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

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

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

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

Supporting information

Table of Contents:

Materials and Methods

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table S1. Structural data processing and refinement statistics.

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

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

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

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

[d] Defined using MolProbity (31).

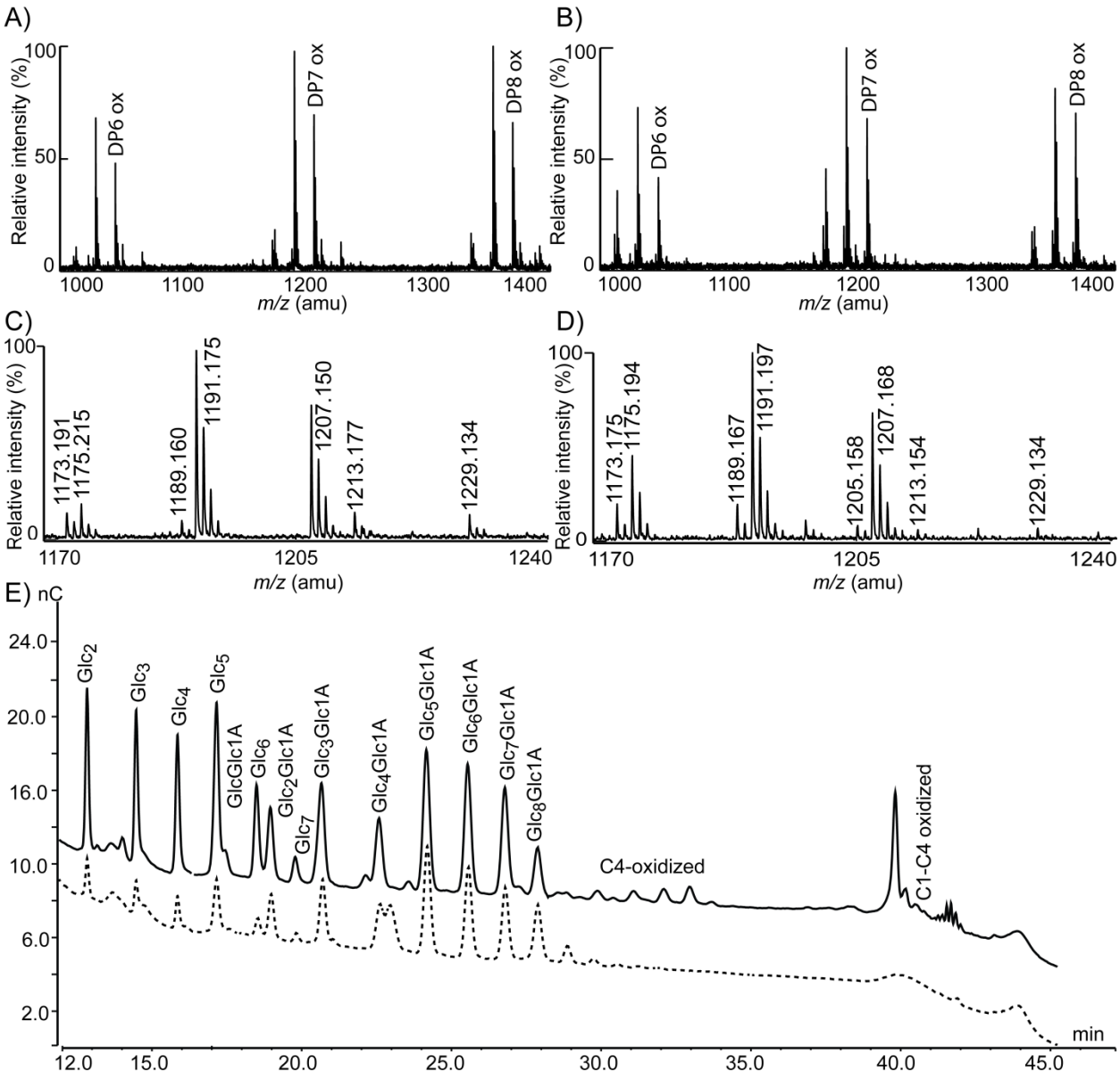

Figure S1. Product profiles from PASC degradation experiments with the *Thermobifida fusca* LPMOs E7 and E8-N, analyzed by MALDI-TOF MS (A-D) and HPAEC (E). MALDI TOF MS of E8-N (A and C) and E7 (B and D) shows that E8-N is a strict C1-oxidizing LPMO as CelS2, whereas E7 produces a mixture of C1-, C4-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

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

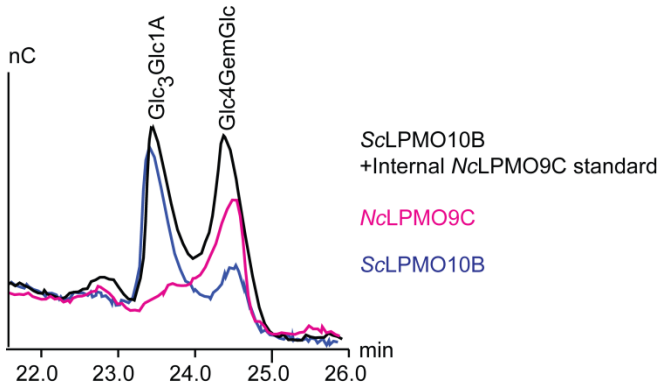

Figure S2. Evidence for C4 oxidation by *Sc*LPMO10B. The chromatogram shows product analysis of the following samples: Magenta, degradation of Glc_5 with *Nc*LPMO9C (sample size 0.2 µL); the product shown has been identified (by MS and, importantly, NMR) as Glc4GemGlc [C4 oxidized dimer; (32)]. Blue: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

