Supplementary information

Lytic xylan oxidases from wood-decay fungi unlock biomass degradation

Marie Couturier¹, Simon Ladevèze¹, Gerlind Sulzenbacher²,³, Luisa Ciano⁴, Mathieu Fanuel⁵, Céline Moreau⁵, Ana Villares⁵, Bernard Cathala⁵, Florence Chaspoul⁶, Kristian Frandsen¹, Aurore Labourel¹, Isabelle Herpoël-Gimbert¹, Sacha Grisel¹, Mireille Haon¹, Nicolas Lenfant², Hélène Rogniaux⁵, David Ropartz⁵, Gideon Davies⁴, Marie-Noëlle Rosso¹, Paul H. Walton⁴, Bernard Henrissat²,³,⁷, Jean-Guy Berrin¹,*
Supplementary Figure 1. Consensus sequence alignment of 283 AA14 sequences reveals the first Histidine as a conserved residue amongst the family.
Supplementary Figure 2. Phylogenetic tree of the AA14 family. AA14 family members strongly cluster together and are very distant from AA9, AA10, AA11 and AA13 sequences. Sequences belonging to ascomycetes and basidiomycetes fungi are denoted in red and blue, respectively.
Supplementary Figure 3. SDS-PAGE analysis of purified PcAA14 enzymes. Lane A, PcAA14A (5 µg); lane B, PcAA14B (5 µg); M, Molecular weight protein ladder.
Supplementary Figure 4. SDS-PAGE analysis of *PcAA14B* control samples (lanes 1-2) and endoHf-deglycosylated *PcAA14B* samples (lanes 3-4) as described in material and methods.
Supplementary Figure 5. Structural analysis of PcAA14B. A. Stereo-view of the active site of PcAA14B. The protein backbone is represented as a cartoon in grey. The active site residues are represented as stick in yellow. A weighted 2Fo-Fc electron density map contoured at 1.0 σ is shown in blue. B. View of the five N-glycans (Asn13, Asn76, Asn133, Asn183, Asn217) attached to PcAA14B. C. View of the PcAA14B surface with a clamp formed by two prominent surface loops. D. Comparison of PcAA14B (grey and red) and TaAA9A (grey and blue) with the
His braces shown as sticks. The L2 loop region of *TaAA9A* (residue 1-41) comprises secondary structure elements. In contrast, the equivalent region of *PcAA14B* forms an extended loop devoid of secondary structure elements. E. View of the extended L3 loop region of *PcAA14B* (cyan) compared with the short loop of *NcAA9C* (green), with the cystine (C67-C90) bordering the insertion.
**Supplementary Figure 6.** A. Continuous wave Q-band EPR spectrum (34.7 GHz, 113 K) with simulation (red) of PcAA14A in 50 mM sodium acetate buffer pH 5.2. B. Line diagram of PcAA14 LPMO active site.
Supplementary Figure 7. HPAEC-PAD profiles of the polysaccharide substrates assayed for PcAA14A activity screening. Oligosaccharide release and control curves of PcAA14A assayed on wheat arabinoxylan (A), arabinan (B), barley β-glucan (C), curdlan (D), konjac glucomannan (E), tamarind xyloglucan (F), starch (G), Avicel (H), chitin (I), laminarin (J), PASC (K) and birchwood xylan (L). Control curves of the substrates with 1 mM ascorbic acid are shown in black, substrates in the presence of 1 µM PcAA14A are shown in red and substrates in presence of 1 µM PcAA14A with ascorbic acid are shown in blue. Analysis of chromatograms revealed that PcAA14A was not active on any of these substrates. Similar experiments were carried out with PcAA14B. Peaks eluted at 10-12 min and 26 minutes are mainly due to the ascorbic acid.
Supplementary Figure 8. Additional saccharification assays. Saccharification assays in the presence and absence of ascorbate on pine (A), poplar (B), and wheat straw (C) where PcAA14A and PcAA14B were added to a concentration of 1 µM. (D) Assays on pretreated pine using a T. reesei cocktail enriched in AA9 LPMO. Assays were run in the presence of either 0.5 µM of PcAA14A or 0.5 µM of PaLPMO9E or 0.5 µM of PcAA14A and 0.5 µM of PaLPMO9E. (E) Comparative analysis of saccharification assays on pretreated pine in the presence of 1 µM PcAA14A or 1 µM PcAA14A bearing a CBM1 module (PcAA14A-CBM1). Error bars indicate
standard errors of the mean from triplicate independent experiments. Data points are shown as dots.
Supplementary Figure 9. Microscopy analyses of enzyme-treated wood fibers. TEM images of negatively stained fibers from birchwood (A) and pine (C), and AFM topography images of birchwood (B) and pine (D) fibers treated with PcAA14B.
Supplementary Figure 10. Solid-state $^{13}$C CP/MAS NMR analysis. Deconvolution of the C-1 and C4 region with crystalline forms Cr (Iα), Cr (Iβ) and Cr (Iα +Iβ) (black), para-crystalline form (PCr) (grey), accessible fibril surfaces (AS) (green), inaccessible fibril surface (IAS) (dark green) and hemicelluloses (HC) (xylan) (light orange) signals for reference cellulose fibers (A) and after enzyme treatment with PcAA14A (B) and PcAA14B (C). The table indicates the relative proportion of cellulose and hemicellulose content from C1 and C4 region and values of crystallinity, accessible/total fibril surface ratio (AS/(AS + IAS)), lateral fibril dimensions (LFD) and lateral fibril aggregate dimensions (LFAD) calculated from the C4 region deconvolution of the solid-state $^{13}$C CP/MAS NMR spectra of reference and PcAA14-treated cellulose samples.

