



HAL
open science

Lytic xylan oxidases from wood-decay fungi unlock biomass degradation

Marie Couturier, Simon Ladeveze, Gerlind Sulzenbacher, Luisa Ciano, Mathieu Fanuel, Céline Moreau, Ana Villares, Bernard Cathala, Florence Chaspoul, Kristian Frandsen, et al.

► To cite this version:

Marie Couturier, Simon Ladeveze, Gerlind Sulzenbacher, Luisa Ciano, Mathieu Fanuel, et al.. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nature Chemical Biology*, 2018, 14 (3), pp.306-310. 10.1038/nchembio.2558 . hal-02188478

HAL Id: hal-02188478

<https://hal.science/hal-02188478>

Submitted on 19 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Lytic xylan oxidases from wood-decay fungi unlock biomass degradation**

2 Marie Couturier^{1,*}, Simon Ladevèze^{1,*}, Gerlind Sulzenbacher^{2,3}, Luisa Ciano⁴, Mathieu Fanuel⁵,
3 Céline Moreau⁵, Ana Villares⁵, Bernard Cathala⁵, Florence Chaspoul⁶, Kristian E. Frandsen¹,
4 Aurore Labourel¹, Isabelle Herpoël-Gimbert¹, Sacha Grisel¹, Mireille Haon¹, Nicolas Lenfant²,
5 Hélène Rogniaux⁵, David Ropartz⁵, Gideon Davies⁴, Marie-Noëlle Rosso¹, Paul H. Walton⁴,
6 Bernard Henrissat^{2,3,7}, Jean-Guy Berrin^{1,**}

7 ¹INRA, Aix Marseille Univ., Biodiversité et Biotechnologie Fongiques (BBF), UMR1163, F-
8 13009 Marseille, France

9 ²Architecture et Fonction des Macromolécules Biologiques (AFMB), CNRS, Aix-Marseille
10 Univ., F-13009 Marseille, France

11 ³INRA, USC1408 Architecture et Fonction des Macromolécules Biologiques (AFMB), F-13009
12 Marseille, France

13 ⁴Department of Chemistry, University of York, York, UK, YO10 5DD

14 ⁵INRA, Unité de Recherche Biopolymères Interactions Assemblages (BIA), F-44316 Nantes,
15 France

16 ⁶IMBE Aix Marseille Univ., IRD CNRS UAPV, Faculté de Pharmacie, F-13005 Marseille,
17 France.

18 ⁷Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

19 *These authors contributed equally to this work

20 **Correspondence and request to materials should be addressed to J.G.B.

21 (jean-guy.berrin@inra.fr)

22

23 **Abstract**

24 Wood biomass is the most abundant feedstock envisioned for the development of modern
25 biorefineries. However, the cost-effective conversion of this form of biomass to commodity
26 products is limited by its resistance to enzymatic degradation. Here we describe a new family of
27 fungal lytic polysaccharide monooxygenases (LPMOs) prevalent amongst white-rot and brown-
28 rot basidiomycetes, which is active on xylans - a recalcitrant polysaccharide abundant in wood
29 biomass. Two AA14 LPMO members from the white-rot fungus *Pycnoporus coccineus*
30 significantly increase the efficiency of wood saccharification through oxidative cleavage of
31 highly refractory xylan-coated cellulose fibers. The discovery of this unique enzyme activity
32 advances our knowledge on the degradation of woody biomass in nature and offers an innovative
33 solution to improve enzyme cocktails for biorefinery applications.

34

35

36 **Introduction**

37 Wood is the most abundant organic source of biomass on Earth, with an annual production of
38 about 5.64×10^{10} tons of carbon¹. Its widespread nature has allowed humans to use it in many
39 contexts, most notably as a building material due to its exceptional mechanical properties and
40 resistance to decay. In bio-based industries, the utilization of wood is taking on a new
41 importance as it constitutes the most promising source for advanced biofuels and plant-derived
42 products. Notwithstanding its potential, however, the cost-effective conversion of woody
43 feedstocks is limited by a single key factor, the recalcitrance of the lignocellulosic matrix to
44 degradation by enzyme cocktails². To overcome this recalcitrance, biorefineries utilize energy-
45 demanding pretreatment processes to solubilize the inaccessible biomass components before
46 enzymatic saccharification. The recalcitrant fraction reflects its heteroxylan content which is
47 known to be particularly resistant to xylanases due to extensive decoration and because these
48 xylans can adopt a flat conformation with their chains solidly adhering *via* hydrogen-bonds to the
49 surface of cellulose microfibrils^{3,4}. Finding sustainable means of overcoming this resistance to
50 degradation is one of the main challenges faced by modern biorefineries. Indeed, the xylan
51 problem is so severe that consideration is being given to engineering energy crops modified to
52 contain fewer recalcitrant xylans⁵.

53 In nature, fungi play a vital role in the terrestrial carbon cycle and dominate wood decomposition
54 in boreal forests⁶. Wood-decaying basidiomycetes classified as white-rot and brown-rot fungi,
55 naturally degrade cellulose and hemicelluloses using a large diversity of carbohydrate-active
56 enzymes (CAZymes; www.cazy.org)⁷ and Fenton-type chemistry⁸. In this context, understanding
57 of plant-cell wall deconstruction was recently overturned by the discovery of lytic
58 polysaccharide monooxygenases (LPMOs) enzymes which cleave polysaccharides through an

59 oxidative as opposed to hydrolytic mechanism⁹⁻¹¹. Such is their importance, that industrial
60 enzyme mixtures for the conversion of agricultural residues to biofuels now incorporate
61 cellulose-active LPMOs¹², helping biorefineries move towards environmental and economic
62 sustainability. Despite the significant efficiencies that LPMOs have brought to biomass
63 degradation, industrial enzyme cocktails are still unable to degrade woody biomass completely
64 and there is a major need to identify new enzymes capable of effecting this breakdown. From
65 this perspective, there are three fungal LPMO families (termed AA9, AA11 and AA13 in the
66 CAZy classification)⁷, which were discovered from genome sequences by virtue of their modular
67 structure where the catalytic LPMO domain is sometimes appended to known substrate-targeting
68 carbohydrate-binding modules (CBMs). Each fungal LPMO family is associated with the
69 oxidative cleavage of distinct polysaccharides with AA9 acting mainly on cellulose and
70 xyloglucan¹⁰, AA11 on chitin¹³ and AA13 on starch^{14,15}: a solely xylan-acting LPMO is
71 conspicuous by its absence.

