}$	CelS2 K= 0.3, $E°$ = -31 mV $E°$ = 273 mV	*Sc*LPMO10B K= 0.43, $E°$ = -22 mV $E°$ = 273 mV
$LPMO\text{-}Cu^{2+} + e^{-} \longrightarrow LPMO\text{-}Cu^{+}$	$E°$ = 242 mV	$E°$ = 251 mV
II) $Cu^{2+} + e^{-} \longrightarrow Cu^{+}$ $LPMO\text{-}Cu^{+} \longrightarrow LPMO\text{-}Cu^{2+} + e^{-}$	$E°$ = 160 mV $E°$ = –242 mV	$E°$ = 160 mV $E°$ = –251 mV
$Cu^{2+} + LPMO\text{-}Cu^{+} \rightleftharpoons Cu^{+} + LPMO\text{-}Cu^{2+}$	$E°$ = –0.082 V, $\Delta G_r°$ = 1.9 kcal/mol	$E°$ = –0.091 V, $\Delta G_r°$ = 2.1 kcal/mol
III) $Cu^{2+} + LPMO\text{-}Cu^{+} \rightleftharpoons Cu^{+} + LPMO\text{-}Cu^{2+}$ $LPMO\text{-}Cu^{2+} \rightleftharpoons LPMO + Cu^{2+}$	$E°$ = –0.082 V, $\Delta G_r°$ = 1.9 kcal/mol K_d = 31 nM, $\Delta G_r°$ = 9.7 kcal/mol	$E°$ = –0.091 V, $\Delta G_r°$ = 2.1 kcal/mol K_d = 12 nM, $\Delta G_r°$ = 10.3 kcal/mol
$LPMO\text{-}Cu^{+} \rightleftharpoons LPMO + Cu^{+}$	$\Delta G_r°$ = 11.6 kcal/mol, K_d = 1.1 nM	$\Delta G_r°$ = 12.4 kcal/mol, K_d = 0.3 nM

Figure S12. Calculation of the dissociation constant for the LPMO-Cu^{+} complex at 10 °C (283 K). The constant for binding of Cu^{+} can be calculated by combining three thermodynamic relations (21). First, I) the cell potential for LPMO-Cu^{2+}/LPMO-Cu^{+} was measured experimentally by determining the equilibrium constant from the electron transfer reaction between the mediator TMP_{red}/TMP_{ox} and LPMO-$Cu^{2+/}$LPMO-Cu^{+} and obtained from the relation $RT\ln K = nFE°$. Second, II) the $E°$ derived from relation I) is combined with the known cell potential for reduction of Cu^{2+} in aqueous conditions to yield the free energy change ($\Delta G_r°$) for the reduction of aqueous Cu^{2+} by LPMO-Cu^{+} using the relation $\Delta G_r° = -nFE°$. Third, III) by combining the $\Delta G_r°$ for binding of Cu^{2+} derived from ITC (Table 1) with the $\Delta G_r°$ for the electron transfer between aqueous Cu^{2+} and LPMO-Cu^{+}, the dissociation of LPMO-Cu^{+} can be calculated using the relation $\Delta G_r° = RT\ln K_d$.

