

RESEARCH

Open Access



Functional characterization of fungal lytic polysaccharide monoxygenases for cellulose surface oxidation

Yann Mathieu^{1,2†}, Olanrewaju Raji^{3†}, Annie Bellemare⁴, Marcos Di Falco⁴, Thi Truc Minh Nguyen⁴, Alexander Holm Viborg¹, Adrian Tsang⁴, Emma Master^{3,5*} and Harry Brumer^{1,2,6,7,8*}

Abstract

Background Microbial lytic polysaccharide monoxygenases (LPMOs) cleave diverse biomass polysaccharides, including cellulose and hemicelluloses, by initial oxidation at C1 or C4 of glycan chains. Within the Carbohydrate-Active Enzymes (CAZy) classification, Auxiliary Activity Family 9 (AA9) comprises the first and largest group of fungal LPMOs, which are often also found in tandem with non-catalytic carbohydrate-binding modules (CBMs). LPMOs originally attracted attention for their ability to potentiate complete biomass deconstruction to monosaccharides. More recently, LPMOs have been applied for selective surface modification of insoluble cellulose and chitin.

Results To further explore the catalytic diversity of AA9 LPMOs, over 17,000 sequences were extracted from public databases, filtered, and used to construct a sequence similarity network (SSN) comprising 33 phylogenetically supported clusters. From these, 32 targets were produced successfully in the industrial filamentous fungus *Aspergillus niger*, 25 of which produced detectable LPMO activity. Detailed biochemical characterization of the eight most highly produced targets revealed individual C1, C4, and mixed C1/C4 regiospecificities of cellulose surface oxidation, different redox co-substrate preferences, and CBM targeting effects. Specifically, the presence of a CBM correlated with increased formation of soluble oxidized products and a more localized pattern of surface oxidation, as indicated by carbonyl-specific fluorescent labeling. On the other hand, LPMOs without native CBMs were associated with minimal release of soluble products and comparatively dispersed oxidation pattern.

Conclusions This work provides insight into the structural and functional diversity of LPMOs, and highlights the need for further detailed characterization of individual enzymes to identify those best suited for cellulose saccharification versus surface functionalization toward biomaterials applications.

[†]Yann Mathieu and Olanrewaju Raji contributed equally to this work

*Correspondence:

Emma Master
emma.master@utoronto.ca

Harry Brumer
brumer@mssl.ubc.ca

¹ Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC V6T 1Z4, Canada

² BioProducts Institute, University of British Columbia, 2385 East Mall, Vancouver, BC V6T 1Z4, Canada

³ Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, Toronto, ON M5S 3E5, Canada

⁴ Centre for Structural & Functional Genomics, Concordia University, 7141 Sherbrooke-West Street, Montreal H4B 1R6, Canada

⁵ Department of Bioproducts and Biosystems, Aalto University, Kemistintie 1, 02150 Espoo, Finland

⁶ Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC V6T 1Z1, Canada

⁷ Department of Biochemistry and Molecular Biology, University of British Columbia, 2350 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada

⁸ Department of Botany, University of British Columbia, 3200 University Boulevard, Vancouver, BC V6T 1Z4, Canada



Background

Biomass is a promising alternative to petroleum-based feedstocks for the production of valuable fuels, chemicals, and materials. Terrestrial plant biomass, in particular, is an abundant and renewable source of both complex carbohydrates and hydrocarbons (i.e., polysaccharides, lignin, and low molecular weight extractives) [1–5]. The development of chemical and enzymatic processes to convert lignocellulose into value-added products continues to expand, especially over the past decade [6–9]. However, the fractionation and transformation of lignocellulose remains challenging, due to its inherently heterogeneous and recalcitrant nature [10–12]. Significant inspiration for the development of efficient biocatalytic methods to overcome these challenges comes from the microbial world: Diverse bacteria and fungi have evolved specialized hydrolytic and oxidative enzymatic systems to target individual biomass polymers for nutrient acquisition [13–20].

Lytic polysaccharide monooxygenases (LPMOs) have attracted attention for biomass deconstruction and biofuels production, due to their ability to potentiate the activity of cellulases and thereby enhance lignocellulose saccharification [21–24]. More recently, LPMOs have been harnessed for nanocellulose production [25–31] and the surgical introduction of carboxylate groups as reactive handles on cellulose and chitin surfaces [32, 33]. LPMOs are mononuclear copper enzymes, which have been classified into eight sequence-related Auxiliary Activity families (AA9, AA10, AA11, AA13, AA14, AA15, AA16 and AA17) in the Carbohydrate-Active Enzymes (CAZy) database [34–36]. First identified due to their ability to oxidatively cleave insoluble cellulose or chitin, which remain the predominant reported activities, certain LPMOs preferentially target xyloglucan [37], pectin [36] and starch [38]. LPMOs use molecular oxygen or hydrogen peroxide [39, 40] and external electron donors originating from enzymatic or non-enzymatic reactions to perform C–H bond activation leading to hydroxylation of C1 or C4 of glycosyl residues within glycan chains. Subsequent elimination results in chain cleavage and production of the corresponding C1 lactone (which may hydrolyze to the carboxylic acid) or C4 ketone (which may hydrate to the diol), respectively [41]. Based on this regioselectivity, LPMOs can be divided into three types: C1-oxidizing, C4-oxidizing, and C1/C4-oxidizing [42, 43].

Structurally, LPMOs are small β -sheet proteins that tightly coordinate the catalytic copper in a highly conserved histidine-brace motif [44]. LPMOs can also comprise multi-modular architectures containing non-catalytic carbohydrate-binding modules (CBM, e.g., CBM1 or CBM18, specific for cellulose and chitin,

respectively), or other modules of unknown function [45, 46]. CBMs can enhance catalysis by diverse carbohydrate-active enzymes (CAZymes) through substrate targeting, particularly at low substrate consistency [47–49]. Notably, the deletion or replacement of native CBMs with those from different CBM families was detrimental to LPMO activity on insoluble substrates [28, 50–54].

AA9 comprises the first, and currently largest, family of fungal LPMOs, which has a rich history of study [34, 55]. The large size of AA9 and apparent sequence diversity of its members pose a challenge for the comprehensive, definitive functional characterization of this family. Here, we sought to increase the biochemical coverage of AA9 to enable more confident bioinformatic prediction of function within the family. We mined over 17,000 AA9 homologs from publicly available databases and used Sequence Similarity Network analysis to guide target selection from distinct clusters. Ultimately, 32 AA9-encoding genes were successfully produced heterologously in *Aspergillus niger* and LPMO activity was confirmed by HPAEC-PAD for 25 targets. Eight of these were chosen for purification and detailed biochemical characterization, with emphasis on cellulose functionalization through the introduction of C1 carboxylate and/or C4-keto groups and direct visualization of activity on cellulose surfaces using fluorescent labeling.

Results and discussion

Selection and production of AA9 LPMO targets

After removing identical sequences, and those with an arginine as the first amino acid instead of the catalytically essential histidine [56], the 17,575 AA9 homologs collected from several sequence databases were reduced to 5328 unique sequences. Scalable Sequence Similarity Network (SSN) analysis [57, 58] based on pairwise BLASTP analysis defined 33 subgroups (clusters) of AA9 modules using a bit-score threshold of 250 (Fig. 1). The choice of this threshold was based on inspection of the number of clusters and the number of sequences within them. A bit-score threshold of 225 produced fewer clusters (25 in total); however, some clusters were not well resolved. On the other hand, application of a bit-score threshold of 275 doubled the number of clusters from 33 to 66. To further validate this delineation, a maximum-likelihood phylogenetic tree was constructed from 10 sequences randomly selected from each of the 33 SSN clusters. The monophyletic groups inferred from the phylogenetic tree were identical to the subfamilies identified in the SSN and were color coded accordingly (Fig. 1). Importantly, we could correlate SSN clusters and phylogenetic groups using a bit-score threshold of 250, while we could not with thresholds of 225 or 275. Eleven of