72 Using comparative post-genomic approaches among fungal wood decayers, we identified the
73 existence of a previously unknown family of LPMO. This new family to be termed AA14 in the
74 CAZy classification differs phylogenetically and structurally from the previous AA9, AA10,
75 AA11 and AA13 families. The first characterized members from the white-rot basidiomycete
76 fungus *Pycnoporus coccineus* target xylan chains covering wood cellulose fibers thus unlocking
77 the enzymatic degradation of wood biomass.

78 **Results**

79 **Discovery of the AA14 family among fungal wood decayers**

80 The white-rot basidiomycete *Pycnoporus coccineus* is an efficient degrader of both hardwood
81 and softwood¹⁶. While studying the effect of different types of biomass on *P. coccineus* growth
82 using transcriptomics and secretomics, we identified a gene encoding a protein of unknown
83 function that was highly up-regulated on pine and poplar as compared to control¹⁶. The
84 corresponding protein (JGI ID 1372210; GenBank ID #KY769370) was secreted only during
85 growth on pine and poplar suggesting a role in wood decay. A BLAST search against public
86 sequence databases identified more than 300 proteins with significant similarity to #KY769370
87 from *P. coccineus*, many of which from well-known saprotrophic fungi. Sequence alignment
88 revealed a conserved N-terminal histidine (**Supplementary Fig. 1**), commensurate with a
89 copper-binding histidine brace active site¹⁰, which is a hallmark of known LPMOs. A
90 phylogenetic analysis shows that the newly identified sequences strongly cluster together with
91 high bootstrap values and are very distant from AA9, AA10, AA11 and AA13 sequences
92 (**Supplementary Fig. 2**), thereby defining a new LPMO family designated AA14 in the CAZy
93 database. AA14 members are found in all well-known white-rot (*Pleurotus ostreatus*,
94 *Phanerochaete chrysosporium*, *Trametes versicolor*) and brown-rot (*Serpula lacrymans*,
95 *Coniophora puteana*, *Postia placenta*) basidiomycetes and in some wood-inhabitants
96 ascomycetes within the *Xylariaceae* and *Hypocreaceae* families. A slight gene family expansion
97 is observed in wood-decaying basidiomycetes (average number per species 3.35 in
98 basidiomycetes and 1.28 in ascomycetes) (**Fig. 1; Supplementary Data Set 1**). None of the
99 AA14 members identified in fungal genomes harbors a carbohydrate-binding module (CBM)

100 explaining why this family was not previously discovered together with AA11 and AA13
101 through the “module walk” approach^{13,15}.

102 **Expression and biochemical characterization of *PcAA14***

103 Two *P. coccineus* proteins, *PcAA14A* (#KY769369) and *PcAA14B* (#KY769370), displaying
104 65% sequence identity were produced to high yield in *Pichia pastoris*, purified to homogeneity
105 and biochemically characterized (**Supplementary Table 1; Supplementary Fig. 3 and 4**). We
106 confirmed the correct processing of the native signal peptide, which exposed the N-terminal
107 histidine residue at position 1 in the mature polypeptide chain (**Supplementary Table 1**). Mass
108 spectrometry analyses revealed that both proteins contained ~ one copper atom per protein
109 molecule and treatment with EDTA led to partial apo forms (~0.1 copper atom per protein
110 molecule). *PcAA14A* and *PcAA14B* were both able to produce hydrogen peroxide in the
111 presence of ascorbate, cysteine or gallate as electron donors (**Supplementary Table 2**).

112 **Crystal structure of *PcAA14***

113 The structure of *PcAA14B* was solved by multiple-wavelength anomalous dispersion data
114 recorded at the gadolinium edge, and refined at 3.0 Å resolution. The core of the protein folds
115 into a largely antiparallel immunoglobulin-like β -sandwich (**Fig. 2a**), a fold globally similar to
116 that seen in LPMOs from other families. The active site of *PcAA14B* constituted by His1, His99
117 and Tyr176 forming the canonical histidine brace is exposed at the surface (**Fig. 2b**). In contrast
118 to the flat substrate-binding surfaces observed in AA9 LPMOs¹⁷, the *PcAA14B* surface has a
119 rippled shape with a clamp formed by two prominent surface loops (**Supplementary Fig. 5**).
120 Both loops are located in the N-terminal half of *PcAA14B*, and are equivalent to the L2 and L3
121 loop regions in AA9 LPMOs. Conventionally, the N-terminal part of AA9 LPMOs upstream of
122 the L2 loop region makes up a β -strand segment (single β -strand or a β -hairpin). No equivalent

123 β -strands are found in the *PcAA14B* structure, which, in contrast, forms loop segments
124 immediately after the N-terminal His (**Supplementary Fig. 5**). The *PcAA14B* structure also
125 reveals a cystine (Cys67-Cys90) in the L3-equivalent region, which borders an extension not
126 present in AA9 LPMOs (**Supplementary Fig. 5**). It is highly interesting to note that the two
127 loops making up the clamp in *PcAA14B* correspond to modified L2 and L3 loop regions, as
128 these have been shown to be involved in LPMO-substrate interactions¹⁷. For AA9 LPMOs a
129 conserved Tyr has been shown to be involved in substrate interactions at the active site surface¹⁷.
130 Interestingly, *PcAA14B* possesses equally a conserved tyrosine residue at the edge of the
131 substrate-binding surface, Tyr240, albeit located on a different loop region, which could
132 potentially make substrate interactions. Overall the crystal structure of *PcAA14B* reveals novel
133 features within its putative substrate binding site, which may suggest differences in terms of
134 substrate specificity compared to known LPMOs.

135 **EPR spectroscopic analysis of the copper site of *PcAA14***

136 Multi-frequency Electron Paramagnetic Resonance (EPR) analysis was carried out on both
137 *PcAA14A* and *PcAA14B* to determine the nature of the copper active site (Figure 2C;
138 **Supplementary Fig. 6**). The spin Hamiltonian parameters (**Supplementary Table 3**) displayed
139 axial parameters ($g_x \approx g_y < g_z$) with a $d(x^2-y^2)$ SOMO, placing the copper active site squarely
140 within a type 2 Peisach-Blumberg classification¹⁸. Simulations required the addition of two ($I=1$)
141 nitrogen atoms (coupling in the range of 30 to 36 MHz), as would be expected from the
142 coordinating histidine side chains. Overall, these spin-Hamiltonian parameters are similar to
143 those obtained for AA9 LPMOs confirming the presence of the copper(II) ion within the
144 histidine brace coordination environment¹⁹. These data support the hypothesis that *PcAA14s*
145 display LPMO characteristics and that copper is their native metal cofactor.