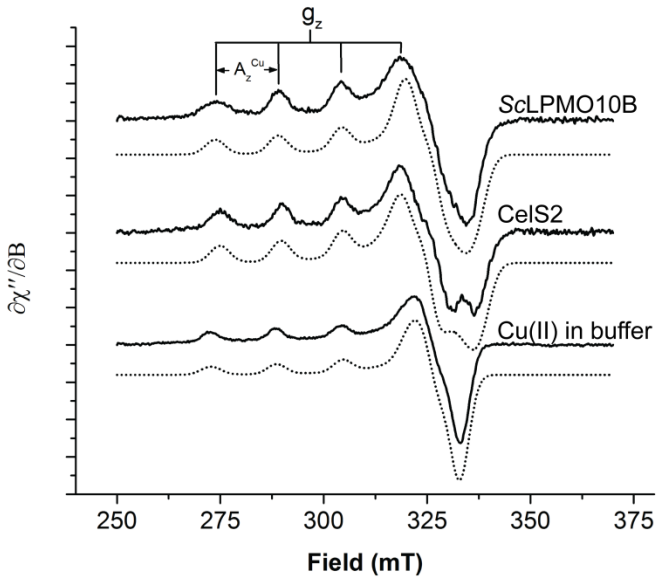


Figure S13. X-band EPR spectra (solid lines) with simulations (dotted lines) for *Sc*LPMO10B, full length CelS2 and Cu(II) in buffer. The EPR spectra were recorded at 77 K using a microwave power of 0.5 mW.

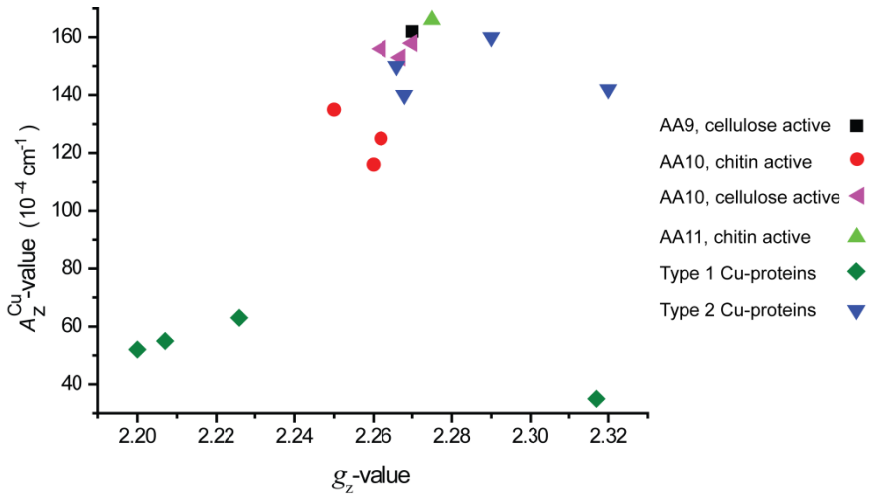


Figure S14. A Peisach – Blumberg plot of different LPMOs compared to selected typical type 1 and type 2 copper proteins. Shown are: chitin-active *Sm*LPMO10A (CBP21) (35), *Bl*LPMO10A (35) and presumably chitin-active *Ba*LPMO10A (36) (●); cellulose-active *Ta*LPMO9A (37) (■), *Tf*LPMO10B (E8-N) (35), *Sc*LPMO10B and *Sc*LPMO10C (CelS2) (35) (▲); chitin-active *Ao*LPMO11 (38) (▲); the type 2 proteins bovine erythrocyte superoxide dismutase (39), pig plasma benzylamine oxidase (40), pig kidney diamine oxidase (41), spinach diphosphate carboxylase (42) (▼); the type 1 proteins horseradish umecyanin (43), plastocyanin (44), pseudoazurin (45), and cucumber basic protein (46) (♦).