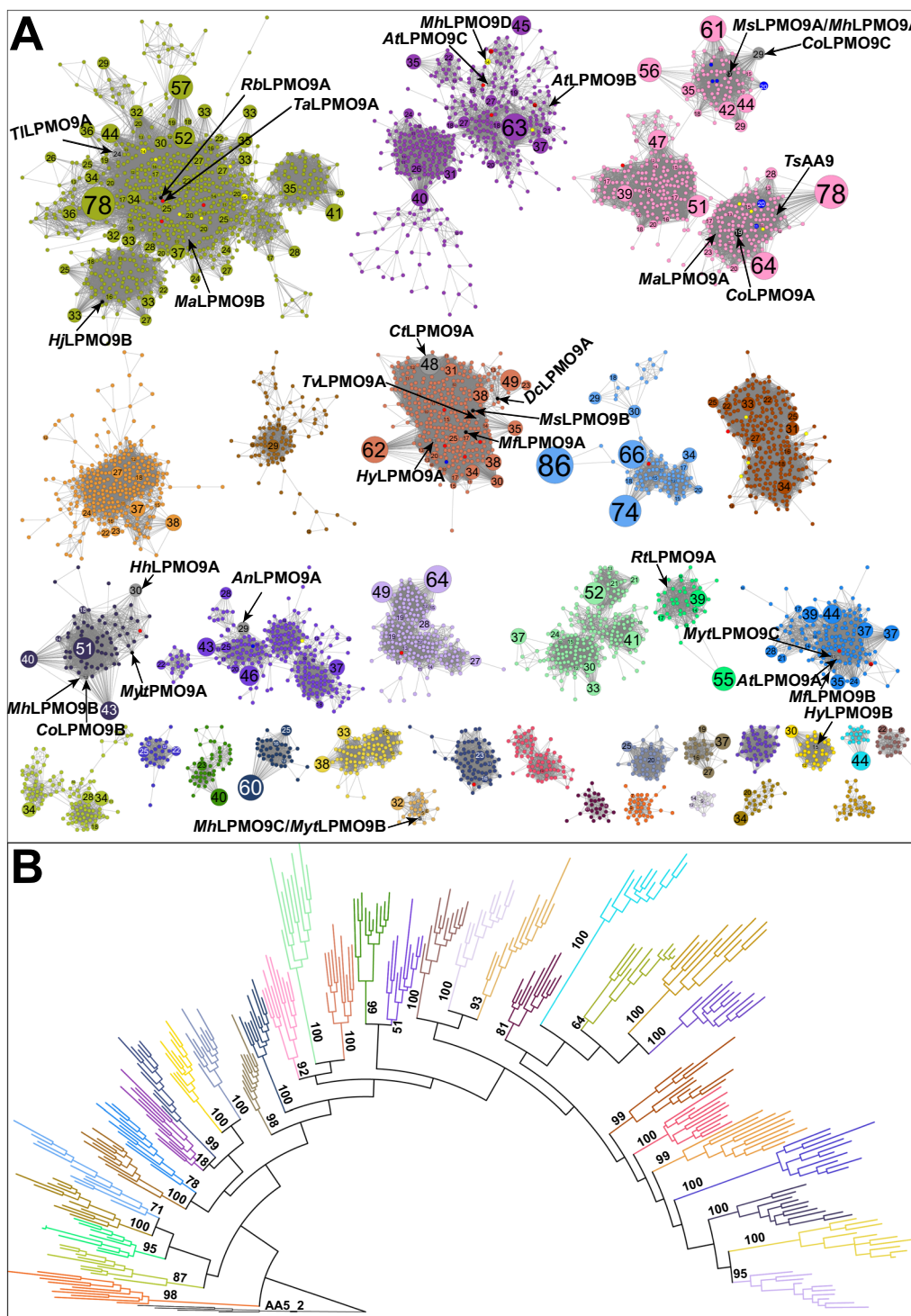


Fig. 1 Sequence relationship of 5,328 AA9 catalytic modules. **A** Sequence similarity network (SSN) created with SSNpipe [58] and displayed in Cytoscape with yFiles Organic layout (Shannon et al. 2003a). Each node corresponds to one of the curated 5328 catalytic modules used as an input to build the SSN. Edges represent an alignment bit-score threshold of 250 that clusters the sequences into subgroups. AA9 members whose regioselectivity is available are colored in blue, yellow and red for C1, C4 and C1/C4 oxidizing enzymes, respectively (functional annotations were obtained from the CAZy database [20]). Sequences expressed in *Aspergillus niger* are colored in black. **B** Maximum-likelihood phylogenetic tree. 10 representative sequences for each subgroup defined by the sequence similarity network in **A**. Bootstrap values based on 100 replicates are shown

the 33 SSN clusters contained at least 1 characterized member.

From a broad selection of 404 targets spanning the major SSN clusters, the production campaign yielded 32 LPMOs from 10 clusters that were produced at quantities detectable by SDS-PAGE (Additional file 2: Table S1). These targets represent the ascomycete classes Sordariomycetes, Eurotiomycetes, and Dothideomycetes. Initial product analysis by HPAEC-PAD, using phosphoric acid-swollen cellulose (PASC) as a substrate and ascorbate as a reducing agent, indicated that 14 of these generated C1-oxidized products, 5 produced C4-oxidized products, 6 produced both C1 and C4-oxidized products, and 7 were apparently inactive (Additional file 2: Table S1), despite possessing intact histidine brace active-site motifs [56] (see the recent review by Vandhana et al. for an in-depth discussion of the potential functions of non-canonical LPMO-like proteins [59]). Of the 25 active LPMOs identified herein, 8 were selected for detailed characterization based on clear and detectable activity, and success in protein production and purification (Additional file 2: Table S1): *CoLPMO9A*, *DcLPMO9A*, *HjLPMO9B*, *MjLPMO9A*, *MsLPMO9A*, *MsLPMO9B*, *MytLPMO9A*, and *RbLPMO9A*. To our knowledge, except for *HjLPMO9B*, all selected sequences represented the first LPMO to be characterized from the source organism. Like *HjLPMO9A* (formerly *TrCel61A* [53]), *HjLPMO9B* displayed C1/C4-oxidizing activity; however, it does not possess a CBM. *MytLPMO9A* shares 91% sequence similarity with the previously characterized *MtLPMO9D* [60] and both display C1-oxidizing activity. *MytLPMO9A* is from *Myriococcum thermophilum* and is referred to here with a three letter acronym denoting the source organism to distinguish it from *MtLPMOs* that originate from *Myceliophthora thermophila*.

Impact of cellulose type on LPMO activity

Following the activity screens using PASC (Additional file 2: Table S1), the eight purified LPMOs were compared using both PASC and Avicel as cellulose substrates having different crystallinity. Sulphanilic acid-treated Avicel (SA-Avicel) was also included in the comparative analysis as this derivatized form of Avicel would be later used for fluorescent labeling studies. *MjLPMO9A*, *MsLPMO9B*, *MytLPMO9A* and *DcLPMO9A* consistently released C1-oxidized cello-oligosaccharides from PASC, Avicel, and SA-Avicel (Fig. 2). By contrast, peaks corresponding to C4-oxidized products were consistently detected when treating the cellulosic substrates with *CoLPMO9A* and *MsLPMO9A*, whereas *RbLPMO9A* and *HjLPMO9B* consistently produced both C1 and C4-oxidized products (Fig. 2). Commensurate with previous studies [61], non-oxidized cello-oligosaccharides were also released

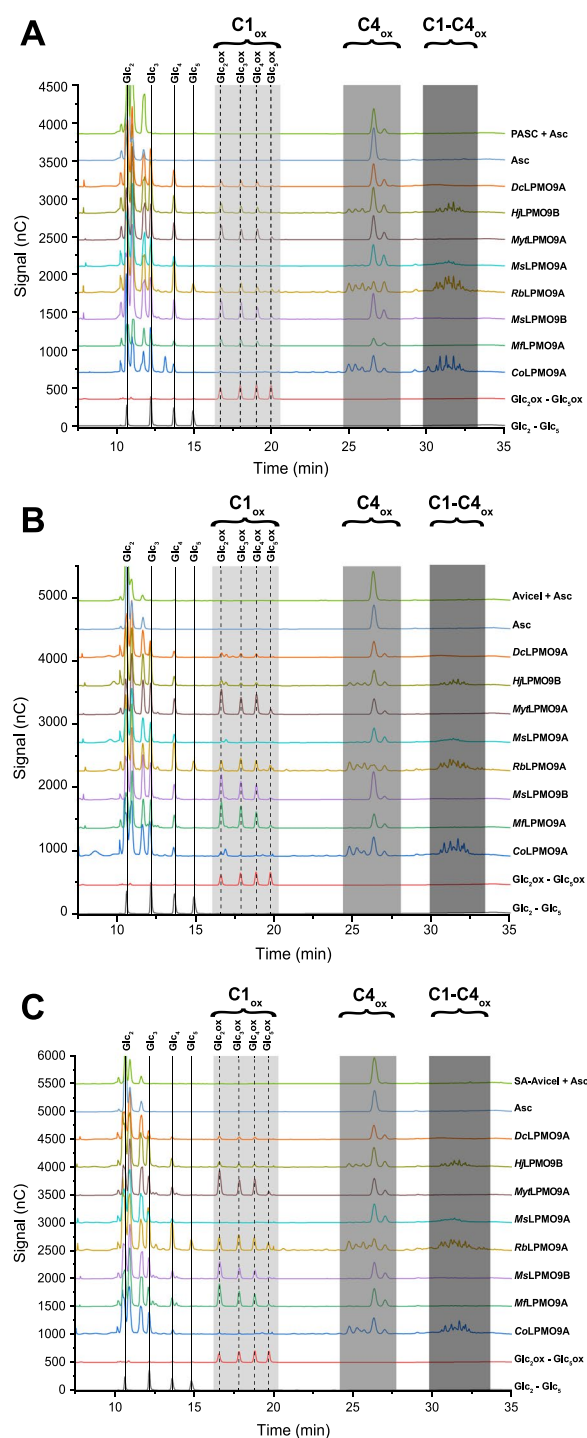


Fig. 2 HPAEC-PAD analysis of LPMO regioselectivity towards three cellulosic substrates. Chromatograms over the retention times of 7.5–35 min. Native and C1-oxidized cello-oligosaccharide standard peaks are annotated. All activity assays were conducted using 5 μ M of enzyme, 0.1% PASC (A), 1% Avicel (B) or 1% SA-Avicel (C) and 1 mM ascorbic acid for 16 h at 50 °C. The negative controls (substrate + electron donor) contained all assay components except enzyme

by all LPMOs. To simplify analyses, total peak areas of all C1-oxidized products, all C4-oxidized products, and all C1/C4-oxidized products were used as semi-quantitative estimates of cellulose deconstruction following overnight reactions (Fig. 3). Notably, both *DcLPMO9A* and *MsLPMO9A* lack a CBM and generated lowest amounts of soluble products from all substrates. Crucially, pre-treatment of Avicel with sulphanic acid to block intrinsic carboxylate groups did not alter LPMO regioselectivities nor significantly alter LPMO performance.

Among the eight LPMOs biochemically characterized herein, the CBM-containing LPMOs generally released more soluble products regardless of substrate type (Fig. 3). The exception was *MytLPMOA*, which lacks a CBM and yet performed similar to CBM-containing LPMOs tested in this study. Of note, low (0.1–1%) cellulose consistency was used in this study and similar impacts of cellulose binding modules on LPMOs with different regioselectivity were previously reported [62, 63].