146 **Substrate specificity of *PcAA14***

147 Activity assays were initially carried out with *PcAA14A* and *PcAA14B* on a wide range of
148 polysaccharides including cellulose and xylans in the presence of ascorbic acid, which is widely
149 used as electron donor for LPMOs. Using standardized methods previously employed to
150 characterize AA9 LPMOs²⁰, no activity could be detected on these polysaccharides. Next, we
151 performed saccharification assays on pretreated biomass including poplar, pine and wheat straw
152 using a *Trichoderma reesei* CL847 cocktail mainly composed of cellulases and xylanases²¹. A
153 boost of glucose release from poplar and pine was observed upon addition of either of the AA14
154 enzymes to the cocktail (**Fig. 3a**). When the reactions were conducted in absence of a reductant
155 the boost effect was maintained (**Supplementary Fig. 8**), suggesting that one of the components
156 from the biomass (e.g. lignin) may act as an electron donor²². This improvement in glucose
157 release was dose-dependent yielding up to ~100% increase on pretreated softwood (**Fig. 3b**).
158 However, no significant boost was observed on wheat straw (**Supplementary Fig. 8**), which
159 differs in terms of hemicellulose composition compared to wood, indicating that AA14 enzymes
160 specifically target one of the components of woody biomass. In a finding with important
161 consequences for biorefinery use of woody biomass as feedstock, the *T. reesei* CL847 cocktail
162 enriched in AA9 LPMO acting on cellulose was also boosted by *PcAA14A*, suggesting that AA9
163 and AA14 enzymes may act on different regions within the lignocellulosic matrix
164 (**Supplementary Fig. 8**). Because AA14 members do not harbor any CBM module, we
165 artificially attached a fungal CBM1 module targeting crystalline cellulose to *PcAA14A*. The
166 resulting modular *PcAA14A*-CBM1 enzyme performed less efficiently than the catalytic module
167 alone (**Supplementary Fig. 8**), suggesting that AA14 enzymes may not require specific binding
168 to the flat crystalline cellulose surface.

169 To discern which polymer was attacked by AA14 enzymes, we used birchwood cellulosic fibers,
170 consisting of 79% cellulose and 21% xylan, as a substrate. After incubation with *PcAA14A* or
171 *PcAA14B*, wood fibers were disrupted (**Fig. 4a**) uncovering cellulose structures visualized at
172 different scales using transmission electron microscopy and atomic force microscopy
173 (**Supplementary Fig. 9**). These observations suggest a weakening of the cohesive forces that
174 link the wood fibers together in a manner similar to that previously described with AA9
175 enzymes²³. Samples treated with AA14 enzymes were further analyzed using solid-state Cross-
176 Polarization Magic Angle Spinning ¹³C Nuclear Magnetic Resonance (¹³C CP/MAS NMR). The
177 impact of AA14 enzymes on the fibers was different to that recently observed for AA9
178 LPMOs²³. In the case of *PcAA14* enzymes, no meaningful change was observed on cellulose
179 signals (**Fig. 4b; Supplementary Fig. 10**). Interestingly, however, significant changes in signal
180 areas corresponding to hemicelluloses located at 101 ppm and 82 ppm were observed when the
181 NMR spectra were deconvoluted in the C-1 and the C-4 regions (**Supplementary Fig. 10**).
182 These results suggest that AA14 enzymes act on xylans bound to cellulose, which have a rigidity
183 and a conformation similar to that of the underlying cellulose chains⁴. The specific attack of
184 *PcAA14* on xylan substrates differentiates this new class of enzymes from all other LPMOs^{24,25},
185 none of which has previously been reported to oxidize xylan in such a selective and efficient
186 manner.

187 To further substantiate the idea that AA14 enzymes act on xylan bound to cellulose, we
188 performed synergy assays of AA14 enzymes in combination with a fungal GH11 xylanase using
189 birchwood cellulosic fibers. Addition of *PcAA14A* to a GH11 xylanase significantly increased
190 by 40% the release of xylo-oligomers from birchwood cellulosic fibers (Figure 4C;
191 **Supplementary Fig. 11**). Additionally, no improvement of xylan conversion was observed on

192 birchwood cellulosic fibers when the xylanase was combined with a cellulose-acting AA9
193 LPMO (**Fig. 4c**).

194 We further investigated the nature of soluble products generated after synergistic action of
195 *PcAA14A* and the GH11 xylanase. Using ionic chromatography, a range of oligosaccharides
196 eluted at similar retention time to C1-oxidized oligosaccharides (**Supplementary Fig. 11**). Mass
197 spectrometry analyses performed on the same samples allowed the identification of several
198 putative oxidative species with masses corresponding to C1-oxidized xylotriose (X3ox) and C1-
199 oxidized xylo-tetraose (X4ox) and non-oxidized xylo-oligosaccharides substituted with
200 glucuronic acid (X₃MeGlcA, X₄MeGlcA, X₅MeGlcA) (**Supplementary Fig. 12**). The structure
201 of the C1-oxidized xylotriose with an aldonic acid on the reducing end (**Fig. 4d**) was confirmed
202 by fragmentation of the species observed at 429 *m/z* by tandem MS (MS/MS) (**Supplementary**
203 **Fig. 12**). The identification of oxidative products demonstrates that AA14 enzymes are LPMOs.