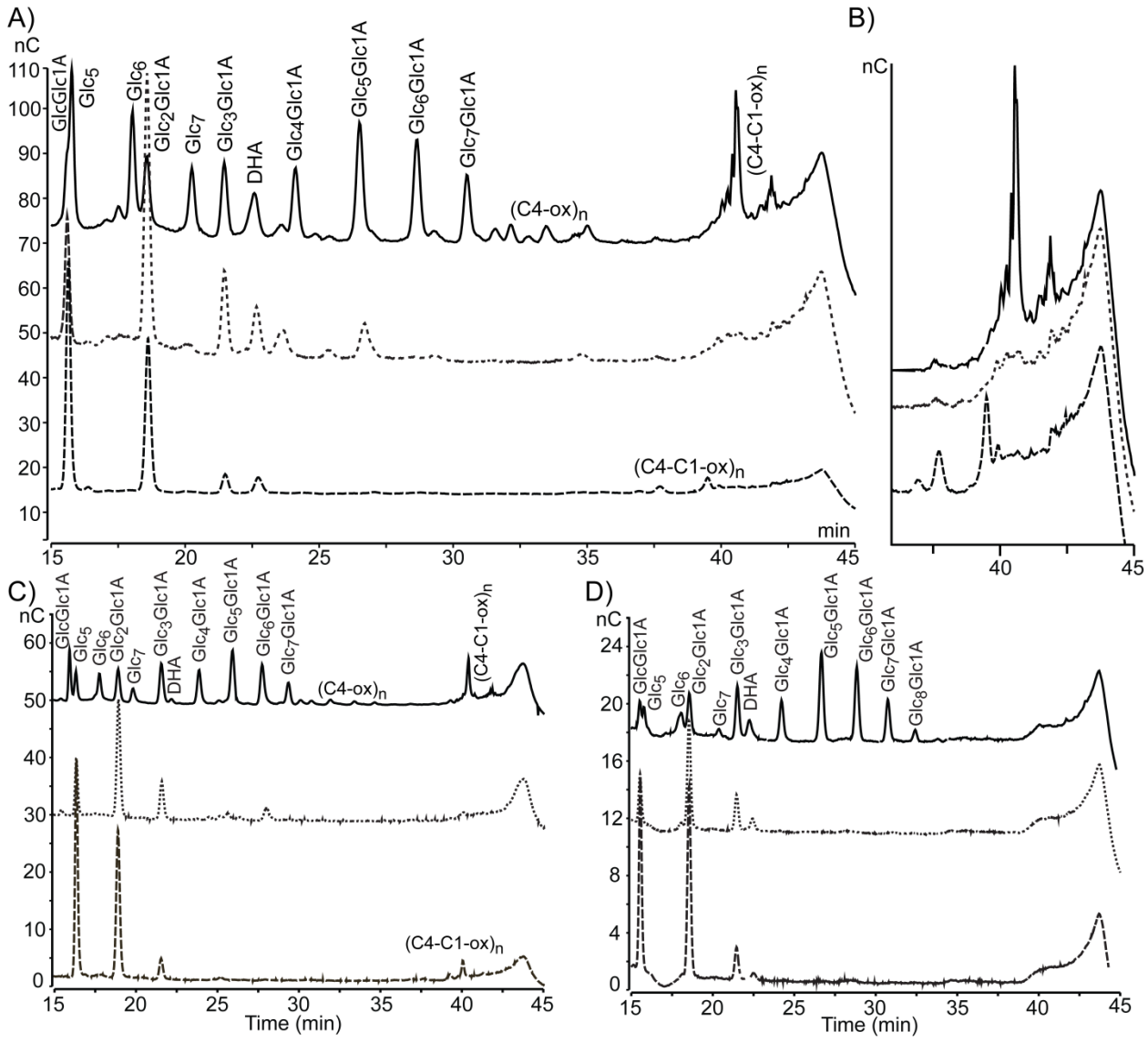

Figure S3. HPAEC chromatograms to verify the occurrence of double oxidized products formed by *Sc*LPMO10B (A-B) and E7 (C). First, *Sc*LPMO10B or E7 was incubated with PASC over night in the presence of ascorbate; subsequently the soluble fraction was taken and analyzed by HPAEC (panel A and C, top chromatogram).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

