Structural diversity of lytic polysaccharide monooxygenases

Gustav Vaaje-Kolstad[1]*, Zarah Forsberg[1], Jennifer S.M. Loose[1], Bastien Bissaro[1,2], and Vincent G. H. Eijsink[1]

[1]Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences (NMBU), 1432 Ås, Norway

[2]INRA, UMR792, Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

*Corresponding author: Vaaje-Kolstad, Gustav (gustav.vaaje-kolstad@nmbu.no), P.O. Box 5003, Department of Chemistry, Biotechnology, and Food Science, The Norwegian University of Life Sciences (NMBU), 1432 Ås, Norway

Abstract

Lytic polysaccharide monooxygenases (LPMOs) catalyze the oxidative cleavage of glycosidic bonds and represent a promising resource for development of industrial enzyme cocktails for biomass processing. LPMOs show high sequence and modular diversity and are known, so far, to cleave insoluble substrates such as cellulose, chitin and starch, as well as hemicelluloses such as beta-glucan, xyloglucan and xylan. All LPMOs share a catalytic histidine brace motif to bind copper, but differ strongly when it comes to the nature and arrangement of residues in the substrate-binding surface. In recent years, the number of available LPMO structures has increased rapidly, including the first structure of an enzyme-substrate complex. The insights gained from these structures is reviewed below.

Introduction

Lytic polysaccharide monooxygenases (LPMOs; also called PMOs by some) represent a unique group of copper-dependent enzymes that perform catalysis on crystalline surfaces, oxidizing ordered polysaccharide chains in e.g. cellulose and chitin [1,2•,3•,4•]. Based on sequence similarity, LPMOs are classified in four families in the auxiliary activities of the CAZy database (AA9, AA10, AA11 and AA13; [5]). Most LPMOs characterized to date display relatively flat substrate binding-surfaces [6•,7••,8••] that are thought to interact with the flat surfaces of crystalline substrates. However, as discussed below, some LPMOs are also capable of cleaving soluble polysaccharides.

The reaction mechanism of LPMOs is still unclear, but several plausible scenarios have been suggested [4,9,10,11], as recently reviewed [12,13]. A shared view is that the resting redox state of the LPMO copper center is Cu(II) that undergoes an initial reductive activation step to Cu(I), which allows the enzyme to subsequently activate dioxygen. Then, the redox state alternates between Cu(II) and Cu(I) along the reaction pathway, depending on which mechanism is considered. The mechanisms entail hydrogen abstraction from one of the carbons in the scissile glycoside bond (C1 or C4 in the case of cellulose), followed by hydroxylation of the resulting substrate radical, which then leads to destabilization of the glycosidic linkage and bond cleavage via an elimination reaction [4•,10]. The reaction requires two electrons delivered by an external

electron donor (Figure 1), which may be of an enzymatic or non-enzymatic nature; the impact of these various electron donors on LPMO activity is currently receiving considerable attention [14•,15,16]. Cellulose-active LPMOs show different regioselectivity, producing either C1 oxidized products (i.e. lactones, that spontaneously convert to aldonic acids), or C4 oxidized products (i.e. ketones that spontaneously convert to gemdiols), or a mixture of the two (Figure 1).

The solvent-exposed active site consists of two fully conserved histidines, one of which is the N-terminal residue. The two histidine side chains and the N-terminal amino group coordinate a copper ion in an arrangement called the histidine brace ([3•]; Figure 2). In fungal LPMO, the N-terminal histidine is post translationally methylated at the Nε2 (Fig. 2b), but the significance of this modification for enzyme function is not known. The coordination sphere of the copper varies between LPMOs and is related to the copper oxidation state. Due to X-ray photoreduction, most LPMO crystal structures display the reduced state, where the copper is coordinated by three nitrogen ligands (from the two histidine side chains and the N-terminal amino group) in a T-shaped geometry [17,18] (Figure 2). Use of low radiation dosages during data collection of LPMO10s have showed that in the Cu(II) state, the copper has five ligands organized in a trigonal bipyramidal geometry [18,19]. In LPMO9s, -11s and -13s, the copper is associated with a somewhat distant buried tyrosine and the oxidized state could thus be considered to have six ligands that coordinate the copper in an octahedral geometry (Figure 2b). In many, but not all LPMO10s, this tyrosine is a phenylalanine, at about 3.5 Å from the copper (Figure 2a), which cannot be considered a true copper ligand.

Structural diversity

Before the discovery of LPMO activity in 2010, only three structures of these enzymes had been determined (CBP21 [20], *Hj*GH61 [21] and *Tt*GH61 [22]). Today there are more than 20 unique LPMO structures deposited in the protein data bank, spanning all four LPMO families. The structural diversity of LPMOs becomes visible when clustering the enzyme structures based on structural similarity (Figure 3).

Common to all LPMOs is a slightly distorted Fibronectin-like/ Immunoglobulin-like β-sandwich core structure consisting of two β-sheets comprising seven or eight β-strands in total (Figure 3)

and the catalytic histidine brace involved in copper coordination (Figure 2). Structural diversity is generated by the helices and loops that connect the core β-strands, giving rise to the variable dimensions and topologies of the substrate-binding surface (Figures 3 and 4). In LPMO10s, most structural variability is found in the region located between β-strand 1 and 3 of the core β-sandwich, also called "loop 2" (abbreviated "L2") or motif 1 [23,24]. A similar highly variable L2 region occurs in LPMO9s between β-strands 1 and 2 [23]. The L2 region consists of varying numbers of loops and short helices, and accommodates one, or in few cases two (for some LPMO9s), surface-located aromatic amino acids. The L2 region is believed to influence substrate recognition and specificity as it constitutes large parts of the substrate-binding surface and shows great variation [23,24,25,26••,27•,28•]. Some LPMO9s (cluster 7) show a characteristic insertion between β-strands 3 and 4, referred to as L3 [28], that interacts with the L2 loop.

Variation of the substrate-binding surface on the opposite side of L2 includes regions referred to as LS (loop short) and LC (long C-terminal loop) [13•,23]. The LS and LC regions are exclusive to LPMO9s and LPMO13s (i.e. clusters 5-8 in Figure 3 and 4) and often contain one or more solvent-exposed aromatic residues that have their side chains positioned flat on the binding surface and which could be involved in substrate binding (Figure 4) [23,25]. Notably, while the substrate-binding surface of LPMOs are generally thought to be "flat", they do show topological variability that could be related to substrate specificity (e.g. [26••]). Different from most other LPMOs, the starch-degrading LPMO13s (cluster 5) possesses a shallow groove that includes the active site and which could accommodate an amylose chain [29••].