<table>
<thead>
<tr>
<th></th>
<th>Reference</th>
<th>PcAAxxA</th>
<th>PcAAxxB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cellulose (%)</strong></td>
<td>97.34 (0.56)</td>
<td>98.34 (0.34)</td>
<td>98.11 (0.13)</td>
</tr>
<tr>
<td><strong>Hemicellulose (%)</strong></td>
<td>2.66 (0.56)</td>
<td>1.66 (0.34)</td>
<td>1.89 (0.13)</td>
</tr>
<tr>
<td><strong>Crystallinity</strong></td>
<td>53.04 (0.89)</td>
<td>57.10 (2.19)</td>
<td>51.95 (0.96)</td>
</tr>
<tr>
<td><strong>AS/(AS + IAS)</strong></td>
<td>40.65 (2.42)</td>
<td>40.18 (5.95)</td>
<td>45.75 (4.38)</td>
</tr>
<tr>
<td><strong>LFD (nm)</strong></td>
<td>2.15 (0.04)</td>
<td>2.00 (0.08)</td>
<td>2.19 (0.04)</td>
</tr>
<tr>
<td><strong>LFAD (nm)</strong></td>
<td>11.39 (0.88)</td>
<td>11.41 (1.62)</td>
<td>9.84 (0.93)</td>
</tr>
</tbody>
</table>

Nature Chemical Biology: doi:10.1038/nchembio.2558
Supplementary Figure 11. Synergy assays between *PcAA14A* and a GH11 xylanase on birchwood cellulose fibers. Control birchwood cellulose fibers are shown in black; C1-oxidized X2, X3 and X4 standards in blue; birchwood cellulose fibers incubated with 0.1 µM GH11 xylanase and 1 mM L-cysteine in red; birchwood cellulose fibers incubated with 0.1 µM GH11 xylanase, 1 µM *PcAA14A* and 1 mM L-cysteine in green. L-cystein was a preferred electron donor as it reduced peaks generated by ascorbic acid. All assays were carried out in triplicate.
Supplementary Figure 12. MS analysis of products released after enzymatic treatment of cellulosic fibers. Positive and negative polarity ESI-MS spectra of the degradation products generated from birchwood cellulosic fibers by *Pc*AA14A in synergy with a GH11 xylanase (A: positive ion mode, B: negative ion mode). The putative oxidized species (X$_3^{ox}$) observed at 429 m/z in panel B (indicated by a red arrow) was fragmented by tandem MS (MS/MS). Panel C shows the corresponding fragmentation spectrum (top trace), and compares this spectrum to that
obtained from a C1 oxidized Xyl₃ (bottom trace). ▽: water losses. □: H₂CO losses. The two fragmentation patterns unambiguously match, which implies that *PcAA14A* produces C1-oxidized xylo-oligosaccharides. The structure of the C1 oxidized standard was confirmed by two different ways. Firstly, by comparing its fragmentation pattern with the one shown in⁴⁹ to be a C1 oxidized oligosaccharide with an aldonic acid on the reducing end. Secondly, by observing, on its positive MS/MS spectrum, fragments reported in⁵⁰ as characteristic of an acidic function at the C1 of the reducing end. All the observed fragments are represented on the structure in panel D.
Supplementary Figure 13. Fermentation profile of the production of PcAA14A in Pichia pastoris. Time profiles of dissolved oxygen and dry cell weight (DCW) in a 1.3-L bioreactor fermentation. The inset represents SDS-PAGE analysis of secreted PcAA14A enzyme. 30 µL of culture supernatant were loaded at 124 h (lane 1), 148 h (lane 2), 165 h (lane 3) and 172 h (lane 4); M, Molecular weight protein ladder. The yield of PcAA14A enzyme produced was estimated at ~ 1 g.L⁻¹.
**Supplementary Tables**

**Supplementary Table 1. Biochemical features of \( PcAA14A \) and \( PcAA14B \) LPMOs.**
Glycosylation sites were predicted using the Hirst group glycosylation prediction server.\(^5\) Experimental mass and N-terminal sequences were determined using Maldi-MS and Edman degradation, respectively, as described in materials and methods.