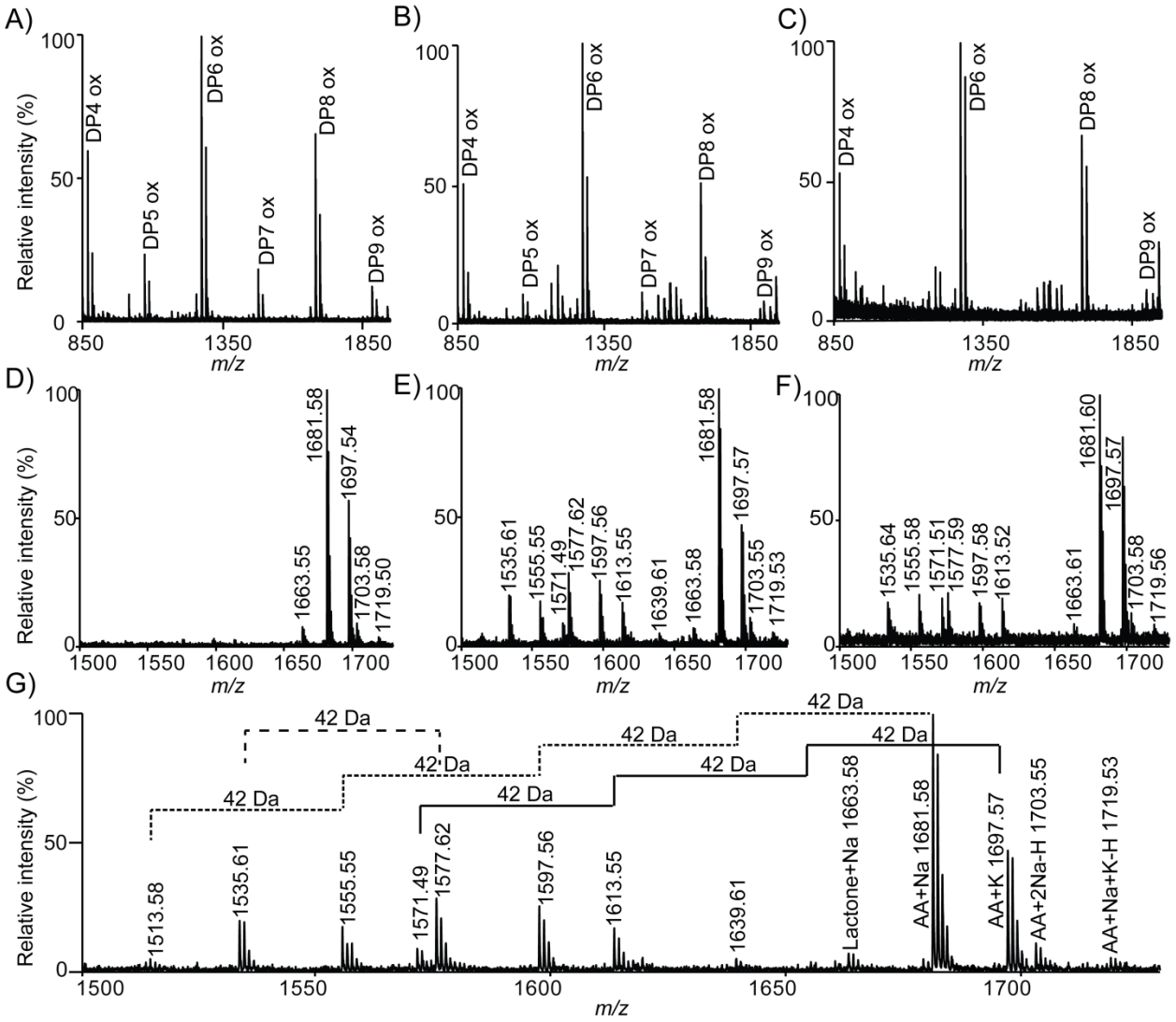

Figure S4. MALDI-TOF MS spectra showing LPMO activity on squid pen β-chitin for *S. marcescens* CBP21 (A, D), *Sc*LPMO10B (B, E), and E7 (C, F). Panels A-C show oxidized products ranging from DP 4-9, whereas panels D-F show a zoom in on the octamer ion cluster. Panel G shows more details of the *Sc*LPMO10B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

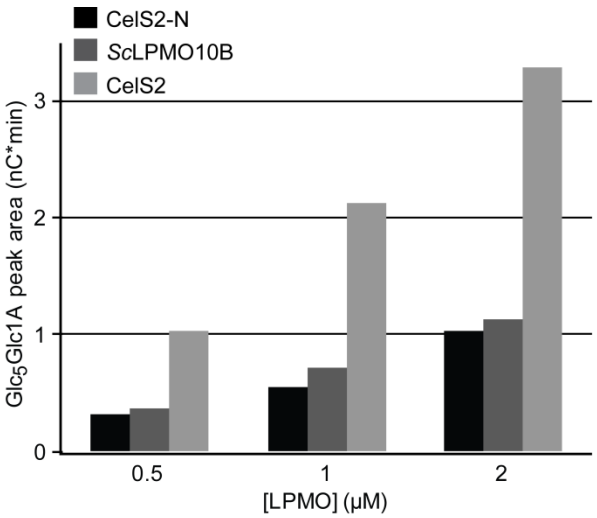

Figure S5. Dose-response effects for CelS2-N (black), *Sc*LPMO10B (dark grey) and full length CelS2 (light grey). Products were generated by incubation of PASC (2 g/L) with the indicated LPMO concentrations, in a 20 mM ammonium acetate buffer pH 6.0, at 50 °C, followed by cellobiohydrolase digestion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

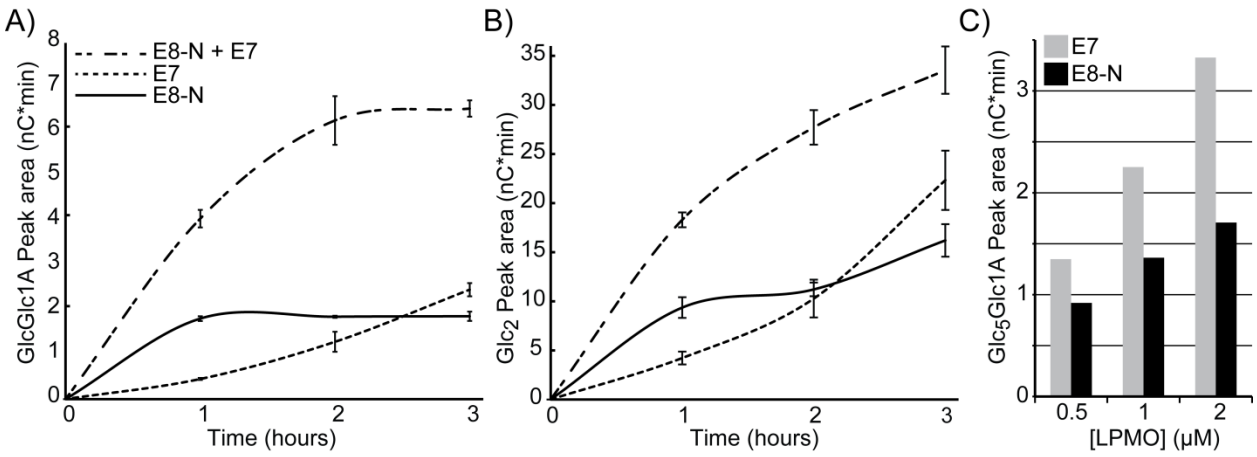

Figure S6. Time course of products released from PASC by the *T. fusca* LPMO pair. The figures show (A) cellobionic acid and (B) cellobiose obtained after incubation of 1 μM LPMO (E8-N or E7) or a mixture of 0.5 μM E8-N and 0.5 μM E7 with 2 g/L PASC and 2 mM ascorbic acid in 20 mM ammonium acetate buffer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