Figure 4 shows that there is high sequence variation in the substrate-binding surfaces of LPMOs, even within the clusters shown in Figure 3d and also close to the catalytic center. This variation suggests that LPMOs may display a wide variety of substrate specificities, not only in terms of what glycosidic bond they break, but also in terms of varying substrate topologies, as they may occur in different types of plant cell walls. Notably, plant cell walls comprise complex composite structures and even a "homogenous" compound such as cellulose may occur in various crystal forms. There are several studies showing that LPMOs act on, and even may be optimized for composite polysaccharide structures [31•,32].

A closer look at the catalytic centers of LPMOs in Figure 4 reveals a structurally highly conserved glutamate at approximately 5 Å from the active site copper, in all chitin-active LPMOs (LPMO10 and 11) and all cellulose-active C1-specific LPMO10s. This glutamate, located in the red region for clusters 1 and 3 and in the black region for clusters 2, 4 and 9 (Figure 4, marked by arrow), points towards the copper active site and its presence does not seem correlated with substrate-specificity. Thus, this residue may be involved in the LPMO general mechanism. Interestingly, a highly conserved glutamine, always located in the black region, is found at an approximately equivalent position in all other LPMOs. Experiments show that this glutamate [33] and glutamine [22] are essential for catalysis.

Structural basis of substrate specificity

Since the original discovery of LPMO activity towards chitin, LPMOs with activities towards various plant polysaccharides have been described, including cellulose [3•,4•,34], soluble cello-oligosaccharides [35•], xyloglucan and other β-glucans containing β-1,4-linkages [36], starch [29••,37•] and xylan [31•]. LPMO9s that act on β-glucan hemicelluloses vary in terms of specificity and also differ in the extent to which they can handle substitutions of the xyloglucan backbone [32,38,39]. Early work on CBP21, the chitin-active LPMO10 from *S. marcescens*, has shown that substrate binding primarily involves polar interactions and includes a contribution from the single aromatic amino acid in the substrate-binding surface (located in the L2 loop) ([6•,20]; Figure 5).

In contrast to bacterial LPMO10s, the fungal LPMO9s display more than one aromatic amino acid on the substrate-binding surface (Figure 3, clusters 6-8), at least one or two in the LC loop (Figure 4, left-hand yellow-shaded areas) and optionally one or two in the L2 loop. Such arrangements are often found in proteins that bind to carbohydrates, where the interaction between the substrate and the protein is mediated by CH-π stacking interactions. The spacing between the aromatic residues on the substrate binding surface is equal to one, two to three times the distance separating the monosaccharides in a polysaccharide chain, suggesting that these aromatic amino acids interact with the substrate [13•,23,25].

The majority of LPMOs hitherto characterized are only active towards insoluble substrates, making in-depth investigation of enzyme-substrate interactions challenging. However, the discovery of LPMOs active on soluble substrates [35•] paved the way for use of X-ray crystallography and NMR to study the structural basis of substrate specificity. A milestone in LPMO research was reached by the X-ray crystallographic structures of an LPMO9 in complex with cello-oligosaccharide substrates [8••]. The structure revealed that the enzyme-ligand interactions are dominated by polar interactions between the enzyme and the substrate and that the N-terminal histidine stacks with the +1 sugar (Figure 5a,b). An NMR study on substrate binding by a very similar enzyme showed that the interacting area comprised the histidine brace (Figure 5c) as well as neighboring residues Ala80 and His155 [7••]. Docking studies constrained by the NMR data for cellohexaose and interaction studies with other substrates (xyloglucan and xyloglucan oligosaccharides), revealed additional interacting residues (His64 and Tyr204; Figure 5c). Notably, His155 is a highly conserved residue among LPMO9s (Figure 4, black region). Interestingly, the combined results of the recent X-ray crystallographic and NMR studies on enzyme-substrate interactions show that chloride and cyanide, both potential mimics of a negatively charged reactive oxygen species, enhance substrate binding. Cyanide is a known copper-binding analogue of superoxide [40] and its effect on substrate affinity suggests that the formation of a copper-oxygen species couple contributes to the LPMO-substrate interaction, rather than the copper itself. While EPR studies have shown that substrate-binding has influence on copper coordination [8••,28•], Courtade *et al.* have shown that the presence of copper(II) alone hardly affects substrate affinity.

Despite recent progress, the structural determinants of LPMO substrate specificity remain largely unknown. There are data indicating that the L3 loop may play a role in activity on xyloglucan [7••], but recently xyloglucan activity was detected in an LPMO lacking this loop [41••]. Surface topological features could discriminate between chitin and cellulose in LPMO10s [26••], and may be important for activity of starch [29••]. Carbohydrate-binding modules (CBMs) could obviously also play a role ([42]; see “Modular diversity”). However, the fact is that we really do not know; there are no examples of engineered LPMOs with changed substrate specificity.

LPMO stability

It is known that copper-binding stabilizes the LPMO structure [17], but apart from this, little is known about the structural determinants of LPMO stability. Importantly, inspection of available kinetic data, and our own unpublished results, show that LPMO stability deserves attention, because the enzymes tend to be unstable under process conditions [15,43]. A recent study by Loose *et al.* clearly showed that the LPMO rapidly loses activity under certain conditions [15]. These authors showed that the nature of the reductant affects the rate of activity loss, suggesting that the interplay between the redox systems in the reaction influences LPMO stability. Considering the very powerful redox species generated in the LPMO active site [9,12•,13•], it is conceivable that protection against destructive oxidative side reactions has been a driving force in LPMO evolution and could explain some of the active site features of today's LPMOs. Destructive oxidative side reactions may be reduced by binding to the substrate, as suggested by the observation that LPMOs generate H_2O_2 in the absence of substrate [35•,44]. Thus, CBMs could indirectly play a role in determining LPMO stability. Some of the apparent activity changes that have been observed upon removing or adding CBMs [26••,42] are perhaps related to stability effects of changes in substrate-affinity.

Structural basis of oxidative regioselectivity

LPMOs acting on chitin (LPMO10 and 11; clusters 1-3 and 9) and starch (LPMO13, cluster 5) have only been shown to oxidized the C1-position. On the other hand, LPMO9s include strict C1-, strict C4- and mixed C1/C4-oxidizers sometimes referred to as Type 1, Type 2 and Type 3 [27•]. For LPMO10s active on cellulose, only strict C1- and mixed C1/C4-oxidizers have been described [26••]. The surface analysis of Figure 4 shows a plethora of structural variations that could affect the precise positioning of either the substrate or the reactive oxygen species, with a possible effect on which of the glycosidic carbons is attacked. So far, there is hardly any experimental data addressing the structural basis of oxidative regioselectivity. However, Vu *et al.* have shown that an LPMO9 mutant lacking a small helix in the L2 region (containing a conserved Tyr; see cluster 8 in Figure 4), generated no C4-oxidized products compared to the C1/C4-oxidizing wild type [27•], indicating the importance of this region for C4-specificity for cluster 8 LPMOs.