204

205 **Discussion**

206 Our findings that xylans are susceptible to AA14 oxidative cleavage only when adsorbed onto
207 crystalline cellulose and not when in solution are supported by reports that showed that xylans
208 exist in different contexts within the cell wall^{4,26}. Recalcitrant xylans bound to cellulose
209 microfibrils display a two-fold screw axis conformation aligned parallel to the cellulose chain
210 direction⁴ that is compatible with the proper orientation of the carbohydrate H1 and H4 atoms
211 with respect to the LPMO catalytic center²². Unravelling the substrate specificity of AA14s has
212 been challenging as these enzymes are not active on xylans in solution most probably due to the
213 three-fold helical screw conformation of the substrate²⁷. Using multidisciplinary approaches, we
214 reveal that AA14 LPMOs probably target specifically the protective shield made by heteroxylans

215 that cover cellulose microfibrils in wood. The conformation of xylan in this context contributes
216 to wood recalcitrance and glycoside hydrolases are not able to access such a sterically restricted
217 substrate. The cleavage of these rare motifs by AA14 LPMOs unlocks the accessibility of xylan
218 and cellulose chains to glycoside hydrolases therefore improving the overall saccharification of
219 woody biomass. These results not only greatly enhance our knowledge of wood superstructure,
220 they also contribute to understand and better exploit biomass deconstruction by fungal
221 saprotrophs.

222

223 **Acknowledgements.** We thank the European Synchrotron Radiation Facility (Grenoble), and the
224 synchrotron Soleil (Gif-sur-Yvette) for beam time allocation and assistance. We thank S. Tapin
225 for providing cellulose fibers, E. Bonnin and J. Vigouroux for compositional analyses, G. Toriz
226 and P. Gatenholm for providing purified wood xylan, L. Foucat and X. Falourd for their valued
227 assistance with treatments of the NMR data, E. Perrin for the excellent technical support for
228 TEM images, B. Seantier for the access and assistance to AFM facilities, D. Hartmann and E.
229 Bertrand for their help with enzyme production in bioreactor, and D. Navarro and G. Anasontzis
230 for insightful discussions. MC was funded by a Marie Curie International Outgoing Fellowship
231 within the 7th European Community Framework Program (328162). SL, MNR and JGB were
232 funded by the Microbio-E A*MIDEX project (ANR-11-IDEX-0001-02). AL was funded by a
233 Marie Curie Individual Fellowship within the Horizon 2020 Research and Innovation
234 Framework Programme (748758). This work was supported in part by the CNRS and the French
235 Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. NL and BH
236 were supported by Agence Française de l'Environnement et de la Maîtrise de l'Energie
237 (1201C102). PHW, GJD and LC thank the UK Biotechnology and Biological Sciences Research

238 Council (BB/L001926/1 and BB/L021633/1) for funding. GJD is the Royal Society Ken Murray
239 Research Professor.

240 **Author contributions.** MC and SL contributed equally to this work. MC identified the new
241 enzymes and performed biochemical characterization. MNR was in charge of transcriptomic and
242 proteomic analyses. MC, SL, SG, IG and MH performed production of proteins in flasks and
243 bioreactors. FC performed ICP-MS analysis. SL and SG performed synergy assays with xylanase
244 and protein crystallization. SL and GS solved the crystal structure. BH and NL performed
245 bioinformatic analyses. MC, SL, SG performed HPAEC analyses. MF, DR and HR identified
246 oxidized products using mass spectrometry. MC, SG and IG performed saccharification assays.
247 AV, CM and BC carried out microscopy and NMR analyses. LC performed the EPR study under
248 the direction of PHW and GJD. JGB supervised the work and organized the data. The manuscript
249 was written by JGB with contributions from BH and PHW. All authors made comments on the
250 manuscript and approved the final version. Figures were prepared by JGB, KEF, AL, NL, SL,
251 LC, MF, SG and IG.

252

253 **Competing Financial Interests Statement.** The authors declare no competing financial
254 interests.

255

256

257 **References**

- 258 1. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the
259 biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237-240 (1998)
- 260 2. Himmel, M. E. *et al.* Biomass recalcitrance: engineering plants and enzymes for biofuels production.
261 *Science* **315**, 804-807 (2007)
- 262 3. Biely, P., Singh, S. & Puchart, V. Towards enzymatic breakdown of complex plant xylan structures:
263 State of the art. *Biotechnol. Adv.* **34**, 1260-1274 (2016)
- 264 4. Simmons, T. J. *et al.* Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state
265 NMR. *Nat. Commun.* **7**, 13902 (2016)
- 266 5. Loqué, D., Scheller, H. V. & Pauly, M. Engineering of plant cell walls for enhanced biofuel
267 production. *Curr. Opin. Plant Biol.* **25**, 151-161 (2015)
- 268 6. Hibbett, D. S. & Donoghue, M. J. Analysis of character correlations among wood decay
269 mechanisms, mating systems, and substrate ranges in homobasidiomycetes. *Syst. Biol.* **50**, 215-242
270 (2001)
- 271 7. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-
272 active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-495 (2014)
- 273 8. Riley, R. *et al.* Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-
274 rot/brown-rot paradigm for wood decay fungi. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9923-9928 (2014)
- 275 9. Vaaje-Kolstad, G. *et al.* An oxidative enzyme boosting the enzymatic conversion of recalcitrant
276 polysaccharides. *Science* **330**, 219-222 (2010)
- 277 10. Quinlan, R. J. *et al.* Insights into the oxidative degradation of cellulose by a copper metalloenzyme
278 that exploits biomass components. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15079-15084 (2011)
- 279 11. Kracher, D. *et al.* Extracellular electron transfer systems fuel cellulose oxidative degradation.
280 *Science* **352**, 1098-1101 (2016)

- 281 12. Johansen, K. S. Discovery and industrial applications of lytic polysaccharide mono-oxygenases.
282 *Biochem. Soc. Trans.* **44**, 143-149 (2016)
- 283 13. Hemsworth, G. R., Henrissat, B., Davies, G. J. & Walton, P. H. Discovery and characterization of a
284 new family of lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **10**, 122-126 (2014)
- 285 14. Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R. & Marletta, M. A. A family of starch-active
286 polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13822-13827 (2014)
- 287 15. Lo Leggio, L. *et al.* Structure and boosting activity of a starch-degrading lytic polysaccharide
288 monooxygenase. *Nat. Commun.* **6**, 5961 (2015)
- 289 16. Couturier, M. *et al.* Enhanced degradation of softwood versus hardwood by the white-rot fungus
290 *Pycnoporus coccineus*. *Biotechnol. Biofuels* **8**, 216 (2015)
- 291 17. Frandsen, K. E. *et al.* The molecular basis of polysaccharide cleavage by lytic polysaccharide
292 monooxygenases. *Nat. Chem. Biol.* **12**, 298-303 (2016)
- 293 18. Peisach, J. & Blumberg, W. E. Structural implications derived from the analysis of electron
294 paramagnetic resonance spectra of natural and artificial copper proteins. *Arch. Biochem. Biophys.*
295 **165**, 691-708 (1974)
- 296 19. Garajova, S. *et al.* Single-domain flavoenzymes trigger lytic polysaccharide monooxygenases for
297 oxidative degradation of cellulose. *Sci. Rep.* **6**, 28276 (2016)
- 298 20. Bennati-Granier, C. *et al.* Substrate specificity and regioselectivity of fungal AA9 lytic
299 polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnol. Biofuels* **8**, 90 (2015)
- 300 21. Herpoël-Gimbert, I. *et al.* Comparative secretome analyses of two *Trichoderma reesei* RUT-C30 and
301 CL847 hypersecretory strains. *Biotechnol. Biofuels* **1**, 18 (2008)
- 302 22. Westereng, B. *et al.* Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer.
303 *Sci. Rep.* **5**, 18561 (2015)
- 304 23. Villares, A. *et al.* Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. *Sci.*
305 *Rep.* **7**, 40262 (2017)