<table>
<thead>
<tr>
<th></th>
<th>( PcAA14A ) with (His)(_6)-tag (no tag)</th>
<th>( PcAA14B ) with (His)(_6)-tag (no tag)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical molecular</td>
<td>32.64 (29.94)</td>
<td>32.99 (30.29)</td>
</tr>
<tr>
<td>weight (kDa)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental mass (kDa)</td>
<td>43.45</td>
<td>46.45</td>
</tr>
<tr>
<td>Number of theoretical</td>
<td>10</td>
<td>9 – 5 observed in structure</td>
</tr>
<tr>
<td>N-glycosylation sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of theoretical</td>
<td>14</td>
<td>12 – 0 observed in structure</td>
</tr>
<tr>
<td>O-glycosylation sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical pI</td>
<td>4.7 (4.5)</td>
<td>5.07 (4.88)</td>
</tr>
<tr>
<td>Experimental pI</td>
<td>4.4</td>
<td>4.5 – 4.8</td>
</tr>
<tr>
<td>First 10 N-terminal</td>
<td>HAAFWDKSMY</td>
<td>HIAFWHNSMY</td>
</tr>
<tr>
<td>residues after signal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>peptide cleavage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental N-terminal</td>
<td>HAAFW</td>
<td>HIAFW</td>
</tr>
<tr>
<td>N-terminal sequence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Table 2. \( \text{H}_2\text{O}_2 \) production in presence of various electron donors. Values are expressed as specific activities (U.g\(^{-1}\)) and relative specific activities (in \%) with reference to the ascorbate activity are shown in parenthesis.

<table>
<thead>
<tr>
<th>Electron Donor</th>
<th>( \text{PcAA14A} )</th>
<th>( \text{PcAA14B} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ascorbate</td>
<td>1782 (100)</td>
<td>1609 (100)</td>
</tr>
<tr>
<td>L-cysteine</td>
<td>1356 (76)</td>
<td>1336 (83)</td>
</tr>
<tr>
<td>gallic acid</td>
<td>81 (4)</td>
<td>84 (5)</td>
</tr>
<tr>
<td>epigallocatechin gallate</td>
<td>25 (1)</td>
<td>26 (2)</td>
</tr>
<tr>
<td>p-coumaric acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>caffeic acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cinapic acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vanillic acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>menadione</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>tannic acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>syringic acid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3-hydroxyanthranilic acid</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Supplementary Table 3. EPR spin Hamiltonian parameters from simulations of cw X-band and cw Q-band spectra for PcAA14A.

<table>
<thead>
<tr>
<th></th>
<th>X-band</th>
<th>Q-band</th>
</tr>
</thead>
<tbody>
<tr>
<td>g values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$g_x$</td>
<td>2.052</td>
<td>2.058</td>
</tr>
<tr>
<td>$g_y$</td>
<td>2.066</td>
<td>2.068</td>
</tr>
<tr>
<td>$g_z$</td>
<td>2.273</td>
<td>2.273</td>
</tr>
<tr>
<td>$A_x$</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>$A_y$</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>$A_z$</td>
<td>508</td>
<td>508</td>
</tr>
<tr>
<td>$A_{Cu}$ (MHz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{Cu}$ strains (MHz)</td>
<td>30, 70, 50</td>
<td>30, 70, 80</td>
</tr>
<tr>
<td>Line widths</td>
<td>1.5, 1.8</td>
<td>4.5, 4.5</td>
</tr>
</tbody>
</table>
Supplementary Table 4. Data collection and refinement statistics.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Native</th>
<th>Gadoteridol complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P4;2;2</td>
<td>P4;2;2</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a, b, c) (Å)</td>
<td>203.7, 203.7, 110.5</td>
<td>204.1, 204.1, 109.7</td>
</tr>
<tr>
<td>(\alpha, \beta, \gamma) (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>49.47-3.01</td>
<td>44.06-3.15</td>
</tr>
<tr>
<td>(R_{merge})</td>
<td>0.091 (1.039)</td>
<td>0.193 (1.379)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.998 (0.622)</td>
<td>0.995 (0.823)</td>
</tr>
<tr>
<td>(I / \sigma I)</td>
<td>15.9 (2.0)</td>
<td>11.4 (2.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.5 (99.4)</td>
<td>100.0 (100.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>8.3 (8.6)</td>
<td>14.4 (13.9)</td>
</tr>
<tr>
<td>Wilson B (Å²)</td>
<td>105.80</td>
<td>85.98</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>49.47-3.01</td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>43427</td>
<td></td>
</tr>
<tr>
<td>(R_{work} / R_{free}) (%)</td>
<td>23.0/24.0</td>
<td></td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>4360</td>
<td></td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>566</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(B)-factors (Å²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>94.76</td>
<td></td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>135.31</td>
<td></td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.107</td>
<td></td>
</tr>
</tbody>
</table>

Number of crystals used: one for the native and one for the MAD data sets.
Values in parentheses are for the highest-resolution shell.
Additional References