Importantly, structural comparisons of the structures of seven well characterized cellulose-active LPMO9s [28•] and two characterized cellulose-active LPMO10s [26••] have revealed a potentially important structural correlation: In C1-oxidizing LPMOs, access to the surface-exposed axial copper coordination site seems somewhat restricted, whereas there do not seem to be any restrictions in strictly C4-oxidizing LPMOs. An intermediate form, in terms of accessibility, is observed in LPMOs with a mixed C1/C4 oxidation pattern. The validity of this intriguing correlation still needs experimental validation.

Modular diversity

Auxiliary modules of carbohydrate-active enzymes may modulate substrate specificity and/or substrate affinity. LPMOs commonly contain additional CBMs [44,45]. Interestingly, it seems that single domain LPMOs that target insoluble substrates have evolved strong substrate binding abilities [20,46,47,48], whereas LPMO modules containing appended CBMs have lost this ability [45,49]. It is well documented that CBMs contribute to substrate binding by LPMOs and that truncation of these domains leads to reduced enzyme performance [26••,28•,42].

As previously noted, the structural diversity of the catalytic LPMO modules is large (Figure 3 and 4). This variation is further expanded by the large variation in auxiliary modules appended to LPMOs [50]. Inspection of the Pfam database families harboring LPMO9s (Pfam ID PF03443) and LPMO10s (Pfam ID PF03067) reveals that both families contain a large variety of auxiliary modules and combinations of these. Closer inspection of the sequences shows that the majority of the appended modules likely promote binding to polysaccharides such as cellulose (CBM1 and CBM2), chitin (CBM1, CBM2, CBM5/12, CBM14 and CBM73) or starch (CBM20). Interestingly, several appended modules with no obvious link to carbohydrates are also observed (e.g. dopamine-monooxygenase-like domains, phosphotyrosine-binding domain, amino transferase domains, epoxide hydrolase domains, heme-binding domains etc.), suggesting that some LPMOs may have roles LPMOs other than polysaccharide oxidation.

The only multidomain LPMO that has been structurally characterized is the *Vibrio cholerae* colonization factor *Vc*LPMO10B (also known as GbpA), which, next to the N-terminal LPMO domain contains three additional modules: one CBM5/12 chitin-binding module and two modules putatively involved in binding to bacterial outer cell wall structures [51]. *Vc*LPMO10B is an

example of one of several LPMOs for which there are indications that they play a role in virulence and infection.

Concluding remarks

The LPMO field has seen a remarkable growth in structures since the discovery of this enzyme activity in 2010. This has exposed large structural diversity and enabled comparative structural analysis [52•]. Interestingly, it seems that only the catalytic copper center is 100% conserved, whereas the second coordination sphere of the copper and the rest of the substrate-binding region are quite diverse. It may be that this variation relates to yet-to-be-discovered variation in substrate preference, perhaps not at the level of the type of glycosidic bond that is cleaved, but rather in terms of the context of this bond (crystalline, amorphous, co-polymeric structures, different faces of a crystal). Anno 2016, we are beginning to get a first glimpse of how LPMOs interact with their substrates. Considering the complexity of the substrates, extensive site-directed mutagenesis studies are likely the best way to get insight into the determinants of substrate specificity. Other important enzyme properties could also be addressed by such studies, including oxidative regioselectivity and stability.

Intriguingly, LPMO10s occur in a wide range of organisms, including bacteria, viruses, fungi, slime molds, insects, algae and various sea animals. The question is whether all these LPMOs are involved in catabolism or whether they may have other functions, such as in regulation of hyphal extension (in fungi), in cell wall remodeling (in fungi and algae) or in moulting (in insects). Moreover, the large variety of domains appended to LPMOs suggest that new LPMO substrates will be identified in future research. Such substrates may be related to host-pathogen barriers since several multimodular LPMOs have been identified as virulence factors, e.g. *Vc*LPMO10B from *Vibrio cholerae* [53] and *Lm*LPMO10A from *Listeria monocytogenes* [54]. The putative roles of LPMO activity in infection are unknown, but certainly of great interest.

In conclusion, it is safe to say that, despite major progress in recent years, research on LPMOs still is in its infancy and much exciting LPMO research is to be expected in the future.

Acknowledgements

This work was supported by the Research Council of Norway (grant 249865 to GV-K and JSML, and grants 214613 & 243663 to VGHE), the French Institut National de la Recherche Agronomique (INRA) and Marie-Curie FP7 COFUND People Programme (AgreenSkills fellowship grant 267196 to BB) and by the Vista program of the Norwegian Academy of Science and Letters (grant 6510, to ZF and VGHE).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest
•• of outstanding interest

1. Eriksson KE, Pettersson B, Westermark U: **Oxidation: an important enzyme reaction in fungal degradation of cellulose**. *FEBS Lett* 1974, **49**:282-285.

2. • Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH: **An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides**. *Science* 2010, **330**:219-222.

The first publication describing an LPMO, its activity, and the need for external electron supply. These authors did not recognize that LPMOs are copper-dependent.

3. • Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JC, Johansen KS, Krogh KB, Jørgensen CI, Tovborg M, Anthonsen A *et al.*: **Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components**. *Proc Natl Acad Sci USA* 2011, **108**:15079-15084.

First detailed structural analysis of the copper-binding site of LPMOs, i.e. the “histidine brace”.

4. • Phillips CM, Beeson WT, Cate JH, Marletta MA: **Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa***. *ACS Chem Biol* 2011, **6**:1399-1406.

This paper was the first to present possible catalytic scenario’s for LPMOs.

5. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B: **Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes**. *Biotechnol Biofuels* 2013, **6**:41.

6. • Aachmann FL, Sørlie M, Skjåk-Bræk G, Eijsink VGH, Vaaje-Kolstad G: **NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions**. *Proc Natl Acad Sci USA* 2012, **109**:18779-18784.

NMR study providing experimental insight into chitin binding by an LPMO. Notably, this is the only study of this kind with insoluble substrate.

7. •• Courtade G, Wimmer R, Røhr AK, Preims M, Felice AK, Dimarogona M, Vaaje-Kolstad G, Sørlie M, Sandgren M, Ludwig R *et al.*: **Interactions of a fungal lytic polysaccharide monooxygenase with beta-glucan substrates and cellobiose dehydrogenase**. *Proc Natl Acad Sci USA* 2016, **113**:5922-5927.

The first NMR study of an LPMO9 and the interaction with its soluble substrate. The study shows that interactions with both the substrate and the protein electron donor (CDH) are centered around the copper-binding active site.

8. •• Frandsen KEH, Simmons TJ, Dupree P, Poulsen JCN, Hemsworth GR, Ciano L, Johnston EM, Tovborg M, Johansen KS, von Freiesleben P *et al.*: **The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases**. *Nat Chem Biol* 2016, **12**:298-303.

Crystal structure of a C4-oxidizing LPMO9 in complex with substrate. The structure also shows a chloride ion that is equatorially bound to the copper and that may mimic a reactive oxygen species.