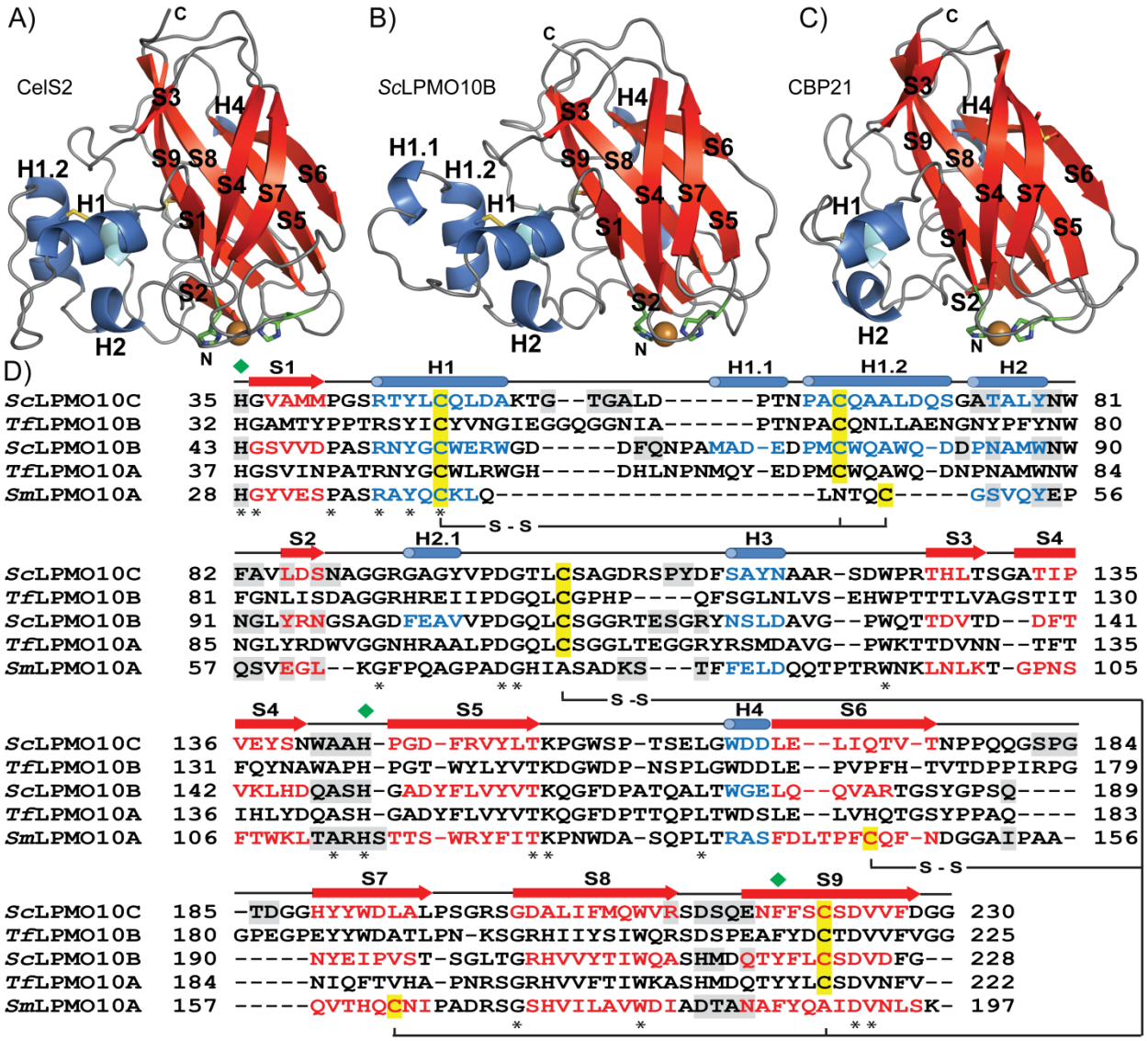

Figure S7. Structural overview of cellulose oxidizing LPMOs from *S. coelicolor* and comparison with the chitin active LPMO from *S. marcescens*. (A-C) Cartoon representations; secondary structure elements are shown in red, for β-strands, and blue, for α-helices. Metal ions are shown as orange spheres,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