Impact of reducing agent on LPMO activity

Ascorbic acid, gallic acid, and cysteine have previously been shown to exhibit differential reactivity as small-molecule electron donors for LPMOs [64, 65]. Hence, the

impact of electron donor on the LPMOs of the present study was evaluated. Due to a lack of pure standards to enable quantitation of C4 products, only C1-oxidizing LPMOs were used for time-course investigations. For the four C1-oxidizing LPMOs, product release from all cellulose substrates approached maximum levels after 2 h when using either ascorbic acid or L-cysteine as the reducing agent (Fig. 4). By comparison, soluble product formation was slower in reactions amended with gallic acid, in which maximum product release was measured after 16 h (Fig. 4). Detailed consideration of the time-course curves for oxidized soluble product formation by each C1 LPMO revealed nuanced impacts of electron donor and cellulose substrate combinations. For example, *MfLPMO9A* released highest levels of soluble product from PASC when using gallic acid as the reductant, whereas using gallic acid reduced total soluble product released from Avicel by 30%. Instead, both *MfLPMO9A* and *MsLPMO9B* released highest levels of soluble product from Avicel when using ascorbic acid as the reductant, whereas using ascorbic acid reduced total soluble products from PASC by over 55%. Notably, the choice of electron donor did not substantially impact soluble product release by *DcLPMO9A* from any of the cellulose substrates, and using cysteine as reductant did not

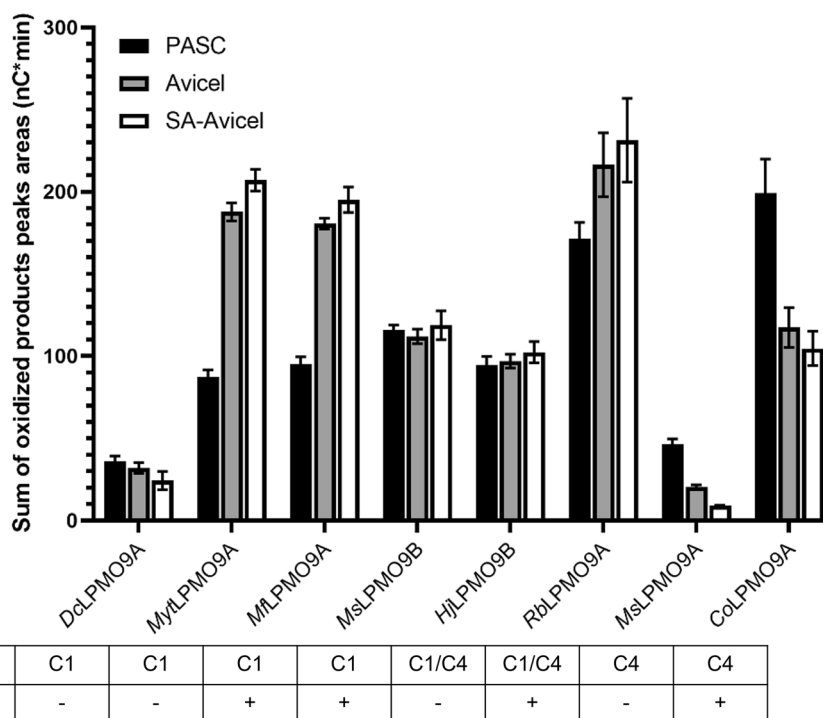


Fig. 3 Activity of 5 μM LPMOs after 16 h on PASC (0.1%), Avicel (1%), and SA-Avicel (1%) with 1 mM ascorbic acid as an electron donor. Semi-quantitative analysis of C1 and C4 oxidized products was completed by summing the peak areas of C1-oxidized and C4-oxidized oligosaccharides. Each bar is the average of three independent assays measured singly by HPAEC-PAD, with error bars indicating the standard error of the mean

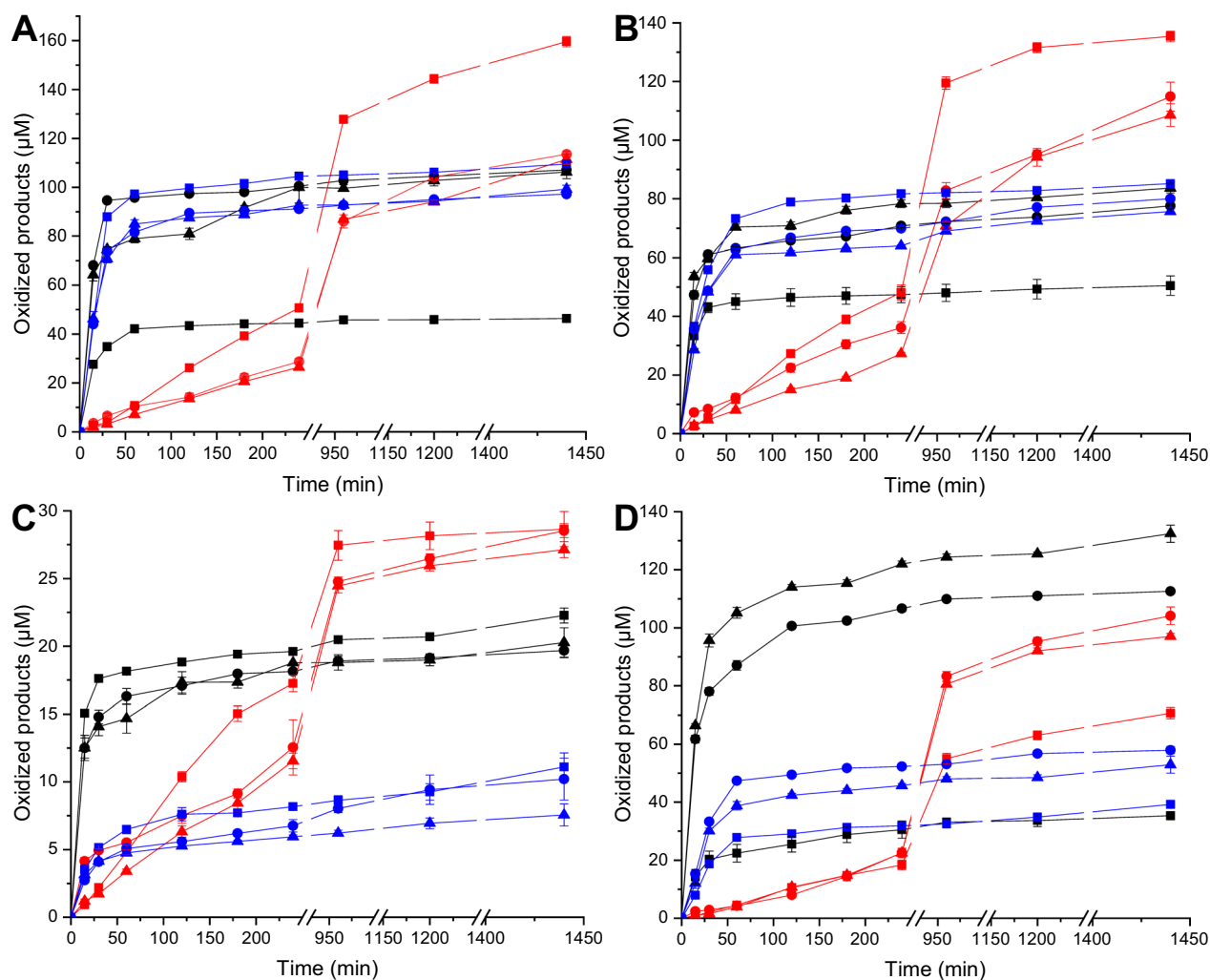


Fig. 4 Time-course release of soluble oxidized products for **A** *MfLPMO9A*, **B** *MsLPMO9B*, **C** *DcLPMO9A*, and **D** *MytLPMO9A*. For each panel, 5 µM LPMOs were incubated at 50 °C for 24 h with 1 mM of either ascorbic acid (black lines), gallic acid (red lines) or cysteine (blue lines) with PASC (squares), Avicel (circles) or SA-Avicel (triangles). For each time point, *T. reesei* cellulase cocktail was used to convert all C1-oxidized products into cellobionic acid and quantified as the total C1-oxidized ends generated (µM). Total oxidized ends were obtained by quantifying cellobionic acid by HPAEC-PAD against a standard curve. Every point is the average of three independent assays measured individually by HPAEC-PAD, with error bars indicating the standard error of the mean

substantially change the LPMO activities on the different cellulose types.

Previous studies investigating the impact of reductant on soluble product release from amorphous cellulose compared three family AA9 LPMOs from *Myceliophthora thermophila* that display different regioselectivities, and in all cases, the preferred reductant was ascorbic acid, followed by gallic acid and then L-cysteine [64]. Conversely, among closely related chitin-active AA10s from *Cellulomonas* species, cysteine, a sulfur containing electron donor, generated more oxidized soluble products [65]. In our case, the time-course experiments clearly demonstrated that each

enzyme favored a specific electron donor/substrate couple for the release of soluble C1-oxidized products. In addition, the protein modularity and structure also impact substrate-dependent preferences for particular reducing agents. Altogether, a growing body of literature, including our present analysis, demonstrates that each LPMO needs to be carefully assessed in light of specific applications (total versus controlled oxidation), with regard to activity and reducing agent specificity for a particular substrate. For instance, using two different electron donors on the same substrate can result in different rates of oxidation and hence surface modification vs. erosion.