- 306 24. Frommhagen, M. *et al.* Discovery of the combined oxidative cleavage of plant xylan and cellulose by
307 a new fungal polysaccharide monooxygenase. *Biotechnol. Biofuels* **8**, 101 (2015)
- 308 25. Fanuel, M. *et al.* The *Podospora anserina* lytic polysaccharide monooxygenase PaLPMO9H
309 catalyzes oxidative cleavage of diverse plant cell wall matrix glycans. *Biotechnol. Biofuels* **10**, 63
310 (2017)
- 311 26. McCartney, L. *et al.* Differential recognition of plant cell walls by microbial xylan-specific
312 carbohydrate-binding modules. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4765-4770 (2006)
- 313 27. Nieduszynski, I. & Marchessault, R. H. Structure of beta-D-(1-->4')xylan hydrate. *Nature* **232**, 46-47
314 (1971)
- 315

316 **Figure legends**

317 **Figure 1. Phylogeny of the AA14 family of LPMOs.** Phylogenetic tree of 283 fungi analyzed
318 for AA14 members. The number of AA14 in fungal species (listed in **Supplementary Data Set**
319 **1**) is represented by red bars (the scale is indicated at the bottom of the figure). When available,
320 the mode of wood decay (brown-rots or white-rots) is specified next to the leaves of the tree.
321 Pictures illustrate the taxonomical diversity of wood-decaying fungi displaying AA14 LPMOs;
322 Genus (order) from top to bottom: *Xylaria* (Xylariales), *Trichoderma* (Hypocreales), *Nectria*
323 (Hypocreales), *Tremella* (Hypocreales), *Pycnoporus* (Polyporales), *Ganoderma* (Polyporales),
324 *Serpula* (Boletales), *Gymnopilus* (Agaricales). Photo credit: C. Lechat (AscoFrance) and A.
325 Favel (CIRM-CF).

326

327 **Figure 2. Structure of AA14 LPMO and organization of the copper active site.** (a) Overall
328 three-dimensional structure of *PcAA14B* (ribbon depiction with active site residues shown as
329 sticks under transparent surface). (b) Active site (Histidine brace) overlay of Cu-*LsAA9A*
330 (magenta) and *PcAA14B* (gold). (c) Continuous wave X-band EPR spectrum (9.3 GHz, 165 K)
331 with simulation (red) of *PcAA14A*. More data are presented in **Supplementary Figures 5 and 6**.

332

333 **Figure 3. Contribution of *PcAA14* enzymes to the saccharification of biomass.** (a) Glucose
334 release upon saccharification of pretreated pine and poplar by the CL847 *Trichoderma reesei*
335 enzyme cocktail in the presence of *PcAA14A* or *PcAA14B* and ascorbic acid. Glucose was
336 quantified using ionic chromatography. More saccharification assays are presented in
337 **Supplementary Figure 8**. (b) Dose effect of the addition of *PcAA14A* on the saccharification of

338 pretreated pine. Concentration of *PcAA14A* was 0.1 μM (+), 0.5 μM (++) and 1 μM (+++).
339 Error bars indicate standard error of the mean from triplicate independent experiments. Data
340 points are shown as dots.

341
342 **Figure 4. Enzymatic activity of *PcAA14* LPMOs.** (a) Morphology of birchwood cellulosic
343 fibers treated with *PcAA14A* and *PcAA14B* LPMOs. Images were recorded after dispersion.
344 Images are representative of the samples analyzed. (b) Solid state ^{13}C CP/MAS NMR analysis of
345 LPMO-treated cellulosic fibers. The differences in hemicellulose content in enzyme-treated
346 fibers were calculated from the C-1 and C-4 region deconvolution of NMR spectra and are
347 indicated in **Supplementary Figure 10**. (c) Assays in the presence of a GH11 xylanase were
348 performed on birchwood cellulose fibers. Xylobiose (X2) and xylotriose (X3) were quantified by
349 ionic chromatography. Error bars indicate standard error of the mean from triplicate independent
350 experiments. (d) Mass spectrometry identification of the X3 oxidized species detected at
351 429.13 m/z generated from birchwood cellulosic fibers by *PcAA14A* in synergy with a GH11
352 xylanase. The fragmentation pattern corresponds to a C1 oxidized species with an aldonic acid at
353 the reducing end. ∇ : water losses. \square : H_2CO losses. An expanded view of the spectrum is
354 provided in **Supplementary Figure 12**.

355

356 **On-line Materials and Methods**

357 **Transcriptomics and secretomics of *Pycnoporus* sp.**

358 Transcriptomic and proteomic data of three-days-old cultures of *Pycnoporus coccineus* BRFM
359 310 and *Pycnoporus sanguineus* BRFM 1264 grown on cellulose (Avicel), wheat straw, pine and
360 aspen are described in^{16;28}.