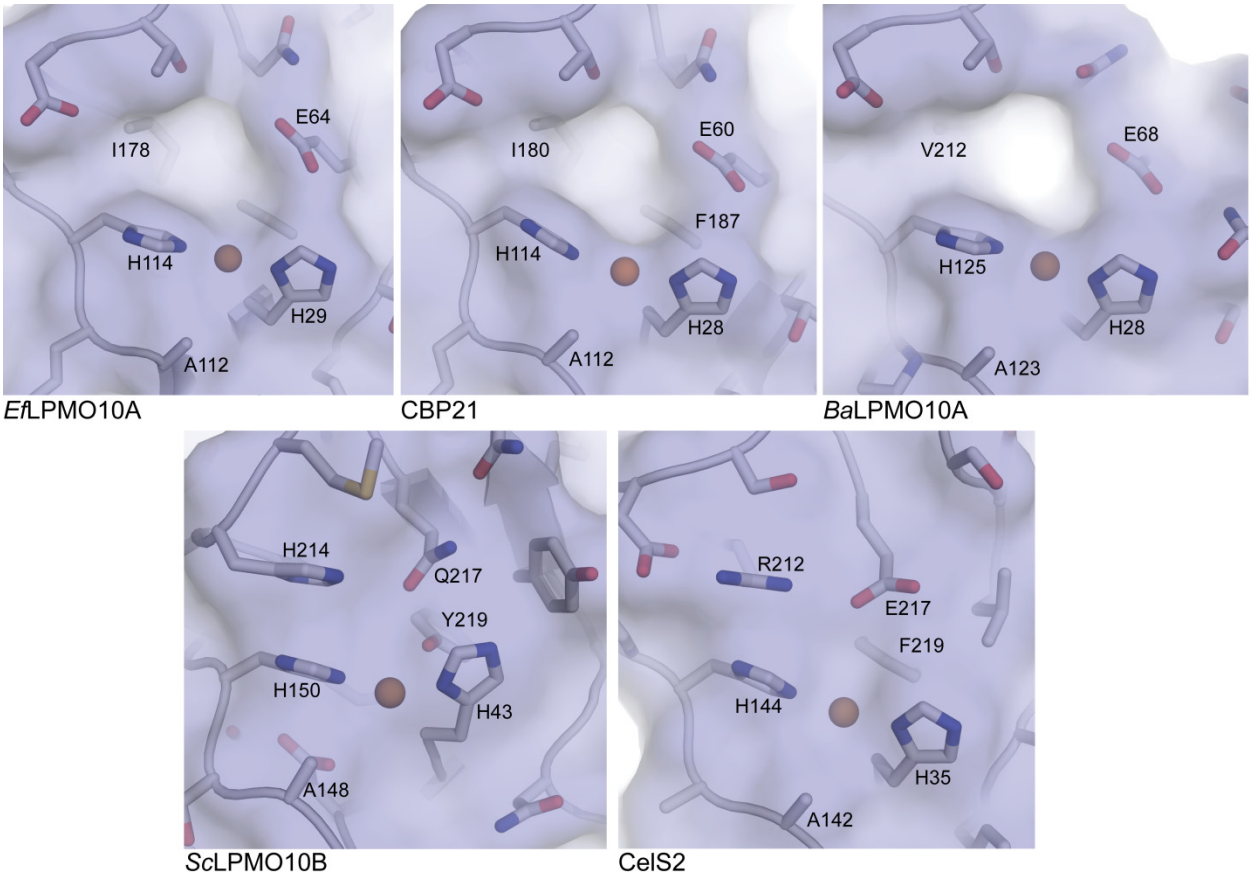

Figure S8. Conserved cavities on the substrate binding surface of chitin active LPMO10s. The picture shows the surfaces of three chitin-active LPMOs (*Ef*LPMO10A, CBP21 (*Sm*LPMO10A), and *Ba*LPMO10A; no activity data have been published for the latter but is known that this protein has high affinity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

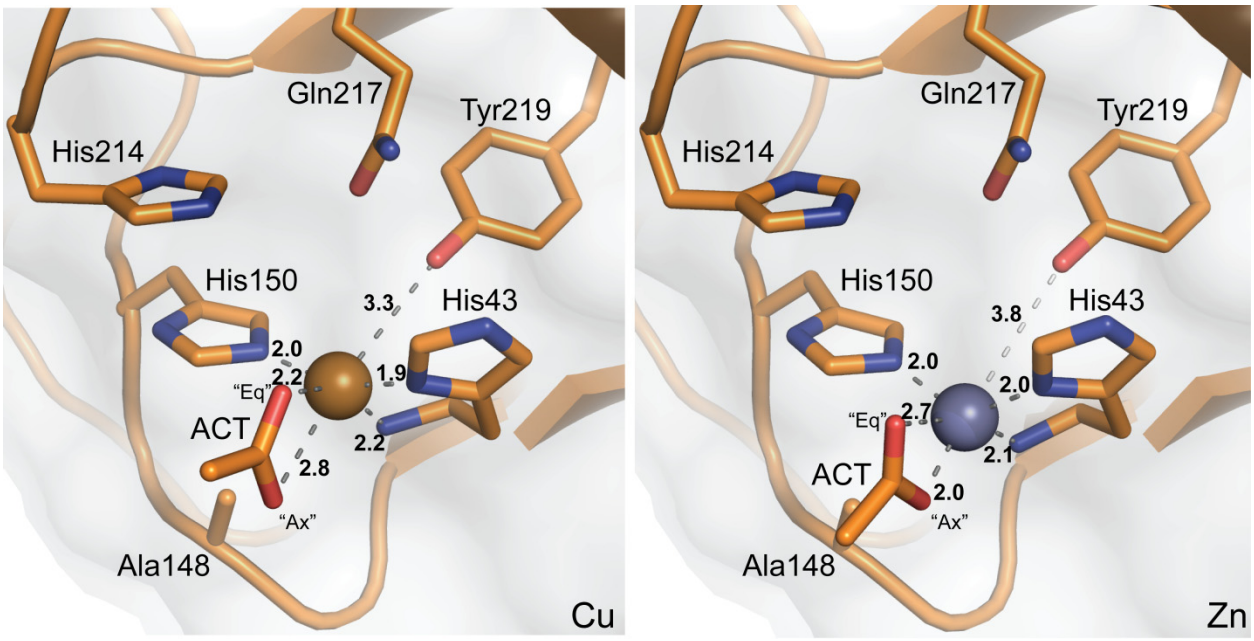

Figure S9. Active site of *Sc*LPMO10B with bound copper (left panel) or zinc (right panel). The two oxygen moieties of the acetate ion (ACT) are bound in the solvent-facing axial position (labeled "Ax") and in a slightly skewed equatorial position (labeled "Eq"). The structural data demonstrate the capability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