LPMO action on cellulose fiber surfaces

SA-Avicel was prepared to block pre-existing carboxyl groups in the cellulose substrate [66] and then used to investigate LPMO-mediated oxidation of cellulose fiber surfaces. Whereas the quantitation of carbonyls introduced by C4 LPMOs remains challenging, carbonyls introduced by C1 LPMO oxidation were quantified through enzymatic hydrolysis of the residual cellulose to glucose, cellobiose and cellobionic acid. Regardless of whether the C1 LPMO comprised a CBM, between 9 and 12 nmol of carboxylate groups per mg of starting material were introduced to the fiber surface (Fig. 5). For material applications requiring fiber surface oxidation, preferred LPMOs would be those that achieve high surface oxidation while minimizing loss in cellulose yield (mass) through release of soluble products [32, 33]. Such a product profile was achieved using *DcLPMO9A* where carboxylate groups on the fiber surface were comparable to the other LPMOs while retaining 2–4 times higher fiber yield (Additional file 1: Figure S1).

To provide additional insight into the distribution of LPMO activity on fiber surfaces, the rhodamine green-fluorescent dye and cyanine 647 aminoxy dye were used to label the enzymatically introduced C1 carboxylate and C4-keto groups, respectively. Although attempts to quantify the resulting fluorescence were unreliable due

to the unavoidable heterogeneity of the substrate suspensions, the pattern of LPMO activity on SA-Avicel surfaces could be visualized by confocal microscopy. Oxidations introduced by all LPMOs lacking a CBM (i.e., *DcLPMO9A*, *MytLPMO9A*, *MsLPMO9A*, and *HjLPMO9B*) were evenly dispersed across the fiber surface (Fig. 6); by comparison, the presence of a CBM in *MflLPMO9A*, *CoLPMO9A*, and *RblLPMO9B* led to a more punctate distribution of oxidized sites across the cellulose fiber surface (Fig. 6). *MsLPMO9B* also comprises a CBM1 domain, however, the oxidation pattern on fiber surfaces was similar to that introduced by the LPMOs lacking a CBM (Fig. 6).

The overall trend whereby CBM-containing LPMOs introduce punctate oxidation sites on cellulose fiber surfaces agrees with earlier predictions drawn from biochemical and dynamic modeling studies of *ScLPMO10C* and a truncated variant lacking the CBM [54]. These authors suggested that the presence of the CBM promotes multiple and localized *ScLPMO10C* action on the cellulose substrate. Likewise, our data provide visual confirmation of the significant impacts that CBMs have on LPMO action.

To investigate the relative distribution of C1 and C4 oxidized sites introduced by C1/C4 type LPMOs, rhodamine green and cyanine 647 aminoxy dyes were used

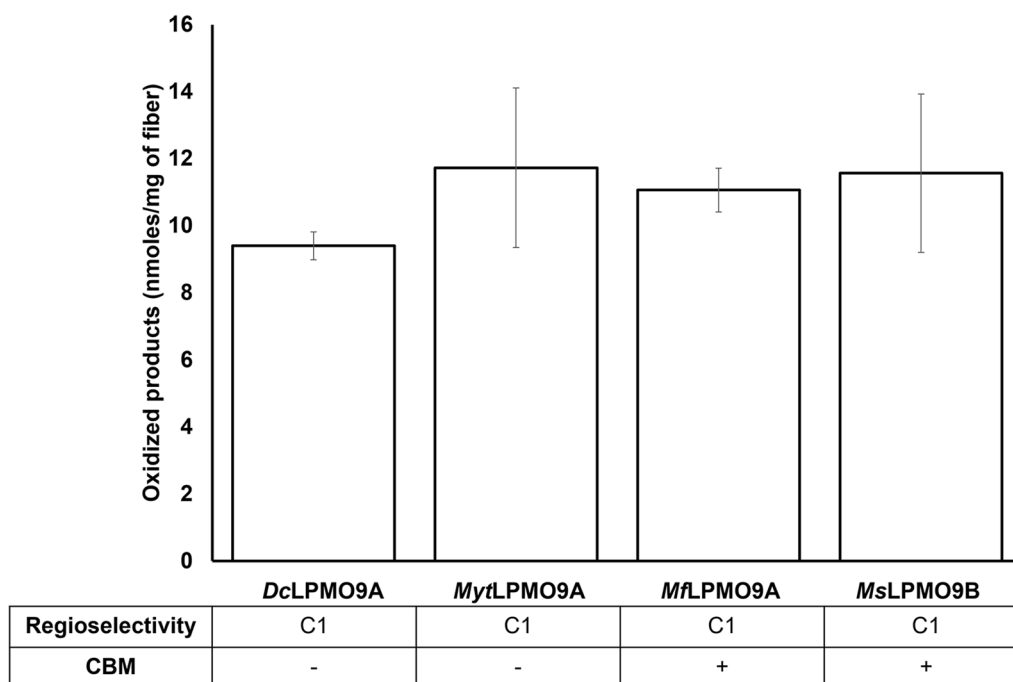


Fig. 5 Assessing LPMO (5 μM) activity on SA-Avicel (1%) fiber after 16 h with 1 mM ascorbic acid as an electron donor. *T. reesei* cellulase cocktail was used to convert all oxidized insoluble products into cellobionic acid and quantify the total C1-oxidized ends introduced to the fiber (nanomoles per mg of starting fiber). Total oxidized ends were obtained by quantifying cellobionic acid by HPAEC-PAD against a standard curve. Each bar is the average of three independent assays measured singly by HPAEC-PAD with error bars indicating the standard error of the mean

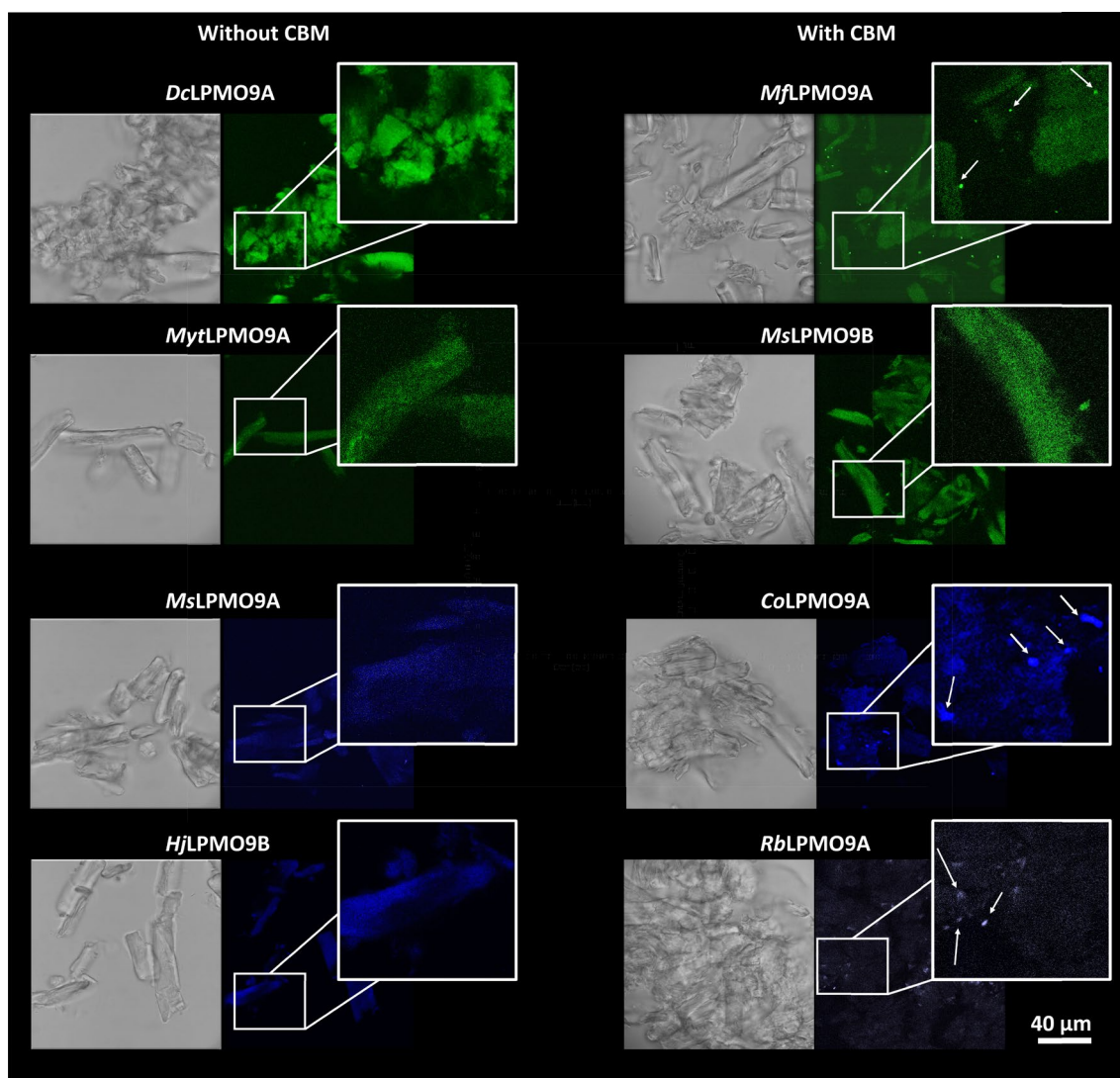


Fig. 6 Brightfield and confocal images of LPMO-treated SA-Avicel labeled using cyanine aminoxy dye (blue) for C4 products and rhodamine chloride (green) for C1 products. For each panel, 5 μ M LPMOs were incubated at 50 $^{\circ}$ C for 24 h with 1% SA-Avicel and 1 mM of gallic acid. Insoluble products were separated, labeled with fluorescent dye, and visualized using a confocal microscope. Untreated substrates did not show fluorescence signals (Additional file 1: Figure S2)

together to label cellulose fibers after treatment with *HjlLPMO9B* and *RblLPMO9A*. The order of the labeling did not influence the analysis; C1 labeling and C4 labeling can be performed simultaneously without cross-reactivity or inhibition of either labeling process. After labeling, cellulose treated with *HjlLPMO9B* or *RblLPMO9A* revealed a blended distribution of the green and blue fluorescence, suggesting co-location of C1- and C4-oxidation events on the fiber surface regardless of the presence of a CBM. While this points to minimal impact of substrate surface characteristics on C1- versus C4-oxidation events, methods to quantify C4 oxidation on cellulose fiber surfaces are needed to confirm this interpretation.