361 **Bioinformatic analysis of AA14 LPMOs**

362 *P. coccineus* AA14 sequences (Genbank ID KY769369 and KY769370) were compared to the
363 NCBI non redundant sequence database using BlastP²⁹ in February 2016. Blast searches
364 conducted with AA14 did not retrieve AA9s, AA10s, AA11s or AA13s with significant scores,
365 and vice-versa. MUSCLE³⁰ was used to perform multiple alignments. To avoid interference from
366 the presence or absence of additional residues, the signal peptides and C-terminal extensions
367 were removed. Bioinformatic analyses were performed on 286 fungal genomes sequenced and
368 shared by JGI collaborators. Protein clusters are available thanks to the JGI
369 (<https://goo.gl/ZAA2NX>) for each of these fungi. A phylogenetic tree has been inferred using 100
370 cleaned and merged alignments of proteins from selected clusters of proteins. Those clusters are
371 present, as much as possible, in all fungi in 1 copy in order to maximize the score $\sum 1/n$ (with n ,
372 the number of copy in the genome). Sequences from clusters were aligned with Mafft³¹, trimmed
373 with Gblocks³² and a phylogenetic tree was built with concatenation of alignments with
374 Fasttree³³. The tree is displayed with Dendroscope³⁴ and Bio::phylo³⁵.

375 **Production of *P.coccineus* AA14 LPMOs**

376 The sequences corresponding to *PcAA14A* (Genbank ID KY769369) and *PcAA14B* (Genbank ID
377 KY769370) genes from *P. coccineus* BRFM310 were synthesized after codon optimization for

378 expression in *P. pastoris* (GenScript, Piscataway, USA). The region corresponding to the native
379 signal sequence was kept while the C-terminal extension region was removed. Synthesized genes
380 were further inserted with into a modified pPICZ α A vector (Invitrogen, Cergy-Pontoise, France)
381 using *Bst*BI and *Xba*I restriction sites in frame with the (His)₆-tag located at the C-terminus of
382 recombinant proteins. Fusion of *PcAA14A* with CBM1 was carried using the CBM1 domain of
383 *PaLPMO9E*, which was added to *PcAA14A* at the end of the catalytic module using the linker
384 sequence of *PaLPMO9E*²⁰. Constructs without (His)₆-tag sequence were also designed by adding
385 a stop codon at the end of the AA14 catalytic module. *P. pastoris* strain X33 and the pPICZ α A
386 vector are components of the *P. pastoris* Easy Select Expression System (Invitrogen), all media
387 and protocols are described in the manufacturer's manual (Invitrogen).

388 Transformation of competent *P. pastoris* X33 was performed by electroporation with PmeI-
389 linearized pPICZ α A recombinant plasmids and zeocin-resistant *P. pastoris* transformants were
390 screened for protein production as described in³⁶. The best-producing transformants were grown
391 in 2 L of BMGY medium containing 1 mL.L⁻¹ *Pichia* trace minerals 4 (PTM₄) salts in shaken
392 flasks at 30°C in an orbital shaker (200 rpm) to an OD₆₀₀ of 2 to 6. Cells were then transferred to
393 400 mL of BMMY medium containing 1 mL.L⁻¹ of PTM₄ salts at 20°C in an orbital shaker (200
394 rpm) for 3 days, with supplementation of 3% (v/v) methanol every day.

395 Bioreactor productions were carried out in 1.3-L New Brunswick BioFlo[®] 115 fermentors
396 (Eppendorf, Hamburg, Germany) following the *P. pastoris* fermentation process guidelines
397 (Invitrogen). Recombinant enzymes were secreted up to ~1 g.L⁻¹ (**Supplementary Figure 13**).

398 **Purification of *PcAA14* LPMOs**

399 The culture supernatants were recovered by pelleting the cells by centrifugation at 2,700 *g* for
400 5 min, 4°C and filtered on 0.45 µm filters (Millipore, Molsheim, France). For (His)₆-tagged
401 enzymes, the pH was adjusted to 7.8 and the supernatants were loaded onto 5 ml His Trap HP
402 columns (GE healthcare, Buc, France) connected to an Akta Xpress system (GE healthcare).
403 Prior to loading, the columns were equilibrated in 50 mM Tris HCl pH 7.8; 150 mM NaCl
404 (buffer A). The loaded columns were then washed with 5 column volumes (CV) of 10 mM
405 imidazole in buffer A, before the elution step with 5 CV of 150 mM imidazole in buffer A.
406 Fractions containing the protein were pooled and concentrated with a 3-kDa vivaspin
407 concentrator (Sartorius, Palaiseau, France) before loading onto a HiLoad 16/600 Superdex 75
408 Prep Grade column (GE Helthcare) and separated in 50 mM sodium acetate buffer pH 5.2. Gel
409 filtration analysis showed that both *PcAA14* proteins are monomeric in solution. For enzymes
410 without (His)₆-tag, salts contained in the culture media were diluted ten-fold in 20 mM Tris-HCl
411 pH 8, then culture supernatants were concentrated with a Pellicon-2 10-kDa cutoff cassette
412 (Millipore) to a volume of approx. 200 mL and loaded onto a 20-mL High Prep DEAE column
413 (GE Helthcare). Proteins were eluted using a linear gradient of 1 M NaCl (0 to 700 mM in 200
414 mL). Fractions were then analyzed by SDS PAGE and those containing the recombinant protein
415 were pooled and concentrated. The concentrated proteins were then incubated with one-fold
416 molar equivalent of CuSO₄ overnight before separation on a HiLoad 16/600 Superdex 75 Prep
417 Grade column in 50 mM sodium acetate buffer pH 5.2.

418 **Biochemical analysis of AA14 LPMOs**

419 Concentration of purified proteins was determined by using the Bradford assay (Bio-Rad,
420 Marnes-la-Coquette, France) or using a nanodrop ND-2000 device with calculated molecular
421 mass and molar extinction coefficients derived from the sequences. Proteins were loaded onto

422 10% SDS-PAGE gels (Thermo Fisher Scientific, IL, USA) which were stained with Coomassie
423 Blue. The molecular mass under denaturing conditions was determined with reference standard
424 proteins (Page Ruler Prestained Protein Ladder, Thermo Fisher Scientific). Native IEF was
425 carried out in the Bio-Rad gel system, using pI standards ranging from 4.45 to 8.2 (Bio-Rad).

426 **N-terminal amino acid sequence determination**

427 The N-terminal amino acid sequences of purified *PcAA14A* and *PcAA14B* were determined
428 according to the Edman degradation. Samples were electroblotted onto a polyvinylidene
429 difluoride membrane (iBlot, Life Technologies). Analyses were carried out on a Procise
430 Sequencing System (ThermoFisher).