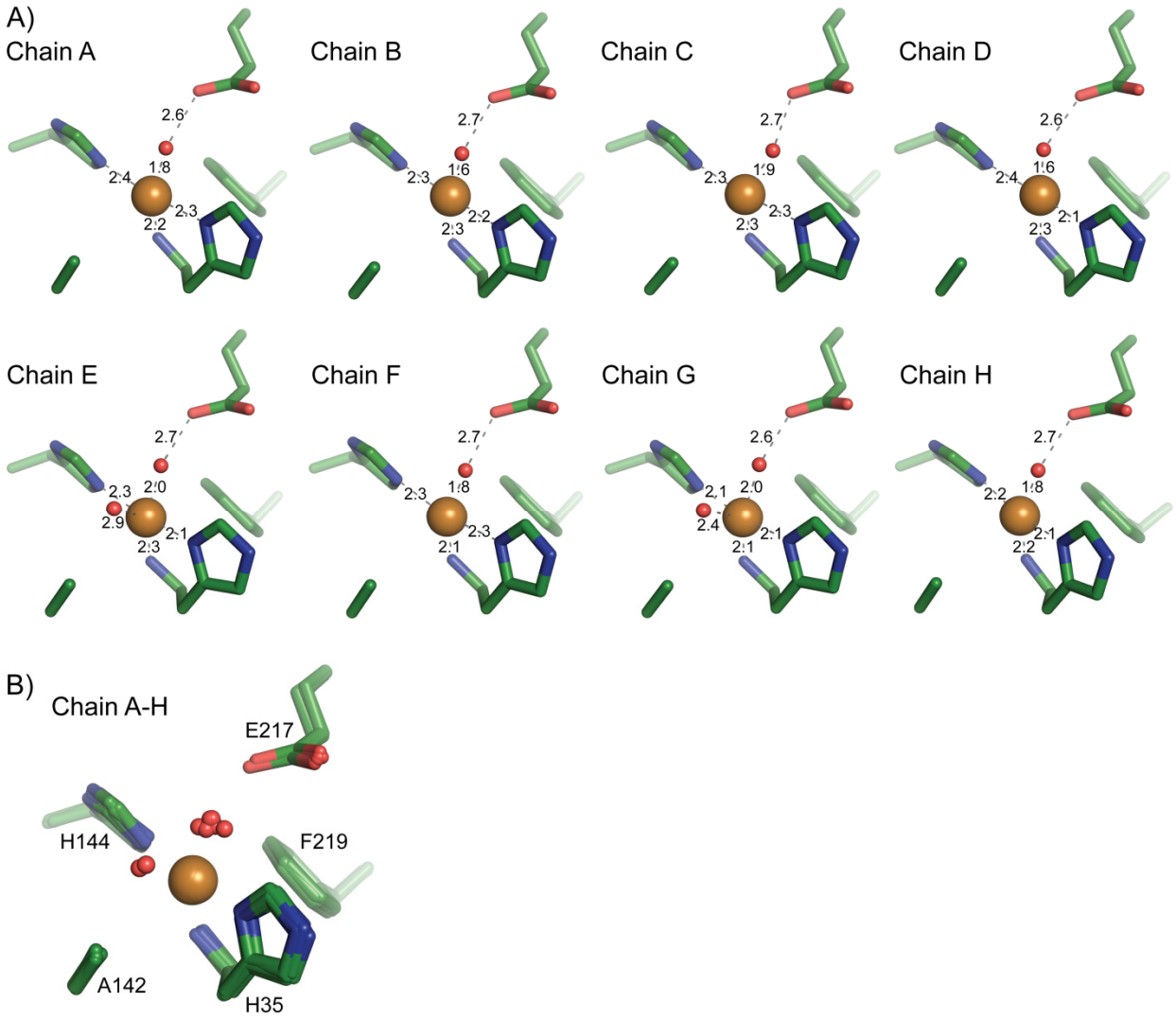

Figure S10. Active sites in all eight CelS2 molecules in the asymmetric unit. Panel (A) shows the individual sites, whereas panel (B) shows a superposition. Amino acid side chains are shown in stick representation. The water molecules and copper ions are shown as red and golden spheres, respectively.

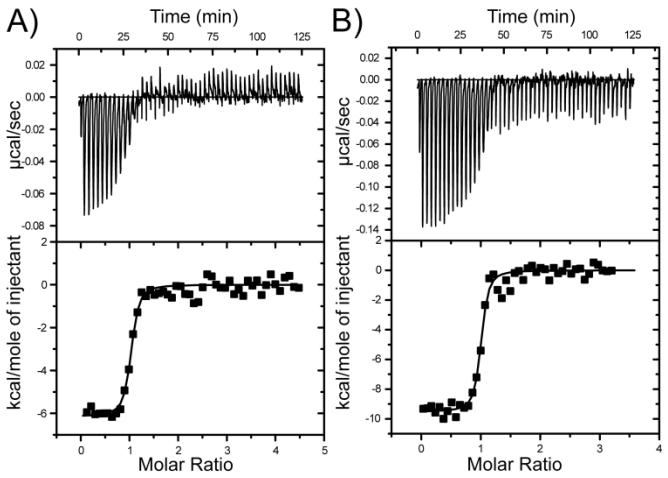

Figure S11. Thermograms (upper) and binding isotherms with theoretical fits (lower) for A) titration of 150 µM Cu^{2+} to 5 µM apo-CelS2 and B) titration of 120 µM Cu^{2+} to 5 µM apo-*Sc*LPMO10B in MES buffer pH 5.5 at 10 °C.