Conclusions

The present study investigated the impact of LPMO sequence, cellulose morphology, and small-molecule reductant on the soluble product profile, extent of substrate degradation, and pattern of cellulose surface oxidation. Although regiospecificity could not be correlated to SSN clusters, the extent of soluble product release and punctate pattern of cellulose surface oxidation was notably affected by the presence of a carbohydrate-binding module (CBM) in tandem. Of further practical relevance, the optimal reductant was LPMO-specific and influenced by cellulose type. The development of LPMOs as tools not only for cellulose deconstruction, but also

to produce functionalized cellulosic materials relies on having a holistic view of LPMO action that tracks both soluble and insoluble product formation. The fluorescent labeling method applied herein not only allows us to qualitatively confirm surface oxidation of fibers, but also to visually observe the pattern of oxidation relevant for material applications. The analyses revealed that while *Myt*LPMO9A and CBM-containing LPMOs characterized herein might be the preferred enzymes for cellulose deconstruction, *Dc*LPMO9A and *Ms*LPMO9A would be preferred for dispersed cellulose surface oxidation leading to carboxyl and keto-functional groups, respectively. Applications benefiting from these surface modifications include fiber dye retention and fiber carding; the known ability to produce the relevant enzymes in an industrial production system (i.e., *Aspergillus niger*) motivates near term development of these application concepts.

Materials and methods

Materials

Avicel[®] was purchased from Sigma Aldrich (St. Louis, MO, USA), glucose was purchased from Fisher Scientific (Hampton, NH, USA), and cellobiose was purchased from Acros Organics (Morris Plains, NJ, USA). Cello-oligosaccharides (Glc₂–Glc₅) were purchased from Megazyme (Bray, Ireland). C1-oxidized cello-oligosaccharides standards were prepared using *Podospira anserina* cellobiose dehydrogenase (GenBank accession number: CAP64840) [67], kindly provided by Dr. Jean-Guy Berin (INRA, Marseille, France). Briefly, 1 μM of CDH was added to 2 mM of each cello-oligosaccharide (Glc₂–Glc₅) in 50 mM sodium acetate buffer pH 5.0, followed by incubation at 37 °C. Full conversion was achieved by supplementation of additional CDH every 24 h for 3 days. Phosphoric acid-swollen cellulose (PASC) was prepared from Avicel[®] following a previously published protocol [68].

Treatment of Avicel with sulphanilic acid (SA-Avicel) was performed to block pre-existing carboxylic acid groups on Avicel. Briefly, Avicel was suspended in dimethylformamide (DMF) (Bioshop Ontario, Canada) at 50 g/L before adding the coupling reagent 12.5 mM benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) (Sigma Aldrich Co. Missouri, USA) and 200 mM N, N-di-isopropylethylamine (DIPEA) (Sigma Aldrich Co. Missouri, USA). After gentle mixing for 30 min, 100 mM of sulphanilic acid was added and the mixing continued for an additional 60 min. The reaction was centrifuged and washed several times with DMF to remove unbound sulphanilic acid. The SA-Avicel product was suspended in MilliQ water to a final concentration of 20 g/L.

Bioinformatics for target selection

A total of 17,575 AA9 sequences were retrieved from Carbohydrate-Active enZYmes (CAZy; www.cazy.org), Joint Genome Institute (JGI; www.jgi.doe.gov), and Concordia Characterized Lignocellulose-Active Enzymes (CLAE; www.clae.fungalgenomics.ca) and Centre for Structural and Functional Genomics (CSFG; www.concordia.ca/research/genomics.html) databases. Where present, signal peptides and additional modules, such as carbohydrate-binding modules, were removed to isolate the catalytic modules for subsequent analyses. Catalytic modules sharing 100% identity were down-sampled to one sequence to eliminate redundancy. Moreover, obviously erroneous sequences generated by incorrect splicing predictions were removed via the identification of unusually long deletions/insertions through visual inspection of a multiple sequence alignment. Initial clustering of highly similar sequences was performed using CD-HIT [69] with an 80% identity cut-off [70], yielding a reduced dataset of 5,328 sequences from the original 17,575, and merged into meta-nodes. Sequence similarity networks (SSNs) were generated by computing all-versus-all pairwise local alignments of the 5,328 curated AA9 catalytic domains using SSNpipe [58], which generated the *E*-value, bit score, alignment length, sequence identity and sequence similarity for all sequence pairs. The data were filtered using a bit-score threshold of 250 to generate the final SSNs. To constitute a subfamily or cluster, connected clusters were required to contain at least 20 sequences from different organisms to avoid over-representation of closely related organisms. Members of each putative subfamily were identified using NetworkX [71] and SSNs were visualized with Cytoscape [72] using the yFiles organic layout.

For each cluster, 10 random sequences were aligned with MAFFT v7.402 using the L-INS-I algorithm [73] on the CIPRES Science Gateway (www.phylo.org) [74]. Four AA5_2 sequences (Genbank accessions AHA90705, EFQ30446, CAP96757 and XP_003719369) were included as an outgroup. The quality of the alignment was validated by visual inspection in AliView [75]. A maximum likelihood phylogenetic tree was estimated using RAxML v.8 [76] with 100 bootstrap replications on the CIPRES Science Gateway portal and formatted using iTOL [77].

Gene cloning, transformation into *Aspergillus niger*

The AA9 sequences selected for expression in *Aspergillus niger* were PCR-amplified from cDNA or genomic DNA of source organisms (Additional file 2: Table S1). Where source organisms were not readily available, genes were synthesized (Integrated DNA Technologies Inc.;

Iowa, USA). When obtained from the source organisms, the fungal strains were cultivated in 25 mL of *Trametes* defined medium (TDM) [78] supplemented with 2% of a mix of alfalfa and barley straws. *Hypocrea jecorina* cultures were incubated at 30 °C and all other fungi were incubated at 45 °C. In all cases, cultivations continued for 24 h with shaking at 220 rpm. Mycelia were harvested, ground into powder and genomic DNA was extracted as described previously [79]. In parallel, total RNA was extracted using the RNeasy® Plant Maxi Kit (Qiagen, USA) and complementary DNA (cDNA) was synthesized using the Superscript™ III reverse transcriptase (Invitrogen, ThermoFisher Scientific, USA) as per manufacturer instructions. The amplified or synthesized genes were cloned between the *A. niger* glucoamylase gene promoter and terminator of the plasmid ANIp4 [80] or its derivatives.

The recombinant LPMO genes were introduced into the *A. niger* genome by replacement of the glucoamylase gene (UniProt: A2QHE1) using CRISPR/Cas9 genome editing. The CRISPR plasmid ANEp8-Cas9-gRNA_glaA used in this study contains the guide RNA sequence 5'-GCAGTGTGACTGTCACCTCG-3' that targets the coding region of the glucoamylase gene. The guide RNA was cloned into the plasmid ANEp8-Cas9-gRNA as described [81] to generate ANEp8-Cas9-gRNA_glaA. The CRISPR plasmid ANEp8-Cas9-gRNA_glaA along with a recombinant LPMO gene were introduced into *A. niger* by protoplast transformation [82, 83]. Strain CSFG_9047 (A1513 $\Delta kusA \Delta [prtT amyC agdA] \Delta bglA \Delta laeA$), engineered for this study, was derived from *A. niger* A1513 (Fungal Genetics Stock Center, USA), an industrial cell factory which is also designated as CBS 513.88 [84]. The gene models of the *A. niger* CBS 513.88 genome sequence [84] were used to guide the construction of the production host CSFG_9047; and the gene IDs of CBS 513.88 were used as references for the deleted genes (see below). Strain CSFG_9047 is a uridine/uracil prototroph ($\Delta pyrG$; ID: An12g03570); defective in non-homologous end-joining DNA repair ($\Delta kusA$; ID: An15g02700) to promote homologous recombination [85]; deficient in extracellular proteases ($\Delta prtT$; ID: An04g06940) to reduce proteolytic degradation of secreted proteins as the gene *prtT* encodes a regulator of extracellular proteases [86]; and defective in citric acid production ($\Delta laeA$; ID: An01g12690) to prevent acidification of the culture medium [87]. During the course of this study, proteomic analysis detected a number of extracellular proteins in the culture filtrates. We removed the genes encoding three of them to reduce the level of background proteins: alpha-glucosidase ($\Delta agdA$; ID: An04g06920); alpha amylase $\Delta amyC$; ID: An04g06930); and beta-glucosidase ($\Delta bglA$; ID: An18g03570). The deletion of the *bglA* gene

is important for this study because the beta-glucosidase activity in the culture filtrate interferes with the LPMO activity assays.

Protein production and purification

Selected *A. niger* transformants were inoculated in liquid modified minimal medium (MMJ) [88] supplemented with 0.1% arginine and screened for protein production by SDS-PAGE. Transformants producing the target proteins were cultured in 200 mL to 1 L liquid MMJ. Culture supernatants were concentrated and buffer-exchanged using Vivaflow® cassettes (Sartorius, Göttingen, Germany) based on the manufacturer's protocol and according to [89]. Once exchanged to 50 mM sodium acetate (pH 5.0), recombinantly produced LPMOs were analyzed on Criterion™ TGX Stain-Free protein gels (Bio-Rad, Canada) and quantified by densitometry using GeneTools v.4.3.9.0 (Syngene, UK) according to the instructions provided by the manufacturer. Bovine serum albumin (BSA) was used as the standard for quantification and the identity of each LPMO was confirmed by mass spectrometry as described below.