9. Kim S, Ståhlberg J, Sandgren M, Paton RS, Beckham GT: **Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism**. *Proc Natl Acad Sci USA* 2014, **111**:149-154.

10. Beeson WT, Phillips CM, Cate JH, Marletta MA: **Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases**. *J Am Chem Soc* 2012, **134**:890-892.

11. Kjaergaard CH, Qayyum MF, Wong SD, Xu F, Hemsworth GR, Walton DJ, Young NA, Davies GJ, Walton PH, Johansen KS *et al.*: **Spectroscopic and computational insight into the activation of O_2 by the mononuclear Cu center in polysaccharide monooxygenases**. *Proc Natl Acad Sci USA* 2014, **111**:8797-8802.

1 12. • Walton PH, Davies GJ: **On the catalytic mechanisms of lytic polysaccharide**
2 **monooxygenases**. *Curr Opin Chem Biol* 2016, **31**:195-207.
3 Insightful review of possible LPMO catalytic mechanisms, including comparisons with the
4 mechanisms of other copper-catalyzed oxygenations.
5 13. • Beeson WT, Vu VV, Span EA, Phillips CM, Marletta MA: **Cellulose degradation by**
6 **polysaccharide monooxygenases**. *Annu Rev Biochem* 2015, **84**:923-946.
7 A comprehensive review on all aspects of LPMOs.
8 14. • Kracher D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K, Haltrich D,
9 Eijsink VGH, Ludwig R: **Extracellular electron transfer systems fuel cellulose oxidative**
10 **degradation**. *Science* 2016, **352**:1098-1101.
11 In-depth bioinformatics and experimental analysis of various electron-donating systems
12 that fungi may employ when using LPMOs to degrade plant biomass.
13 15. Loose JS, Forsberg Z, Kracher D, Scheiblbrandner S, Ludwig R, Eijsink VGH, Vaaje-Kolstad
14 G: **Activation of bacterial lytic polysaccharide monooxygenases with cellobiose**
15 **dehydrogenase**. *Protein Sci* 2016. DOI: 10.1002/pro.3043
16 16. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken JP, van Berkel WJ,
17 Kabel MA, Gruppen H: **Lytic polysaccharide monooxygenases from *Myceliophthora***
18 ***thermophila* C1 differ in substrate preference and reducing agent specificity**. *Biotechnol*
19 *Biofuels* 2016, **9**:186.
20 17. Hemsworth GR, Taylor EJ, Kim RQ, Gregory RC, Lewis SJ, Turkenburg JP, Parkin A, Davies
21 GJ, Walton PH: **The copper active site of CBM33 polysaccharide oxygenases**. *J Am Chem Soc*
22 2013, **135**:6069-6077.
23 18. Gudmundsson M, Kim S, Wu M, Ishida T, Momeni MH, Vaaje-Kolstad G, Lundberg D,
24 Royant A, Ståhlberg J, Eijsink VGH *et al.*: **Structural and electronic snapshots during the**
25 **transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by**
26 **X-ray photoreduction**. *J Biol Chem* 2014, **289**:18782-18792.
27 19. Gregory RC, Hemsworth GR, Turkenburg JP, Hart SJ, Walton PH, Davies GJ: **Activity,**
28 **stability and 3-D structure of the Cu(II) form of a chitin-active lytic polysaccharide**
29 **monooxygenase from *Bacillus amyloliquefaciens***. *Dalton Trans* 2016. DOI:
30 10.1039/C6DT02793H

1 20. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VGH, van Aalten DMF: **Crystal**
2 **structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21**. *J*
3 *Biol Chem* 2005, **280**:11313-11319.
4 21. Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M: **The first structure**
5 **of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å**
6 **resolution**. *J Mol Bio* 2008, **383**:144-154.
7 22. Harris PV, Welner D, McFarland KC, Re E, Poulsen JCN, Brown K, Salbo R, Ding HS,
8 Vlasenko E, Merino S *et al.*: **Stimulation of lignocellulosic biomass hydrolysis by proteins of**
9 **glycoside hydrolase family 61: structure and function of a large, enigmatic family**.
10 *Biochemistry* 2010, **49**:3305-3316.
11 23. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, Himmel ME, Crowley MF,
12 Horn SJ, Westereng B *et al.*: **Crystal structure and computational characterization of the lytic**
13 **polysaccharide monooxygenase GH61D from the Basidiomycota fungus *Phanerochaete***
14 ***chrysosporium***. *J Biol Chem* 2013, **288**:12828-12839.
15 24. Book AJ, Yennamalli RM, Takasuka TE, Currie CR, Phillips GN, Fox BG: **Evolution of**
16 **substrate specificity in bacterial AA10 lytic polysaccharide monooxygenases**. *Biotechnol*
17 *Biofuels* 2014, **7**:109.
18 25. Li X, Beeson WT, Phillips CM, Marletta MA, Cate JH: **Structural basis for substrate**
19 **targeting and catalysis by fungal polysaccharide monooxygenases**. *Structure* 2012, **20**:1051-
20 1061.
21 26. •• Forsberg Z, Mackenzie AK, Sørlie M, Røhr AK, Helland R, Arvai AS, Vaaje-Kolstad G,
22 Eijsink VGH: **Structural and functional characterization of a conserved pair of bacterial**
23 **cellulose-oxidizing lytic polysaccharide monooxygenases**. *Proc Natl Acad Sci USA* 2014,
24 **111**:8446-8451.
25 Crystal structures of two cellulose-active LPMO10s with different oxidative
26 regioselectivity. Experimental evidence of synergy between two LPMOs is also provided.
27 27. • Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA: **Determinants of regioselective**
28 **hydroxylation in the fungal polysaccharide monooxygenases**. *J Am Chem Soc* 2014, **136**:562-
29 565.

This study provides indications that the L2 loop in LPMO9s affects oxidative regioselectivity. Figure S4 provides product profiles for a large series of LPMOs with different oxidative regioselectivities.

28. • Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Várnai A, Røhr AK, Payne CM, Sørlie M, Sandgren M *et al.*: **Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity**. *J Biol Chem* 2015, **290**:22955-22969.

First crystal structure of a xyloglucan-active LPMO with an extensive discussion of possible structural determinants of substrate specificity. See also Courtade *et al.*, 2016.

29. •• Lo Leggio L, Simmons TJ, Poulsen JC, Frandsen KE, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L *et al.*: **Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase**. *Nat Commun* 2015, **6**:5961.

First structure of a starch-active AA13 that reveals topological features that are putatively adapted to starch binding and that are more pronounced than in other LPMOs.

30. Holm L, Rosenström P: **Dali server: conservation mapping in 3D**. *Nucleic Acids Res* 2010, **38**:W545-549.

31. • Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SW, Koetsier MJ, van Berkel WJ, Gruppen H, Kabel MA: **Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase**. *Biotechnol Biofuels* 2015, **8**:101.