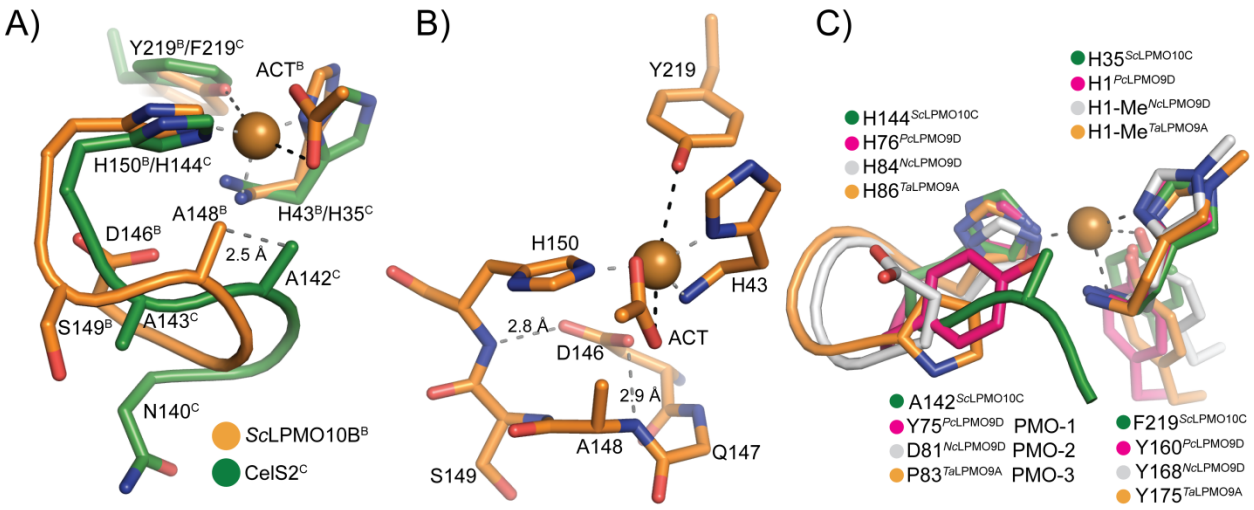


Figure S15. Position of the conserved active site alanine. (A; identical to Fig. 5 in the main manuscript). Structural superposition of CelS2 (green) and *Sc*LPMO10B (orange) active sites (superposed using the "histidine brace" and the metal ion), with side chains shown in stick representation and back bone shown as cartoon. Black and grey dashed lines indicate the axially and equatorially coordinated ligands, respectively. The acetate ion coordinated in the solvent-facing axial position is labeled "ACT" the picture does not show the skewed equatorial contact between the acetate and the copper; see Fig. S9). Note the 2.5 Å relative sift in the position of Ala142/148 between CelS2 and *Sc*LPMO10B; in the former, the alanine (Ala142) is much closer to the acetate oxygen that is bound in the axial position in *Sc*LPMO10B. Panel (B) shows the active site of *Sc*LPMO10B in stick representation, illustrating the hydrogen bond interactions between the side chain of Asp146 and main chain nitrogen atoms. Panel (C) shows a structural superposition of CelS2 (green color) with LPMO9s with C1 (*Pc*LPMO9D in pink color), C4 (*Nc*LPMO9D in white color) and mixed C1/C4 (*Ta*LPMO9A in orange color) activity. The structural superposition was based on the histidine brace and the metal ion. The only other structure of an LPMO9 with presumed C1 activity [inferred for phylogenetic clustering; (47)], *Tt*LPMO9E (48) shows a tyrosine in the same position as in *Pc*LPMO9D. Side chains are shown in stick representation and the back bone is shown as cartoon. The copper ions are shown as a golden colored sphere in all panels.