The heterologously produced proteins were then tested for LPMO activity as described below. Supernatants showing activity were concentrated to 2 mL and filtered through a 0.22- μm PES syringe filter before purification. Protein purification was achieved using size exclusion chromatography on a 16-mm \times 600 mm Superdex® 75 column pre-equilibrated with 50 mM sodium acetate pH 5.0 at 1 mL \cdot min⁻¹; eluted samples were collected in 500- μL fractions. The fractions were analyzed by SDS-PAGE and those containing the target protein with >95% purity were pooled and concentrated before being flash frozen and stored at -80 °C. Protein concentrations were determined by measuring A_{280} using extinction coefficients which were calculated using the ProtParam tool on the ExPASy server (<http://web.expasy.org/protparam/>).

LC-MS/MS-based LPMO identification in culture supernatants

Two volumes of cold methanol were added to 100 μL of cleared culture supernatant and kept on ice for 30 min. Samples were centrifuged for 30 min at 15,000 \times g and 4 °C. The precipitated protein pellet was washed once with 300 μL of cold 60% methanol in water before being suspended in 30 μL of 6 M urea solution containing 100 mM ammonium bicarbonate (pH 8.0) and digested with trypsin as previously described [90]. Aliquots of peptide digest were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) using an Agilent 1260 Infinity II chromatography system connected in-line with a Thermo-Finnigan 7 Tesla LTQ-FT MS system. MS/MS data were processed for protein

identification using the precursor ion quantitation workflow from Proteome Discoverer 2.4. MS/MS fragmentation data were queried against a database of 17,897 protein sequences comprising the *A. niger* NRRL3 protein models plus a collection of recombinant protein sequences, including the LPMOs of interest.

Enzyme activity assays

Unless stated otherwise, all assays were performed in 1.7 mL tubes incubated in an Eppendorf Thermomixer C (Hamburg, Germany) at 50 °C with shaking at 1,000 rpm. For initial screens of culture supernatant, reactions (300 μ L) were performed in 50 mM sodium acetate buffer (pH 5.0) containing 5 μ M of enzyme, 0.1% w/v PASC, and 1 mM ascorbic acid; the ascorbic acid was added last to initiate the reaction. Purified enzymes were tested the same way using three substrates (0.1% w/v PASC, 1.0% w/v Avicel, and 1.0% w/v SA-Avicel) and three electron donors (ascorbic acid, gallic acid or L-cysteine). Reaction aliquots were collected at 15 min, 30 min, 60 min, 120 min, 240 min, 960 min, 1,200 min and 1,440 min. In all cases, reactions were filtered through a 0.2- μ m cellulose acetate spin filter (VWR, P/N: 2994–752) for soluble products analysis.

Reaction products were analyzed by High Performance Anion Exchange Chromatography (HPAEC) using an ICS-5000 system (Dionex, Sunnyvale, CA, USA) coupled to a gold electrochemical detector (Dionex, P/N: 072044) for pulsed amperometric detection (PAD) [91, 92]. For initial screens of LPMO activity in culture supernatants, 25 μ L of the reaction mixture were directly injected on a CarboPac PA1 (2 \times 250 mm) analytical column (Dionex, P/N: 057178) and corresponding guard column (2 \times 50 mm) (Dionex, P/N: 057179) maintained at 30 °C. Product separation was achieved at a constant flowrate of 0.25 mL \cdot min⁻¹ with the following method: initial column equilibration for 10 min with an isocratic flow of 0.1 M NaOH, followed by a linear gradient to 0.25 M NaOAc over 30 min and then a stepwise increase to 0.9 M NaOAc in 1 min. To quantify reaction products from purified LPMOs, reaction mixtures were filtered through a 0.2 μ m cellulose acetate spin filters before adding 0.1 unit of a *T. reesei* cellulase cocktail (Celluclast, Sigma Aldrich, P/N: C2730-30) to 100 μ L of the reaction filtrate and incubation overnight at 37 °C. In parallel, 100 μ L of 50 mM sodium acetate buffer (pH 5.0) was added to the cellulose retentate before adding 0.1 unit of Celluclast and incubating the sample at 37 °C for 48 h. Following enzymatic digestion using Celluclast, 25 μ L of sample were analyzed using HPAEC-PAD as described above. The resulting glucose, cellobiose and cellobionic acid products were quantified based on corresponding

standards. In all cases, reactions were performed as triplicate independent experiments.

Fluorescent labeling and confocal microscopy

To label the C1-oxidized positions in LPMO-treated SA-Avicel, the insoluble cellulose product was recovered by centrifugation at 14,000 \times g for 10 min and then coupled to Rhodamine 110 chloride (Rh110, Sigma Aldrich, St. Louis, MO, USA, P/N: 83,695; excitation max=498 nm; emission max=520 nm) using benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and *N,N*-diisopropylethylamine (DIPEA). Briefly, 4 mM of DIPEA and 1 mM of PyBOP in dimethylformamide (DMF) were gently mixed with the insoluble cellulose fraction (\leq 10 mg) for 5 min before adding 0.25 mM of Rh110, followed by gentle mixing for 5 min. The labeling process involves conventional amide bond formation, where the enzymatically introduced carboxylic acid group reacts with the amine group of Rh110. The reaction was protected from light to prevent bleaching of the fluorophore.

To label C4-oxidized positions in LPMO-treated celluloses, the residual insoluble cellulose fraction (\leq 10 mg) was treated with 0.25 mM cyanine 647 aminoxy dye (Biotium, Fremont, CA, USA; excitation max=650 nm; emission max=665 nm) in DMF for 5 min. The reaction was protected from light to prevent light bleaching of the fluorophore.

The reactions were centrifuged and washed multiple times with DMF to remove unbound dye from the labeled-insoluble product. Washes were monitored using an Infinite M200 spectrophotometer (Tecan, CHE) set to excitation and emission wavelengths of the corresponding fluorophores. Washes continued until unbound dye was not detectable in the washes. Labeled, insoluble products were suspended in 300 μ L MilliQ water and stored away from light at room temperature.

A confocal microscope (Leica TCS SP5 Wetzlar, Germany) equipped with an argon laser at 20% power was used to image the labeled-insoluble products. Images were obtained using a 100 \times oil-immersion lens with a FITC filter for Rh110 excitation at 480–495 nm and emission at 510–535 nm, while a Cy5 Filter for Cyanine 647 Aminoxy excitation at 635–655 nm and emission at 660–675 nm. At least four randomly chosen positions were analyzed per sample. Application Suite (LAS) was used for processing and annotation of captured images.

Abbreviations

AA9	Auxiliary Activity Family 9
CAZyme	Carbohydrate-Active Enzyme
CBM	Carbohydrate-binding module
LPMO	Lytic polysaccharide monoxygenase
SSN	Sequence Similarity Network

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13068-023-02383-3>.

Additional file 1: Figure S1: Product released by 5 μM C1 LPMOs after 16 h on PASC (0.1%), Avicel (1%), and SA-Avicel (1%) with 1 mM ascorbic acid as the electron donor. For each substrate, *T. reesei* cellulase cocktail was used to convert all C1-oxidized products into cellobionic acid, which was then quantified by HPAEC-PAD and reported as the total C1-oxidized ends generated (nanomoles per mg of starting fiber). Each bar is the average of three independent assays measured singly by HPAEC-PAD, with error bars indicating the standard error of the mean. **Figure S2:** Brightfield (A) and confocal (B) images of untreated SA-Avicel labelled using rhodamine chloride. 1% SA-Avicel with 1 mM of gallic acid was incubated at 50 °C for 24 h. Insoluble products were separated, labelled with fluorescent dye, and visualized using confocal microscopy. **Figure S3:** Process schematic for LPMO treatment of cellulosic substrates and subsequent soluble and insoluble product analysis. PASC (0.1%), Avicel (1%), and SA-Avicel (1%) were treated with LPMOs using either ascorbic acid, gallic acid, or cysteine as the electron donor. Insoluble products were separated from soluble products using centrifugation. For the soluble products analysis, HPAEC-PAD was used to annotate native and oxidized cello-oligosaccharide peaks. To quantify the C1-oxidized products, *T. reesei* cellulase cocktail was used to convert all C1-oxidized products into cellobionic acid, which was then quantified by HPAEC-PAD and reported as the total C1-oxidized ends generated (nanomoles per mg of starting fiber). For insoluble product analysis, separated insoluble products were labelled with either C1-specific or C4-specific fluorescent dye and subsequently visualized using confocal microscopy.

Additional file 2: Additional tables.

Acknowledgements

This work was funded by grants awarded to Emma Master, Harry Brumer, and Adrian Tsang from Genome Canada, Ontario Genomics, Genome BC, and G enome Qu ebec for Project Number 10405, "SYNBIOMICS-Functional genomics and techno-economic models for advanced biopolymer synthesis".

Authors contributions

YM: methodology, conceptualization, experimentation, formal analysis, investigation, writing—original draft, writing—review and editing. OR: methodology, conceptualization, experimentation, formal analysis, writing—original draft, writing—review and editing. AB: experimentation, supervision, writing—original draft. MDF: experimentation, writing—original draft. TTMN: data curation, formal analysis, methodology, experimentation, writing—original draft. AHV: data curation, formal analysis, methodology, and writing—review and editing. AT: conceptualization, funding acquisition, project administration, supervision, and writing—review and editing. EM: conceptualization, funding acquisition, project administration, supervision, and writing—review and editing. HB: conceptualization, funding acquisition, project administration, supervision, and writing—review and editing.