431 **Matrix-assisted laser desorption ionization/mass spectrometry**

432 Matrix-assisted laser desorption ionization mass spectra analyses were performed on a Microflex
433 II mass spectrometer (Bruker Daltonics). One μL of matrix [10 mg of 2,5-dihydroxybenzoic acid
434 in 1 mL of $\text{CH}_3\text{CN}/\text{H}_2\text{O}$ 50/50 (v/v), 0.1% formic acid (v/v)] was added to 1 μL of intact
435 *PcAA14A* or *PcAA14B* protein sample (100 pmoles) in the same solution. Then, mixtures were
436 allowed to dry at room temperature. Data acquisition was operated using the Flex control
437 software. External mass calibration was carried out on Peptide calibration standard (Bruker
438 Daltonics).

439 **Deglycosylation assays**

440 To remove N-linked glycans, purified enzymes were treated with EndoHf (New England
441 Biolabs, Ipswich, MA) under denaturing conditions according to the manufacturer's instructions.
442 Briefly, 10 μg of protein were incubated in 0.5% SDS and 40 mM DTT and heated for 10 min at
443 100°C for complete denaturation. Denaturated samples were subsequently incubated with 1,500

444 units of EndoHf in 50 mM sodium acetate pH 6.0 for 1 h at 37°C. Deglycosylated and control
445 samples were analyzed by SDS-PAGE.

446 **Amplex Red assay**

447 A fluorimetric assay based on Amplex Red and horseradish peroxidase was used as described
448 previously³⁷. The reaction (total volume 100 μ L, 30°C, 30 min) was measured in 50 mM sodium
449 acetate buffer pH 6.0 containing 50 μ M Amplex Red (Sigma-Aldrich, Saint-Quentin Fallavier,
450 France), 7.1 U.mL⁻¹ horseradish peroxidase, 0.2 to 4 μ M enzyme, and 50 μ M reductant, i.e.
451 ascorbate, *p*-coumaric acid, caffeic acid, cinapic acid, vanillic acid, menadione, L-cysteine,
452 tannic acid, syringic acid, gallic acid, 3-hydroxyanthranilic acid (3-HAA) and epigallocatechin
453 gallate in water and fluorescence was detected using an excitation wavelength of 560 nm and an
454 emission wavelength of 595 nm using a Tecan Infinite M200 plate reader (Tecan, Männedorf,
455 Switzerland). The specific activity was counted from H₂O₂ calibration curve, and the slope
456 (13,227 counts. μ mol⁻¹) was used to convert the fluorimeters' readout (counts.min⁻¹) into enzyme
457 activity.

458 **ICP/MS Analysis**

459 To obtain apo enzymes, 100 mM EDTA treatment was performed overnight. Prior to the
460 analysis, samples were mineralized in a mixture containing 2/3 of nitric acid (Sigma-Aldrich,
461 65% Purissime) and 1/3 of hydrochloric acid (Fluka, 37%, Trace Select) at 120°C. The residues
462 were diluted in ultra-pure water (2 mL) before ICP/MS analysis. The ICP-MS instrument was an
463 ICAP Q (ThermoElectron, Les Ullis, France), equipped with a collision cell. The calibration
464 curve was obtained by dilution of a certified multi-element solution (Sigma-Aldrich). Copper

465 concentrations were determined using Plasmalab software (Thermo-Electron), at a mass of
466 interest $m/z=63$.

467 **Saccharification assays**

468 Wheat straw, pine and poplar biomass were pretreated under acidic conditions. Sugar
469 composition was determined using the alditol acetate method.³⁸ Wheat straw consisted of 51.98
470 ± 2.02 % (w/v) glucose, 5.70 ± 0.23 % (w/v) xylose and 0.46 ± 0.04 % (w/v) arabinose. Pine
471 consisted of 43.25 ± 1.34 % (w/v) glucose, 0.24 ± 0.01 % (w/v) xylose and 0.15 ± 0.02 % (w/v)
472 arabinose. Poplar consisted of 50.85 ± 0.91 % (w/v) glucose, 0.39 ± 0.01 % (w/v) xylose and
473 0.07 ± 0.01 % (w/v) arabinose. The enzymatic treatments were carried out in sodium acetate
474 buffer (50 mM, pH 5.2) in a final volume of 1 ml at 0.5% consistency (w d.m./v). The LPMO
475 treatment was carried out sequentially with a CL847 *T. reesei* enzyme cocktail²¹ provided by
476 IFPEN (Rueil-Malmaison, France). Each *PcAA14* enzyme was added to the substrate at a
477 concentration of between 0.1 and 1 μM in the presence or absence of 1 mM ascorbic acid for
478 72 h, followed by addition of 1 $\text{mg}\cdot\text{g}^{-1}$ dry matter (d. m.) substrate of commercial cellulases from
479 *T. reesei* for 24 h. Enzymatic treatments were performed in 2-ml tubes incubated at 45°C and
480 850 rpm in a rotary shaker (Infors AG, Switzerland). Then, samples were centrifuged at 14,000 *g*
481 for 5 min at 4°C and the soluble fraction was heated for 10 min at 100°C to stop the enzymatic
482 reaction. Glucose was quantified by high performance anion exchange chromatography coupled
483 with amperometric detection (HPAEC-PAD) as described in²⁰.

484 **Polysaccharides cleavage assays**

485 Avicel was purchased from Sigma-Aldrich and lichenan (from Icelandic moss), curdlan, starch,
486 barley β -1,3/1,4-glucan, konjac glucomannan, wheat arabinoxylan, tamarind xyloglucan were

487 purchased from Megazyme (Wicklow, Ireland). PASC was prepared from Avicel as described
488 previously²⁰ in 50 mM sodium acetate buffer pH 5.2. A similar protocol was used to prepare
489 swollen squid pen chitin provided by Dominique Gillet (Mahtani Chitosan, India).
490 Glucuronoxylans were extracted from birchwood as described previously³⁹.

491 All the cleavage assays contained between 0.5 and 1 μ M of *PcAA14s* in the presence of 1 mM
492 ascorbate and 0.1% (w/v) polysaccharides. The enzyme reactions were performed in 2-mL tubes
493 and incubated in a thermomixer (Eppendorf, Montesson, France) at 45°C and 850 rpm. After 16
494 h of incubation, samples were heated for 10 min at 100°C to stop the enzymatic reaction and then
495 centrifuged at 14,000 *g* for 15 min at 4°C to separate the soluble fraction from the remaining
496 insoluble fraction before determination of soluble products using HPAEC as described above
497 with oligosaccharides standards (Megazyme).