References

1. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VGH, van Aalten DM (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* 280(12):11313-11319.
2. Manoil C, Beckwith J (1986) A Genetic Approach to Analyzing Membrane-Protein Topology. *Science* 233(4771):1403-1408.
3. Wood TM (1988) Preparation of Crystalline, Amorphous, and Dyed Cellulase Substrates. *Method Enzymol* 160:19-25.
4. Harreither W, et al. (2011) Catalytic properties and classification of cellobiose dehydrogenases from ascomycetes. *Appl Environ Microbiol* 77(5):1804-1815.
5. Vaaje-Kolstad G, et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330(6001):219-222.
6. 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.
7. Kabsch W (2010) Xds. *Acta Crystallogr D* 66(2):125-132.
8. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62(1):72-82.
9. Matthews BW (1968) Solvent Content of Protein Crystals. *J Mol Biol* 33(2):491-497.
10. Kantardjieff KA, Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* 12(9):1865-1871.
11. Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D* 66(4):479-485.
12. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of *Coot*. *Acta Crystallogr D* 66(4):486-501.
13. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171-1179.
14. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* 67(4):355-367.
15. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D* 67(4):271-281.
16. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53(3):240-255.
17. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66(2):213-221.
18. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* 68(4):352-367.
19. Lamzin VS, Wilson KS (1993) Automated refinement of protein models. *Acta Crystallogr D* 49(1):129-147.
20. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D* 63(1):32-41.
21. Aachmann FL, Sørlie M, Skjåk-Braek G, Eijsink VGH, 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(46):18779-18784.
22. Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. *Anal Biochem* 287(1):118-125.
23. Liu Y, Seefeldt LC, Parker VD (1997) Entropies of redox reactions between proteins and mediators: the temperature dependence of reversible electrode potentials in aqueous buffers. *Anal Biochem* 250(2):196-202.
24. Wiseman T, Williston S, Brandts JF, Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179(1):131-137.
25. Stoll S, Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J Magn Reson* 178(1):42-55.
26. Bramucci E, Paiardini A, Bossa F, Pascarella S (2012) PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics* 13 Suppl 4:S2.
27. Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 11(9):739-747.
28. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.
29. Dereeper A, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465-4699.
30. Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* 4(4):269-275.
31. Chen VB, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* 66(1):12-21.
32. Isaksen T, et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem* 289(5):2632-2642.
33. Forsberg Z, et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20(9):1479-1483.
34. Chu HH, Hoang V, Hofemeister J, Schrempf H (2001) A *Bacillus amyloliquefaciens* ChbB protein binds beta- and alpha-chitin and has homologues in related strains. *Microbiology* 147(7):1793-1803.
35. Forsberg Z, et al. (2014) Comparative Study of Two Chitin-Active and Two Cellulose-Active AA10-Type Lytic Polysaccharide Monooxygenases. *Biochemistry* 53(10):1647-1656.
36. Hemsworth GR, et al. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc* 135(16):6069-6077.
37. Quinlan RJ, et al. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A* 108(37):15079-15084.
38. Hemsworth GR, Henrissat B, Davies GJ, Walton PH (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* 10(2):122-126.
39. Rotilio G, et al. (1971) Studies of the metal sites of copper proteins. Ligands of copper in hemocuprein. *Biochemistry* 10(4):616-621.

40. Buffoni F, Corte LD, Knowles PF (1968) The nature of copper in pig plasma benzylamine oxidase. *Biochem J* 106(2):575-576.
41. Mondovi B, et al. (1967) Diamine oxidase from pig kidney. Improved purification and properties. *J Biol Chem* 242(6):1160-1167.
42. Wishnick M, Lane MD, Scrutton MC, Mildvan AS (1969) The presence of tightly bound copper in ribulose diphosphate carboxylase from spinach. *J Biol Chem* 244(20):5761-5763.
43. Stigbrand T, Malmstrom BG, Vanngard T (1971) On the state of copper in the blue protein umecyanin. *FEBS Lett* 12(5):260-262.
44. Gewirth AA, Cohen SL, Schugar HJ, Solomon EI (1987) Spectroscopic and Theoretical-Studies of the Unusual Electron-Paramagnetic-Res Parameters of Distorted Tetrahedral Cupric Sites - Correlations to X-Ray Spectral Features of Core Levels. *Inorg Chem* 26(7):1133-1146.
45. Xie XJ, et al. (2009) A variable temperature spectroscopic study on *Paracoccus pantotrophus* pseudoazurin: Protein constraints on the blue Cu site. *J Inorg Biochem* 103(10):1307-1313.
46. LaCroix LB, et al. (1998) Spectroscopic and geometric variations in perturbed blue copper centers: Electronic structures of stellacyanin and cucumber basic protein. *J Am Chem Soc* 120(37):9621-9631.
47. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc* 136(2):562-565.
48. 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.