Availability of data and materials

All data and sequences supporting the conclusions of this article are included within the article and its additional files, deposited in public databases, or available from the authors upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 1 June 2023 Accepted: 14 August 2023

Published online: 07 September 2023

References

- Isikgor FH, Becer CR. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym Chem-Uk*. 2015;6(25):4497–559.
- Ubando AT, Felix CB, Chen WH. Biorefineries in circular bioeconomy: A comprehensive review. *Bioresour Technol*. 2020;299:122585.
- Dessbesell L, Xu CB, Pulkki R, Leitch M, Mahmood N. Forest biomass supply chain optimization for a biorefinery aiming to produce high-value bio-based materials and chemicals from lignin and forestry residues: a review of literature. *Can J Forest Res*. 2017;47(3):277–88.
- Fatih Demirbas M. Biorefineries for biofuel upgrading: a critical review. *Appl Energy*. 2009;86:5151–61.
- Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ Jr, Hallett JP, Leak DJ, Liotta CL, et al. The path forward for biofuels and biomaterials. *Science*. 2006;311(5760):484–9.
- Wu S, Snajdrova R, Moore JC, Baldenius K, Bornscheuer UT. Biocatalysis: enzymatic synthesis for industrial applications. *Angew Chem Int Ed Engl*. 2021;60(1):88–119.
- Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. Engineering the third wave of biocatalysis. *Nature*. 2012;485(7397):185–94.
- Dedes G, Karnaouri A, Topakas E. Novel Routes in Transformation of lignocellulosic biomass to furan platform chemicals: from pretreatment to enzyme catalysis. *Catalysts*. 2020;10(7):743.
- Bender TA, Dabrowski JA, Gagne MR. Homogeneous catalysis for the production of low-volume, high-value chemicals from biomass. *Nat Rev Chem*. 2018;2(5):35–46.
- Sanchez OJ, Cardona CA. Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresour Technol*. 2008;99(13):5270–95.
- Hall M, Bansal P, Lee JH, Realf MJ, Bommarius AS. Cellulose crystallinity—a key predictor of the enzymatic hydrolysis rate. *FEBS J*. 2010;277(6):1571–82.
- Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*. 2007;315(5813):804–7.
- Cragg SM, Beckham GT, Bruce NC, Bugg TDH, Distel DL, Dupree P, Etxabe AG, Goodell BS, Jellison J, McGeehan JE, et al. Lignocellulose degradation mechanisms across the Tree of Life. *Curr Opin Chem Biol*. 2015;29:108–19.
- Bomble YJ, Lin CY, Amore A, Wei H, Holwerda EK, Ciesielski PN, Donohoe BS, Decker SR, Lynd LR, Himmel ME. Lignocellulose deconstruction in the biosphere. *Curr Opin Chem Biol*. 2017;41:61–70.
- Andlar M, Rezac T, Mardetko N, Kracher D, Ludwig R, Santek B. Lignocellulose degradation: an overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng Life Sci*. 2018;18(11):768–78.
- Grondin JM, Tamura K, Dejean G, Abbott DW, Brumer H. Polysaccharide utilization loci: fueling microbial communities. *J Bacteriol*. 2017;199(15):15.
- Hoell IA, Vaaje-Kolstad G, Eijsink VGH: Structure and function of enzymes acting on chitin and chitosan. In: *Biotechnology and Genetic Engineering Reviews*, Vol 27. Edited by Harding SE, vol. 27. Nottingham: Nottingham Univ Press; 2010: 331–366.
- Arnosti C, Wietz M, Brinkhoff T, Hehemann JH, Probandt D, Zeugner L, Amann R: The Biogeochemistry of Marine Polysaccharides: Sources, Inventories, and Bacterial Drivers of the Carbohydrate Cycle. In: *Annual Review of Marine Science*, Vol 13, 2021. Edited by Carlson CA, Giovannoni SJ, vol. 13. Palo Alto: Annual Reviews; 2021: 81–108.
- Hehemann JH, Boraston AB, Czjzek M. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr Opin Struct Biol*. 2014;28:77–86.

20. Drula E, Garron ML, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 2022;50(D1):D571–7.
21. Marjamaa K, Kruus K. Enzyme biotechnology in degradation and modification of plant cell wall polymers. *Physiol Plant.* 2018;164(1):106–18.
22. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sorlie M, Eijsink VG. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science.* 2010;330(6001):219–22.
23. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry.* 2010;49(15):3305–16.
24. Johansen KS. Lytic polysaccharide monoxygenases: the microbial power tool for lignocellulose degradation. *Trends Plant Sci.* 2016;21(11):926–36.
25. Hu J, Tian D, Renneckar S, Saddler JN. Enzyme mediated nanofibrillation of cellulose by the synergistic actions of an endoglucanase, lytic polysaccharide monoxygenase (LPMO) and xylanase. *Sci Rep.* 2018;8(1):3195.
26. Koskela S, Wang SN, Fowler PMP, Tan FC, Zhou Q. Structure and self-assembly of lytic polysaccharide monoxygenase-oxidized cellulose nanocrystals. *ACS Sustain Chem Eng.* 2021;9(34):11331–41.
27. Koskela S, Wang SN, Xu DF, Yang X, Li K, Berglund LA, McKee LS, Bulone V, Zhou Q. Lytic polysaccharide monoxygenase (LPMO) mediated production of ultra-fine cellulose nanofibers from delignified softwood fibres. *Green Chem.* 2019;21(21):5924–33.
28. Chalak A, Villares A, Moreau C, Haon M, Grisel S, d'Orlando A, Herpoel-Gimbert I, Labourel A, Cathala B, Berrin JG. Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monoxygenase on cellulosic substrates. *Biotechnol Biofuels.* 2019;12(1):206.
29. Muraleedharan MN, Karnaouri A, Piatkova M, Ruiz-Caldas MX, Matsakas L, Liu B, Rova U, Christakopoulos P, Mathew AP. Isolation and modification of nano-scale cellulose from organosolv-treated birch through the synergistic activity of LPMO and endoglucanases. *Int J Biol Macromol.* 2021;183:101–9.
30. Valenzuela SV, Valls C, Schink V, Sanchez D, Roncero MB, Diaz P, Martinez J, Pastor FI. Differential activity of lytic polysaccharide monoxygenases on celluloses of different crystallinity. Effectiveness in the sustainable production of cellulose nanofibrils. *Carbohydr Polym.* 2019; 207:59–67.
31. Villares A, Moreau C, Bennati-Granier C, Garajova S, Foucat L, Falourd X, Saake B, Berrin JG, Cathala B. Lytic polysaccharide monoxygenases disrupt the cellulose fibers structure. *Sci Rep.* 2017;7:40262.
32. Solhi L, Li J, Li J, Heyns NMI, Brumer H. Oxidative enzyme activation of cellulose substrates for surface modification. *Green Chem.* 2022;24(10):4026–40.
33. Wang D, Li J, Salazar-Alvarez G, McKee LS, Srivastava V, Sellberg JA, Bulone V, Hsieh YS. Production of functionalised chitins assisted by fungal lytic polysaccharide monoxygenase. *Green Chem.* 2018;20(9):2091–100.
34. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZY database to integrate auxiliary redox enzymes. *Biotechnol Biofuels.* 2013;6(1):41.
35. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoel-Gimbert I, Chevreton D, Fanuel M, Henrissat B, Heiss-Blanquet S, Margeot A, et al. AA16, a new lytic polysaccharide monoxygenase family identified in fungal secretomes. *Biotechnol Biofuels.* 2019;12(1):55.
36. Sabbadin F, Urresti S, Henrissat B, Avrova AO, Welsh LRJ, Lindley PJ, Csukai M, Squires JN, Walton PH, Davies GJ, et al. Secreted pectin monoxygenases drive plant infection by pathogenic oomycetes. *Science.* 2021;373(6556):774–9.
37. Monclaro AV, Petrović DM, Alves GSC, Costa MMC, Midorikawa GEO, Miller RNG, Filho EXF, Eijsink VGH, Várnai A. Characterization of two family AA9 LPMOs from *Aspergillus tamarii* with distinct activities on xyloglucan reveals structural differences linked to cleavage specificity. *PLoS ONE.* 2020;15(7):e0235642.
38. 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 monoxygenase. *Nat Commun.* 2015;6:5961.
39. Tandrup T, Frandsen KEH, Johansen KS, Berrin JG, Lo Leggio L. Recent insights into lytic polysaccharide monoxygenases (LPMOs). *Biochem Soc Trans.* 2018;46(6):1431–47.
40. Bissaro B, Rohr AK, Muller G, Chylenski P, Skaugen M, Forsberg Z, Horn SJ, Vaaje-Kolstad G, Eijsink VGH. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H(2)O(2). *Nat Chem Biol.* 2017;13(10):1123–8.
41. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JC, Johansen KS, Krogh KB, Jorgensen 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(37):15079–84.
42. Vu VV, Beeson WT, Phillips CM, Cate JHD, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monoxygenases. *J Am Chem Soc.* 2014;136(2):562–5.
43. Huttner S, Várnai A, Petrovic DM, Bach CX, Kim Anh DT, Thanh VN, Eijsink VGH, Larsbrink J, Olsson L. Specific Xylan activity revealed for aa9 lytic polysaccharide monoxygenases of the thermophilic fungus *malbranchea cinnamomea* by functional characterization. *Appl Environ Microbiol.* 2019;85(23):e01408-01419.
44. Walton PH, Davies GJ. On the catalytic mechanisms of lytic polysaccharide monoxygenases. *Curr Opin Chem Biol.* 2016;31:195–207.
45. Forsberg Z, Sorlie M, Petrovic D, Courtade G, Aachmann FL, Vaaje-Kolstad G, Bissaro B, Rohr AK, Eijsink VG. Polysaccharide degradation by lytic polysaccharide monoxygenases. *Curr Opin Struct Biol.* 2019;59:54–64.
46. Lenfant N, Hainaut M, Terrapon N, Drula E, Lombard V, Henrissat B. A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res.* 2017;448:166–74.
47. Herve C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knox JP. Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci U S A.* 2010;107(34):15293–8.
48. Gilbert HJ, Knox JP, Boraston AB. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol.* 2013;23(5):669–77.
49. Várnai A, Mäkelä MR, Djajadi DT, Rahikainen J, Hatakka A, Viikari L: Chapter Four - Carbohydrate-Binding Modules of Fungal Cellulases: Occurrence in Nature, Function, and Relevance in Industrial Biomass Conversion. In: *Advances in Applied Microbiology*. Edited by Sariaslani S, Gadd GM, vol. 88: Academic Press; 2014: 103–165.
50. Crouch LI, Labourel A, Walton PH, Davies GJ, Gilbert HJ. The contribution of non-catalytic carbohydrate binding modules to the activity of lytic polysaccharide monoxygenases. *J Biol Chem.* 2016;291(14):7439–49.
51. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Várnai A, Rohr AK, Payne CM, Sorlie M, Sandgren M, et al. Structural and functional characterization of a lytic polysaccharide monoxygenase with broad substrate specificity. *J Biol Chem.* 2015;290(38):22955–69.
52. Danneels B, Tanghe M, Desmet T. Structural features on the substrate-binding surface of fungal lytic polysaccharide monoxygenases determine their oxidative regioselectivity. *Biotechnol J.* 2019;14(3):e1800211.
53. Hansson H, Karkehabadi S, Mikkelsen N, Douglas NR, Kim S, Lam A, Kaper T, Kelemen B, Meier KK, Jones SM, et al. High-resolution structure of a lytic polysaccharide monoxygenase from *Hypocrea jecorina* reveals a predicted linker as an integral part of the catalytic domain. *J Biol Chem.* 2017;292(46):19099–109.
54. Courtade G, Forsberg Z, Heggset EB, Eijsink VGH, Aachmann FL. The carbohydrate-binding module and linker of a modular lytic polysaccharide monoxygenase promote localized cellulose oxidation. *J Biol Chem.* 2018;293(34):13006–15.
55. The CAZyPedia Consortium. Ten years of CAZyPedia: a living encyclopedia of carbohydrate-active enzymes. *Glycobiology.* 2017;28(1):3–8.
56. Frandsen KEH, Tovborg M, Jorgensen CI, Spodsborg N, Rosso MN, Hemsworth GR, Garman EF, Grime GW, Poulsen JCN, Batth TS, et al. Insights into an unusual Auxiliary Activity 9 family member lacking the histidine brace motif of lytic polysaccharide monoxygenases. *J Biol Chem.* 2019;294(45):17117–30.
57. Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS ONE.* 2009;4(2):e4345.
58. Viborg AH, Terrapon N, Lombard V, Michel G, Czjzek M, Henrissat B, Brumer H. A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J Biol Chem.* 2019;294(44):15973–86.