First study to demonstrate LPMO activity towards xylan. The activity is shown to be highest towards xylan tethered to cellulose (i.e. a co-polymeric structure).

32. Kojima Y, Várnai A, Ishida T, Sunagawa N, Petrovic DM, Igarashi K, Jellison J, Goodell B, Alfredsen G, Westereng B *et al.*: **Characterization of an LPMO from the brown-rot fungus *Gloeophyllum trabeum* with broad xyloglucan specificity, and its action on cellulose-xyloglucan complexes**. *Appl Environ Microbiol* 2016. DOI: 10.1128/AEM.01768-16

33. Vaaje-Kolstad G, Horn SJ, van Aalten DMF, Synstad B, Eijsink VGH: **The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation**. *J Biol Chem* 2005, **280**:28492-28497.

34. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ, Eijsink VGH: **Cleavage of cellulose by a CBM33 protein**. *Protein Sci* 2011, **20**:1479-1483.

1 35. • Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, Ludwig R, Haltrich
2 D, Eijsink VGH, Horn SJ: **A C4-oxidizing lytic polysaccharide monooxygenase cleaving both**
3 **cellulose and cello-oligosaccharides**. *J Biol Chem* 2014, **289**:2632-2642.
4 This study was the first to demonstrate oxidative cleavage of a soluble substrate (cello-
5 oligosaccharides) by an LPMO.
6 36. • Agger JW, Isaksen T, Várnai A, Vidal-Melgosa S, Willats WG, Ludwig R, Horn SJ, Eijsink
7 VGH, Westereng B: **Discovery of LPMO activity on hemicelluloses shows the importance of**
8 **oxidative processes in plant cell wall degradation**. *Proc Natl Acad Sci USA* 2014, **111**:6287-
9 6292.
10 This article reports LPMO activity towards hemicellulose. Cleaved substrates were
11 identified using a novel analytical approach involving glycan micro array screening.
12 37. • Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA: **A family of starch-active**
13 **polysaccharide monooxygenases**. *Proc Natl Acad Sci USA* 2014, **111**:13822-13827.
14 This paper reports the first LPMOs active on starch and contains a large amount of
15 analytical data substantiating the findings.
16 38. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, Fanuel M, Ropartz
17 D, Rogniaux H, Gimbert I *et al.*: **Substrate specificity and regioselectivity of fungal AA9 lytic**
18 **polysaccharide monooxygenases secreted by *Podospora anserina***. *Biotechnol Biofuels* 2015,
19 **8**:90.
20 39. Nekiunaite L, Petrovic DM, Westereng B, Vaaje-Kolstad G, Hachem MA, Várnai A, Eijsink
21 VGH: ***Fg*LPMO9A from *Fusarium graminearum* cleaves xyloglucan independently of the**
22 **backbone substitution pattern**. *FEBS Lett* 2016, **590**:3346-3356.
23 40. Paci M, Desideri A, Rotilio G: **Cyanide binding to Cu, Zn superoxide dismutase. An NMR**
24 **study of the Cu(II), Co(II) derivative**. *J Biol Chem* 1988, **263**:162-166.
25 41. •• Cannella D, Möllers KB, Frigaard NU, Jensen PE, Bjerrum MJ, Johansen KS, Felby C:
26 **Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme**.
27 *Nat Commun* 2016, **7**:11134.
28 This paper presents interesting observations on how light can be used to fuel LPMOs and
29 demonstrates the LPMO activity can be modulated by several order of magnitude.

1 42. Crouch LI, Labourel A, Walton PH, Davies GJ, Gilbert HJ: **The contribution of non-catalytic**
2 **carbohydrate binding modules to the activity of lytic polysaccharide monooxygenases**. *J Biol*
3 *Chem* 2016, **291**:7439-7449.
4 43. Scott BR, Huang HZ, Frickman J, Halvorsen R, Johansen KS: **Catalase improves**
5 **saccharification of lignocellulose by reducing lytic polysaccharide monooxygenase-**
6 **associated enzyme inactivation**. *Biotechnol Lett* 2016, **38**:425-434.
7 44. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R: **Production of four *Neurospora***
8 ***crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric**
9 **assay**. *Biotechnol Biofuels* 2012, **5**:79.
10 45. Forsberg Z, Nelson CE, Dalhus B, Mekasha S, Loose JS, Crouch LI, Røhr AK, Gardner JG,
11 Eijsink VGH, Vaaje-Kolstad G: **Structural and functional analysis of a lytic polysaccharide**
12 **monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus***. *J Biol*
13 *Chem* 2016, **291**:7300-7312.
14 46. Zeltins A, Schrempf H: **Specific interaction of the *Streptomyces* chitin-binding protein**
15 **CHB1 with alpha-chitin - The role of individual tryptophan residues**. *Eur J Biochem* 1997,
16 **246**:557-564.
17 47. Suzuki K, Suzuki M, Taiyoji M, Nikaidou N, Watanabe T: **Chitin binding protein (CBP21)**
18 **in the culture supernatant of *Serratia marcescens* 2170**. *Biosci Biotechnol Biochem* 1998,
19 **62**:128-135.
20 48. Nakagawa YS, Kudo M, Loose JSM, Ishikawa T, Totani K, Eijsink VGH, Vaaje-Kolstad G:
21 **A small lytic polysaccharide monooxygenase from *Streptomyces griseus* targeting alpha- and**
22 **beta-chitin**. *FEBS J* 2015, **282**:1065-1079.
23 49. Forsberg Z, Røhr AK, Mekasha S, Andersson KK, Eijsink VGH, Vaaje-Kolstad G, Sørlie M:
24 **Comparative study of two chitin-active and two cellulose-active AA10-type lytic**
25 **polysaccharide monooxygenases**. *Biochemistry* 2014, **53**:1647-1656.
26 50. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH: **Novel enzymes for the degradation**
27 **of cellulose**. *Biotechnol. Biofuels* 2012, **5**:45.
28 51. Wong E, Vaaje-Kolstad G, Ghosh A, Hurtado-Guerrero R, Konarev PV, Ibrahim AF, Svergun
29 DI, Eijsink VGH, Chatterjee NS, van Aalten DMF: **The *Vibrio cholerae* colonization factor**
30 **GbpA possesses a modular structure that governs binding to different host surfaces**. *PLoS*
31 *Pathog* 2012, **8**:e1002373.