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

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

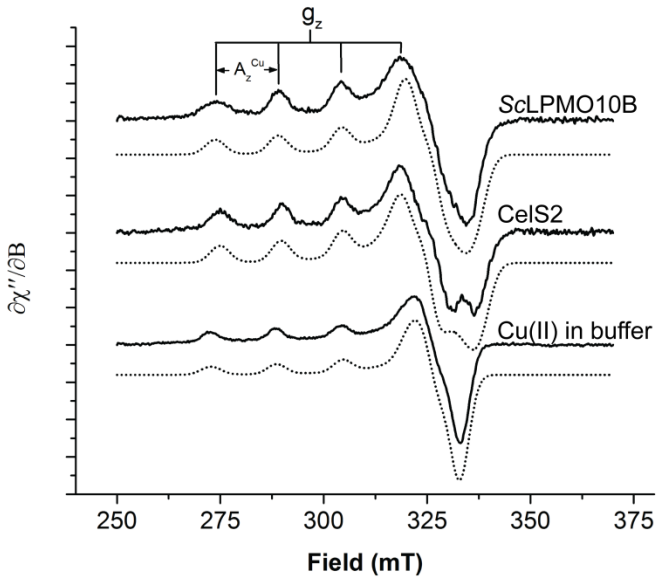

Figure S13. X-band EPR spectra (solid lines) with simulations (dotted lines) for *Sc*LPMO10B, full length CelS2 and Cu(II) in buffer. The EPR spectra were recorded at 77 K using a microwave power of 0.5 mW.

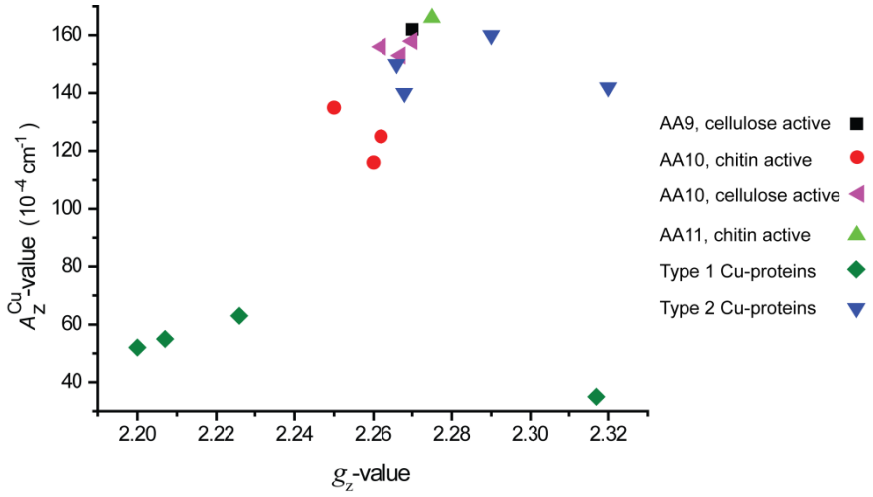

Figure S14. A Peisach – Blumberg plot of different LPMOs compared to selected typical type 1 and type 2 copper proteins. Shown are: chitin-active *Sm*LPMO10A (CBP21) (35), *Bl*LPMO10A (35) and presumably chitin-active *Ba*LPMO10A (36) (●); cellulose-active *Ta*LPMO9A (37) (■), *Tf*LPMO10B (E8-N) (35),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

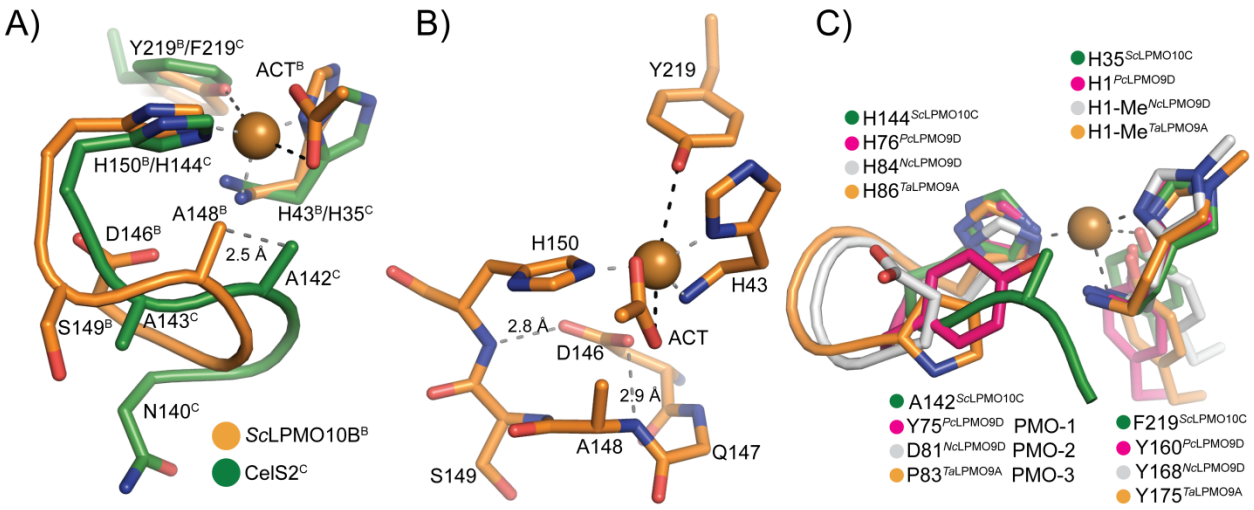

Figure S15. Position of the conserved active site alanine. (A; identical to Fig. 5 in the main manuscript). Structural superposition of CelS2 (green) and *Sc*LPMO10B (orange) active sites (superposed using the "histidine brace" and the metal ion), with side chains shown in stick representation and back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