59. Vandhana TM, Reyre JL, Sushmaa D, Berrin JG, Bissaro B, Madhuprakash J. On the expansion of biological functions of lytic polysaccharide monoxygenases. *New Phytol.* 2022;233(6):2380–96.
60. Frommhagen M, Westphal AH, Hilgers R, Koetsier MJ, Hinz SWA, Visser J, Gruppen H, van Berkel WJH, Kabel MA. Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monoxygenases. *Appl Microbiol Biotechnol.* 2018;102(3):1281–95.
61. Brander S, Tokin R, Ipsen JO, Jensen PE, Hernandez-Rollan C, Norholm MHH, Lo Leggio L, Dupree P, Johansen KS. Scission of Glucosidic Bonds by a *Lentinus similis* Lytic Polysaccharide Monoxygenases Is Strictly Dependent on H₂O₂ while the Oxidation of Saccharide Products Depends on O₂. *ACS Catal.* 2021;11(22):13848–59.
62. Chen KX, Zhang X, Long LK, Ding SJ. Comparison of C4-oxidizing and C1/C4-oxidizing AA9 LPMOs in substrate adsorption, H₂O₂-driven activity and synergy with cellulase on celluloses of different crystallinity. *Carbohydr Polym.* 2021;269:13.
63. Sun PC, Valenzuela SV, Chunkruea P, Pastor FIJ, Laurent C, Ludwig R, Van Berkel WJH, Kabel MA. Oxidized product profiles of aa9 lytic polysaccharide monoxygenases depend on the type of cellulose. *ACS Sustain Chem Eng.* 2021;9(42):14124–33.
64. Frommhagen M, Koetsier MJ, Westphal AH, Visser J, Hinz SW, Vincken JP, van Berkel WJ, Kabel MA, Gruppen H. Lytic polysaccharide monoxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol Biofuels.* 2016;9(1):186.
65. Li J, Goddard-Borger ED, Raji O, Saxena H, Solhi L, Mathieu Y, Master ER, Wakarchuk WW, Brumer H. Chitin-active lytic polysaccharide monoxygenases are rare in cellulomonas species. *Appl Environ Microbiol.* 2022;88(15):e0096822.
66. Vuong TV, Liu B, Sandgren M, Master ER. Microplate-based detection of lytic polysaccharide monoxygenase activity by fluorescence-labeling of insoluble oxidized products. *Biomacromol.* 2017;18(2):610–6.
67. Bey M, Zhou S, Poidevin L, Henriissat B, Coutinho PM, Berrin JG, Sigoillot JC. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monoxygenases (family GH61) from *Podospira anserina*. *Appl Environ Microbiol.* 2013;79(2):488–96.
68. Zhang YH, Cui J, Lynd LR, Kuang LR. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromol.* 2006;7(2):644–8.
69. Li WZ, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006;22(13):1658–9.
70. Mewis K, Lenfant N, Lombard V, Henriissat B. Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. *Appl Environ Microbiol.* 2016;82(6):1686–92.
71. Hagberg A, Swart P, S Chult D. Exploring network structure, dynamics, and function using NetworkX. In: Los Alamos National Lab.(LANL), Los Alamos, NM (United States); 2008.
72. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498–504.
73. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772–80.
74. Miller MA, Pfeiffer W, Schwartz T: Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE): 14–14 Nov. 2010 2010. 1–8.
75. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics.* 2014;30(22):3276–8.
76. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3.
77. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49(W1):W293–6.
78. Addleman K, Archibald F. Kraft pulp bleaching and delignification by dikaryons and monokaryons of *Trametes versicolor*. *Appl Environ Microbiol.* 1993;59(1):266–73.
79. Bellemare A, John T, Marquetteau S. Fungal genomic DNA extraction methods for rapid genotyping and genome sequencing. *Methods Mol Biol.* 2018;1775:11–20.
80. Storms R, Zheng Y, Li HS, Sillaots S, Martinez-Perez A, Tsang A. Plasmid vectors for protein production, gene expression and molecular manipulations in *Aspergillus niger*. *Plasmid.* 2005;53(3):191–204.
81. Song L, Ouedraogo J-P, Kolbusz M, Nguyen TTM, Tsang A. Efficient genome editing using tRNA promoter-driven CRISPR/Cas9 gRNA in *Aspergillus niger*. *PLoS ONE.* 2018;13(8):e0202868.
82. Debets A, Bos C. Isolation of small protoplasts from *Aspergillus niger*. *Fungal Genetics Newsletter.* 1986;33:24.
83. Käfer E. Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. *Adv Genet.* 1977;19:33–131.
84. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K: Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat Biotechnol* 2007, 25(2):221–231.
85. Meyer V, Arentshorst M, El-Ghezal A, Drews A-C, Kooistra R, van den Hondel CA, Ram AF. Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. *J Biotechnol.* 2007;128(4):770–5.
86. Punt PJ, Schuren FH, Lehmebeck J, Christensen T, Hjort C, van den Hondel CA. Characterization of the *Aspergillus niger* prtT, a unique regulator of extracellular protease encoding genes. *Fungal Genet Biol.* 2008;45(12):1591–9.
87. Niu J, Arentshorst M, Nair PDS, Dai Z, Baker SE, Frisvad JC, Nielsen KF, Punt PJ, Ram AF: Identification of a classical mutant in the industrial host *Aspergillus niger* by systems genetics: LaeA is required for citric acid production and regulates the formation of some secondary metabolites. *G3: Genes, Genomes, Genetics* 2016, 6(1):193–204.
88. Master ER, Zheng Y, Storms R, Tsang A, Powlowski J. A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization. *Biochemical Journal.* 2008;411(1):161–70.
89. Venegas FA, Koutaniemi S, Langeveld SM, Bellemare A, Chong S-L, Dilokpimol A, Lowden MJ, Hilden KS, Leyva-Illades JF, Mäkelä MR. Carbohydrate esterase family 16 contains fungal hemicellulose acetyl esterases (HAEs) with varying specificity. *New Biotechnol.* 2022;70:28–38.
90. Budak SO, Zhou M, Brouwer C, Wiebenga A, Benoit I, Di Falco M, Tsang A, de Vries RP. A genomic survey of proteases in *Aspergilli*. *BMC Genomics.* 2014;15(1):1–15.
91. Westereng B, Arntzen MO, Aachmann FL, Varnai A, Eijsink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monoxygenases acting on cellulose. *J Chromatogr A.* 2016;1445:46–54.
92. Li J, Solhi L, Goddard-Borger ED, Mathieu Y, Wakarchuk WW, Withers SG, Brumer H. Four cellulose-active lytic polysaccharide monoxygenases from *Cellulomonas* species. *Biotechnol Biofuels.* 2021;14(1):29.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