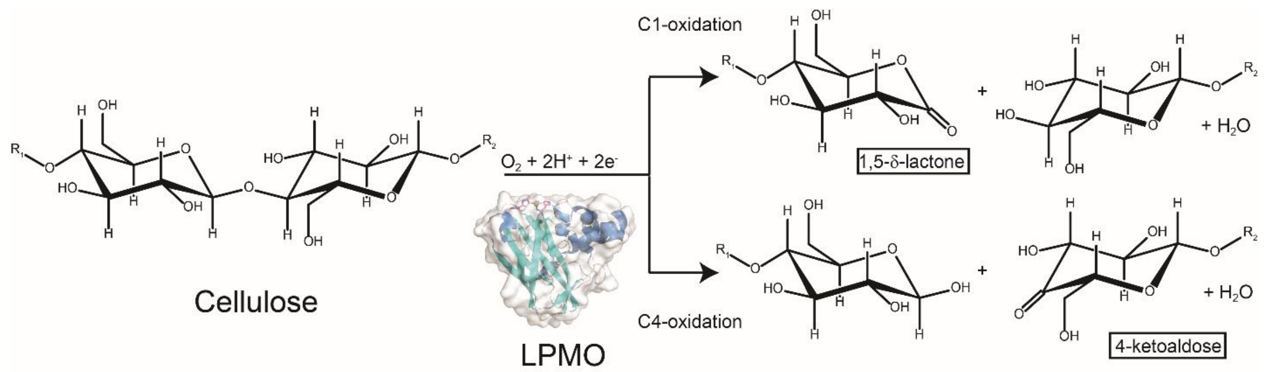
52.• Frandsen KEH, Lo Leggio L: **Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes.** *IUCrJ* 2016, 3. DOI: 10.1107/S2052252516014147

Very recent structure-oriented review providing an overview of the LPMO history and with a detailed listing of all available LPMO structures.

53. Kirn TJ, Jude BA, Taylor RK: **A colonization factor links *Vibrio cholerae* environmental survival and human infection**. *Nature* 2005, **438**:863-866.

54. Paspaliari DK, Loose JS, Larsen MH, Vaaje-Kolstad G: ***Listeria monocytogenes* has a functional chitinolytic system and an active lytic polysaccharide monooxygenase**. *FEBS J* 2015, **282**:921-936.

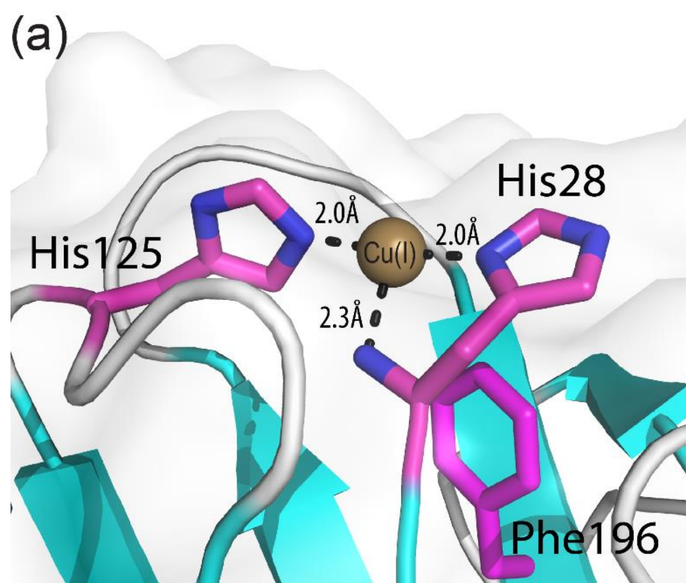
1 # Figure captions

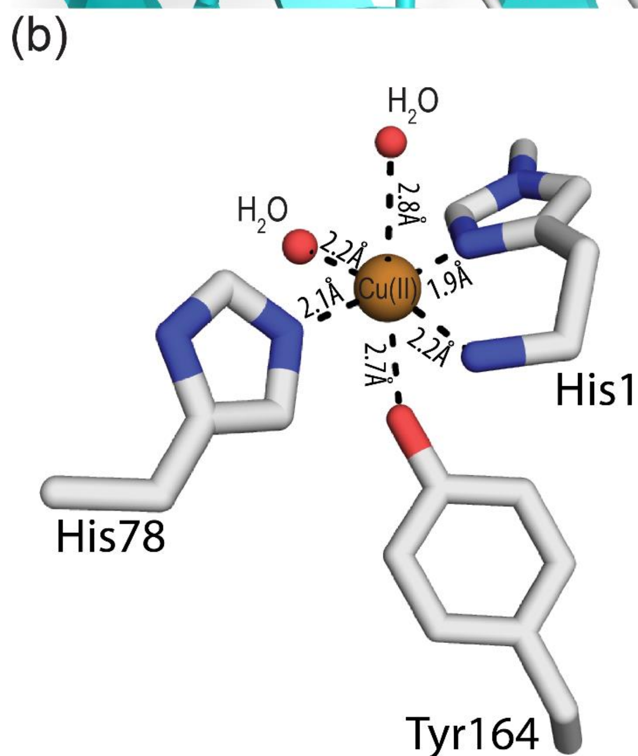


2

3 **Figure 1. Reaction mechanism of LPMOs.** Oxidation of either the C1 or the C4 carbon in
4 cellulose.

5





1

2 **Figure 2. The copper active site of LPMOs.** (a) The solvent exposed copper-containing active
3 site of an LPMO10 (*Ba*LPMO10A) from *Bacillus amyloliquefaciens* (PDB code 2YOX [17]) with
4 copper in its reduced (Cu(I)) state. The accessible surface of the protein is shown by transparent
5 white surface representation. Amino acid side chains are shown in stick representation with
6 magenta colored carbon atoms. The copper ion is shown as a golden sphere. Note the buried
7 phenylalanine close to the copper; in other LPMOs this residue is a tyrosine (see panel b). (b) The

1 active site of an LPMO9 (*Ls*LPMO9A, PDB code 5ACG [8]) in its oxidized state, showing
2 octahedral geometry. Amino acid side chains are shown as yellow sticks and the golden sphere
3 represents the Cu(II) ion. Note that the N-terminal histidine (His1) is methylated at the Nε2
4 nitrogen.