References

1. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VGH, van Aalten DM (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* 280(12):11313-11319.
2. Manoil C, Beckwith J (1986) A Genetic Approach to Analyzing Membrane-Protein Topology. *Science* 233(4771):1403-1408.
3. Wood TM (1988) Preparation of Crystalline, Amorphous, and Dyed Cellulase Substrates. *Method Enzymol* 160:19-25.
4. Harreither W, et al. (2011) Catalytic properties and classification of cellobiose dehydrogenases from ascomycetes. *Appl Environ Microbiol* 77(5):1804-1815.
5. Vaaje-Kolstad G, et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330(6001):219-222.
6. Westereng B, et al. (2013) Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J Chromatogr A* 1271(1):144-152.
7. Kabsch W (2010) Xds. *Acta Crystallogr D* 66(2):125-132.
8. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62(1):72-82.
9. Matthews BW (1968) Solvent Content of Protein Crystals. *J Mol Biol* 33(2):491-497.
10. Kantardjieff KA, Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* 12(9):1865-1871.
11. Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D* 66(4):479-485.
12. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of *Coot*. *Acta Crystallogr D* 66(4):486-501.
13. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171-1179.
14. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* 67(4):355-367.
15. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D* 67(4):271-281.
16. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53(3):240-255.
17. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66(2):213-221.
18. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* 68(4):352-367.
19. Lamzin VS, Wilson KS (1993) Automated refinement of protein models. *Acta Crystallogr D* 49(1):129-147.
20. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D* 63(1):32-41.
21. Aachmann FL, Sørlie M, Skjåk-Braek G, Eijsink VGH, Vaaje-Kolstad G (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding,

protein dynamics, and substrate interactions. *Proc Natl Acad Sci U S A* 109(46):18779-18784.
22. Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. *Anal Biochem* 287(1):118-125.
23. Liu Y, Seefeldt LC, Parker VD (1997) Entropies of redox reactions between proteins and mediators: the temperature dependence of reversible electrode potentials in aqueous buffers. *Anal Biochem* 250(2):196-202.
24. Wiseman T, Williston S, Brandts JF, Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179(1):131-137.
25. Stoll S, Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J Magn Reson* 178(1):42-55.
26. Bramucci E, Paiardini A, Bossa F, Pascarella S (2012) PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics* 13 Suppl 4:S2.
27. Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 11(9):739-747.
28. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.
29. Dereeper A, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465-4699.
30. Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* 4(4):269-275.
31. Chen VB, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* 66(1):12-21.
32. Isaksen T, et al. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem* 289(5):2632-2642.
33. Forsberg Z, et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20(9):1479-1483.
34. Chu HH, Hoang V, Hofemeister J, Schrempf H (2001) A *Bacillus amyloliquefaciens* ChbB protein binds beta- and alpha-chitin and has homologues in related strains. *Microbiology* 147(7):1793-1803.
35. Forsberg Z, et al. (2014) Comparative Study of Two Chitin-Active and Two Cellulose-Active AA10-Type Lytic Polysaccharide Monooxygenases. *Biochemistry* 53(10):1647-1656.
36. Hemsworth GR, et al. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc* 135(16):6069-6077.
37. Quinlan RJ, et al. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci U S A* 108(37):15079-15084.
38. Hemsworth GR, Henrissat B, Davies GJ, Walton PH (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* 10(2):122-126.
39. Rotilio G, et al. (1971) Studies of the metal sites of copper proteins. Ligands of copper in hemocuprein. *Biochemistry* 10(4):616-621.

40. Buffoni F, Corte LD, Knowles PF (1968) The nature of copper in pig plasma benzylamine oxidase. *Biochem J* 106(2):575-576.
41. Mondovi B, et al. (1967) Diamine oxidase from pig kidney. Improved purification and properties. *J Biol Chem* 242(6):1160-1167.
42. Wishnick M, Lane MD, Scrutton MC, Mildvan AS (1969) The presence of tightly bound copper in ribulose diphosphate carboxylase from spinach. *J Biol Chem* 244(20):5761-5763.
43. Stigbrand T, Malmstrom BG, Vanngard T (1971) On the state of copper in the blue protein umecyanin. *FEBS Lett* 12(5):260-262.
44. Gewirth AA, Cohen SL, Schugar HJ, Solomon EI (1987) Spectroscopic and Theoretical-Studies of the Unusual Electron-Paramagnetic-Res Parameters of Distorted Tetrahedral Cupric Sites - Correlations to X-Ray Spectral Features of Core Levels. *Inorg Chem* 26(7):1133-1146.
45. Xie XJ, et al. (2009) A variable temperature spectroscopic study on *Paracoccus pantotrophus* pseudoazurin: Protein constraints on the blue Cu site. *J Inorg Biochem* 103(10):1307-1313.
46. LaCroix LB, et al. (1998) Spectroscopic and geometric variations in perturbed blue copper centers: Electronic structures of stellacyanin and cucumber basic protein. *J Am Chem Soc* 120(37):9621-9631.
47. Vu VV, Beeson WT, Phillips CM, Cate JH, Marletta MA (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc* 136(2):562-565.
48. Harris PV, et al. (2010) Stimulation of Lignocellulosic Biomass Hydrolysis by Proteins of Glycoside Hydrolase Family 61: Structure and Function of a Large, Enigmatic Family. *Biochemistry* 49(15):3305-3316.