5

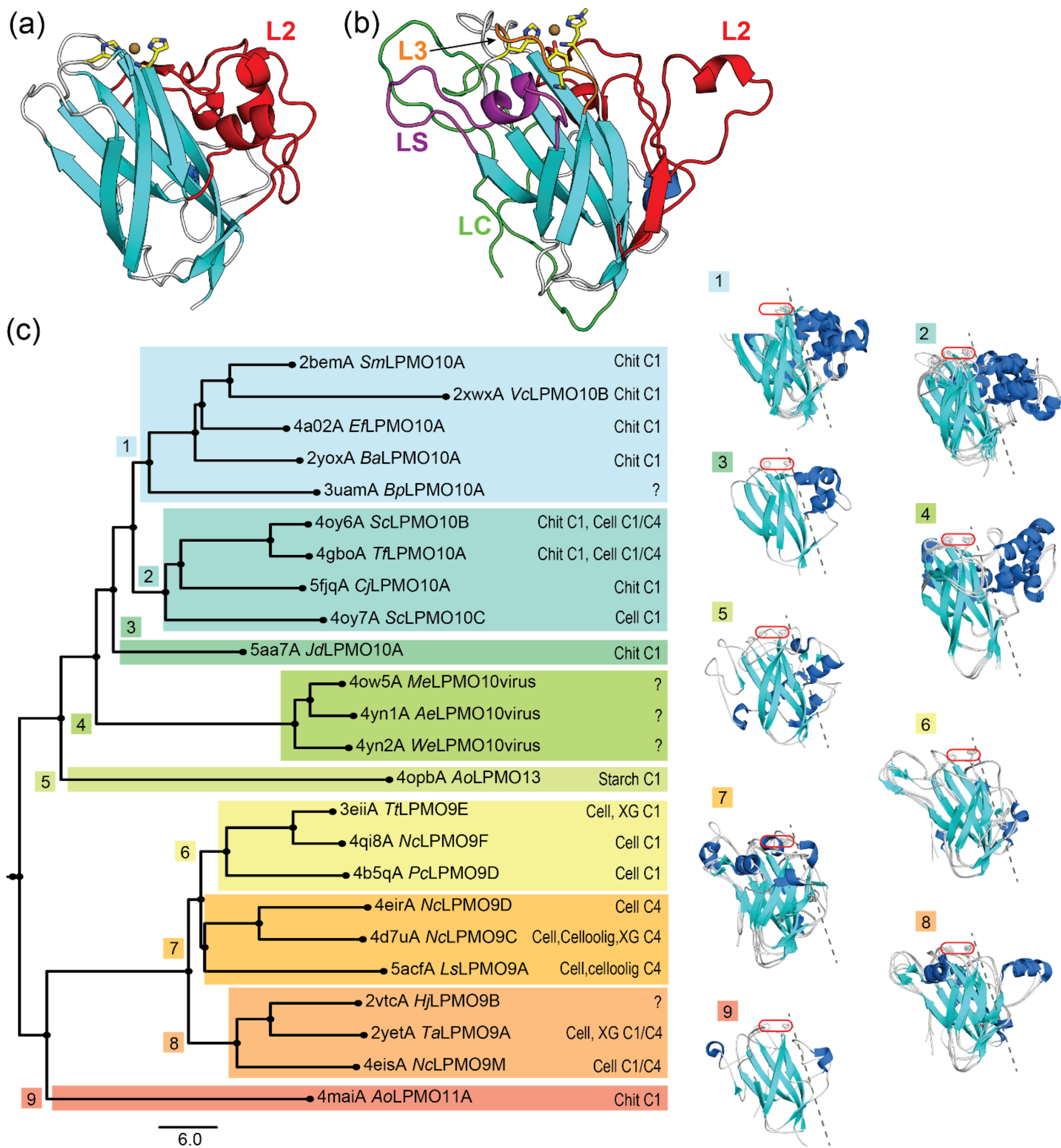


Figure 3. Structural diversity of LPMOs. Panels (a) and (b) show the typical fold of an LPMO10 illustrated by the structure of CBP21 from *Serratia marcescens* (PDB code 2BEM [20]) and an LPMO9 illustrated by *Nc*LPMO9M from *Neurospora crassa* (PDB code 4EIS [25]), respectively. Loops important for forming the substrate-binding surface (L2, LS and LC) are indicated. The *Nc*LPMO9M structure does not contain an L3 loop insert, but the loop hosting this insert in other LPMOs is indicated. (c) Dendrogram showing structural clustering of 24 unique LPMO structures.

Structures are identified by their PDB identifier and the chain ID, followed by the experimentally determined known substrates (note that the absence of a substrate can simply mean that it has never been tested; almost all LPMOs have been tested on crystalline cellulose and chitin). The scale indicates the DALI Z-score. Representatives of each cluster are shown structurally aligned (i.e. in the same orientation) on the right hand side of the dendrogram. Structural clustering was performed using the DALI structural comparison server [30], using the "all against all" option. The location of the two histidine residues of the histidine brace is outlined by a red oval. The dashed line separates the core β-sandwich (left) from the L2 region (right). The experimentally determined substrate is indicate for each enzyme; Cell, cellulose; Celloolig, cello-oligosaccharides; Chit, chitin; XG, xyloglucan. Several of these LPMOs have appeared in the literature under other names, which are given in parenthesis; *Sm*LPMO10A (CBP21), *Vc*LPMO10B (GbpA, *Vc*AA10B), *Ef*LPMO10A (*Ef*CBM33A, *Efa*CBM33), *Ba*LPMO10A (ChbB, *Ba*CBM33), *Tf*LPMO10A (E7), *Sc*LPMO10C (CelS2, *Sc*AA10C), *Ao*LPMO13 (*Ao*(AA13)), *Tt*LPMO9E (*Tt*GH61E), *Pc*LPMO9D (*Pc*GH61D), *Nc*LPMO9D (PMO-2, NCU01050), *Nc*LPMO9C (NCU02916), *Ls*LPMO9A (Ls(AA9)A), *Hj*LPMO9B (EG7, Cel61B), *Ta*LPMO9A (*Ta*GH61A), *Nc*LPMO9M (PMO-3, NCU07898), *Ao*LPMO11 (*Ao*(AA11)).

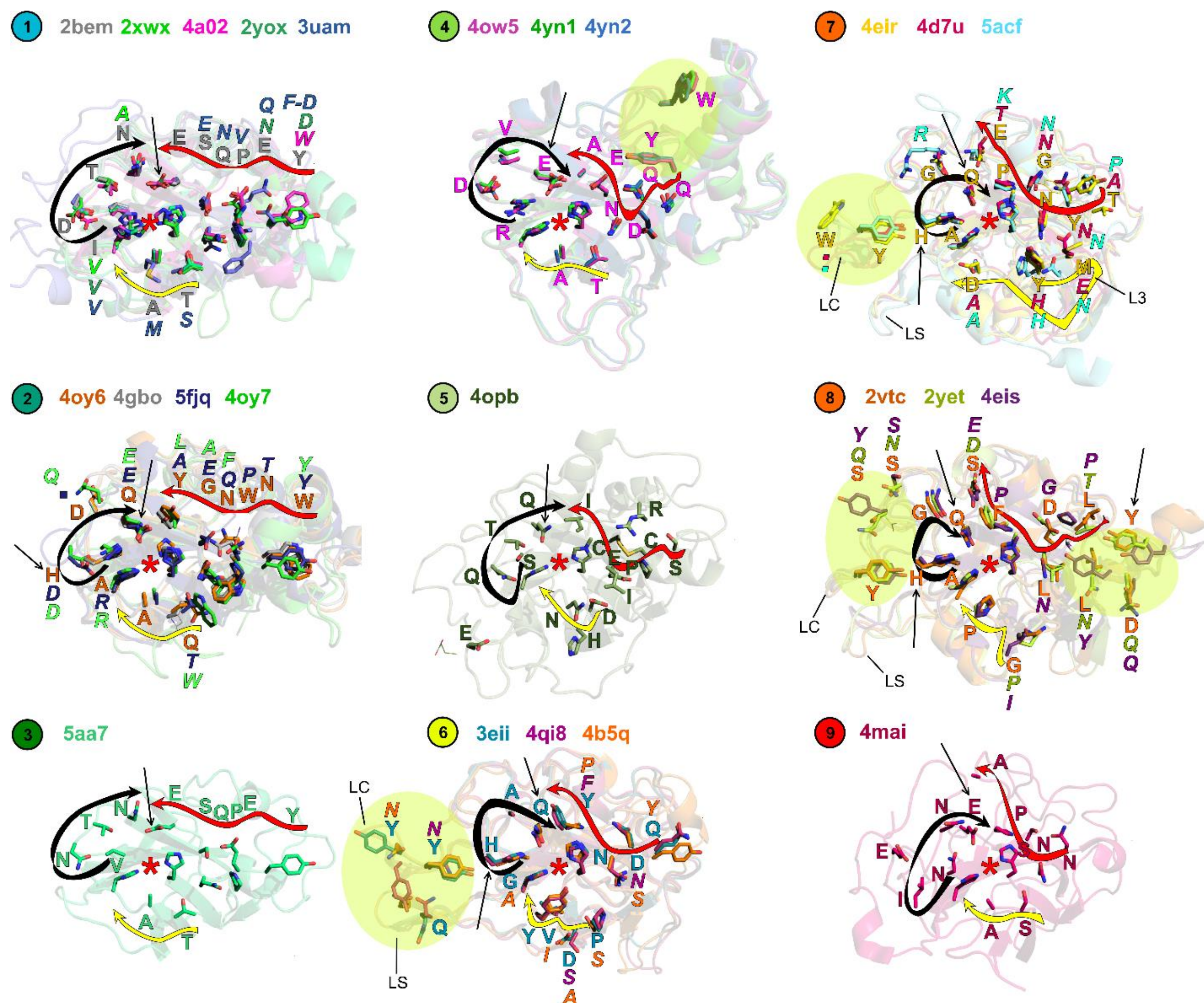


Figure 4. LPMO substrate binding surfaces and their conserved residues. The figure shows a top view of structurally aligned and superimposed LPMOs, grouped according to the clusters defined in Figure 3c. The structures were aligned based on the histidine-brace/copper center, which is represented by a red star. Side chains protruding from the surface are shown as sticks and labeled. Three regions (depicted by black, red and yellow arrows) define the immediate environment of the catalytic center. These regions are globally conserved within each cluster and equivalents can be found in all the clusters. The red region is part of the L2 loop. The yellow region connects β-strands 3 and 4 in the core β-sandwich and contains the second catalytic histidine; some LPMOs have an insertion here, referred to in the text as L3 (only cluster 7). The black region connects the two last β-strands of the β-sandwich in all the nine clusters. Additional conspicuous surface residues more remote from the catalytic centers appear in the yellow-shaded areas that are formed by additional residues in the L2 loop region (only in cluster 8) and/or the LS/LC loops

1 (clusters 6-8, i.e. all LPMO9s). For each labeled side chain, the residue found in the shown
2 structure (the first of the pdb codes) was arbitrarily defined as the reference; alternative residues
3 at this position found in other cluster members (if any) are also indicated using the same color code
4 as for the PDB accession numbers. A dot means that there is no clear structural equivalent. Arrows
5 indicate specific conserved residues discussed in the text.
6

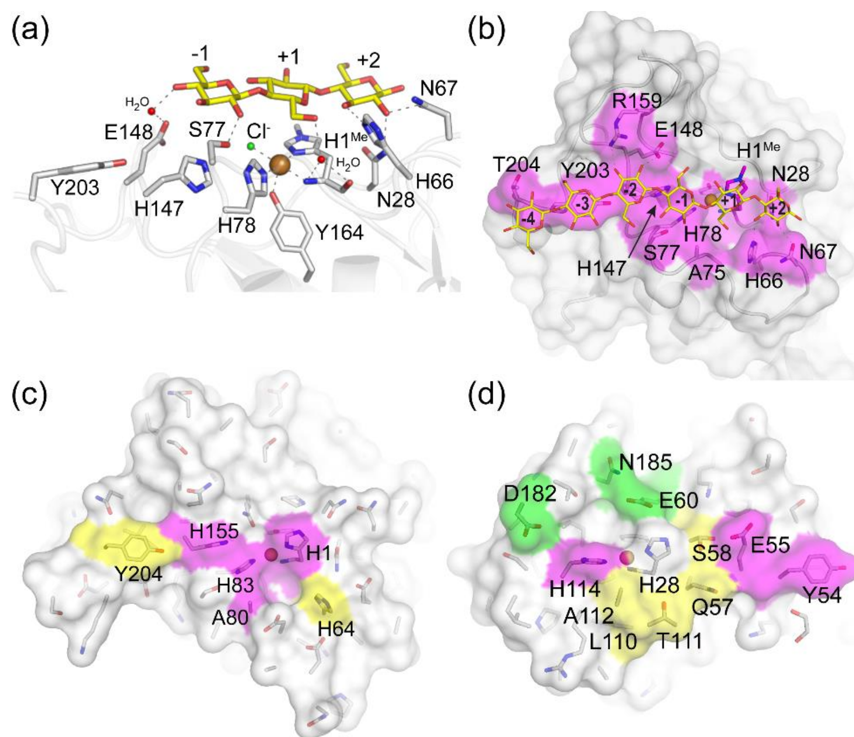


Figure 5. Residues involved in substrate interaction identified by X-ray crystallography (a,b), NMR (c,d), and site-directed mutagenesis (d). (a) *Ls*LPMO9A interacting with cellotriose (PDB code 5ACF [8••]). (b) Top view of *Ls*LPMO9A; residues that interact with cellohexaose are labeled and colored magenta (PDB code 5ACI). (c) Top view of *Nc*LPMO9C (PDB code 4D7U [28•]; residues that show a chemical shift upon addition of substrate are colored magenta and yellow, where the yellow color indicates residues that were more affected by addition of xyloglucan compared to cellohexaose. Note that the NMR analysis is based on exchange of the amide proton (i.e. a backbone proton) only, which reduces the sensitivity of the method (certain side chain substrate interactions, as visible in panel (a) may simply not be detectable by NMR). (d) Top view of CBP21 (PDB code 2BEM [20]); residues that have been shown important for binding to insoluble β-chitin by NMR (yellow), site-directed mutagenesis (green) or both (magenta) are colored. Note that this early study on substrate binding by Aachmann *et al.* (2012) was done with β-chitin, an insoluble and truly crystalline substrate.