498 **Microscopy**

499 Aqueous dispersions of Kraft birchwood cellulosic fibers (kindly provided by Sandra Tapin,
500 FCBA, Grenoble, France) were adjusted to pH 5.2 with acetate buffer (50 mM) in a final
501 reaction volume of 5 mL. Each *PcAA14* enzyme was added to the fibers at a final concentration
502 of 20 mg.g⁻¹ in the presence of 1 mM of ascorbic acid. Enzymatic incubation was performed at
503 40 °C under mild agitation for 48 h. Samples were then dispersed by a Polytron PT 2100
504 homogenizer (Kinematica AG, Germany) for 3 min, and ultrasonicated by means of a QSonica
505 Q700 sonicator (20 kHz, QSonica LLC., Newtown, USA) at 350 W ultrasound power for 3 min
506 as described previously²³. The reference sample was submitted to the same treatment but it did
507 not contain the *PcAA14* enzyme. Birchwood cellulose fibers (reference and *PcAA14*-treated)
508 were deposited onto a glass slide and observed by a BX51 polarizing microscope (Olympus
509 France S.A.S.) with a 4× objective. Images were captured by a U-CMAD3 camera (Olympus

510 Japan). For the atomic force microscopy (AFM) experiments, samples were deposited onto mica
511 substrates from fiber solutions at 0.1 g L^{-1} , and allowed to dry overnight. Topographical images
512 on mica were registered by a Nanoscope III-A AFM (Bruker, Santa Barbara, US). The
513 images were collected in tapping mode under ambient air conditions (temperature and relative
514 humidity) using a monolithic silicon tip (RFESP, Bruker) with a spring constant of 3 N m^{-1} ,
515 and a nominal frequency of 75 kHz. Image processing was performed with the WSxM 5.0
516 software. For transmission electron microscopy (TEM) experiments, fiber solutions at 0.1 g L^{-1}
517 in water were deposited on freshly glow-discharged carbon-coated electron microscope grids
518 (200 mesh, Delta Microscopies, France) and the excess of water was removed by blotting. The
519 sample was then immediately negatively stained with uranyl acetate solution (2%, w/v) for 2 min
520 and dried after blotting. The grids were observed with a Jeol JEM 1230 TEM at 80 kV.

521 **NMR spectroscopy**

522 Solid state ^{13}C NMR experiments were performed on a Bruker Avance III 400 spectrometer
523 operating at a ^{13}C frequency of 100.62 MHz using a 4 mm double-resonance (H/X) magic angle
524 spinning (MAS) probe. Samples were dialyzed against ultrapure water (MWCO 12-14000) for 7
525 days to remove buffer, ascorbate and released soluble sugars. Experiments were conducted at
526 room temperature at a MAS frequency of 9 kHz using a cross-polarization sequence (CP/MAS).
527 The ^{13}C chemical shift was referenced using the carbonyl signal of glycine at 176.03 ppm. The
528 cross polarization pulse sequence parameters were: $3.2 \mu\text{s}$ proton 90° pulse, 2.50 ms contact time
529 at 67.5 kHz, and 10 s recycle time. Typically, the accumulation of 5,120 scans was used. All
530 spectra obtained were processed and analyzed using Bruker Topspin version 3.2. To determine
531 the crystallinity and the general cellulose's morphology of the C-1 and C-4 region of the
532 samples, we used the sophisticated approach⁴⁰ that is described in details in our previous work²³.

533 For the C1-region, this approach used three Lorentzian lines for the crystalline part (Cr (I α) and
534 Cr (I β)) and one Gaussian line for the less ordered cellulose (para-crystalline cellulose, PCr). For
535 the C-4 region, four lines for the crystalline part corresponding to crystalline and para-crystalline
536 (PCr) cellulose and three Gaussian lines for the amorphous part (accessible surfaces, AS, and
537 inaccessible surface, IAS) were used. The cellulosic fibers contained xylan, which was
538 considered in the spectral decomposition: in the C-1 region with one line at 101.4 ppm and in the
539 C-4 region with one broad line centered at 81.6 ppm.

540 **Synergy assays with xylanase**

541 Assays were run on the birchwood cellulose fibers used in microscopy and NMR experiments.
542 Fibers were grinded (< 0.18 mm particle size) and hydrated in water under stirring for 48 h prior
543 to enzymatic assays. One mL reaction volumes containing 0.5% (w/v) birchwood fibers were
544 incubated with 1 μ M of PcAA14s and 0.1 μ M of GH11 xylanase M4 (*Aspergillus niger*) from
545 Megazyme (reference E-XYAN4) in 10 mM sodium acetate pH 5.2 supplemented or not with
546 1 mM L-cysteine. Prior to the reaction, the GH11 xylanase was buffer exchanged with 10 mM
547 sodium acetate pH 5.2 using a PD-10 column (GE Helthcare) to remove any trace of ammonium
548 sulfate. Enzymatic reactions were performed in 2-mL tubes and incubated in a thermomixer
549 (Eppendorf, Montesson, France) at 45°C and 850 rpm for 24 h. Samples were then centrifuged at
550 14,000 g for 5 min at 4°C to separate the soluble fraction from the remaining insoluble fraction.
551 Proteins were removed from the soluble oligosaccharides fraction by filtering the supernatants
552 using Nanosep 3K Omega centrifugal devices (Pall corporation). Soluble oligosaccharides
553 generated were analyzed by HPAEC as described previously and mass spectrometry (see below)
554 using non-oxidized xylo-oligosaccharides (Megazyme) as standards. Corresponding C1-oxidized

555 standards (from DP2 to DP4) were produced from non-oxidized xylo-oligosaccharides by using
556 purified *Pa*CDHB prepared as described previously.²⁰ All assays were carried out in triplicate.

557 **Electrospray mass spectrometry (ESI-MS and MS/MS)**

558 Experiments were performed on a Synapt G2Si high-definition mass spectrometer (Waters
559 Corp., Manchester, UK) equipped with an Electrospray ion (ESI) source. Two types of mass
560 measurements were performed on the samples: firstly, a mass profile was done on a mass range
561 of 300-2000 m/z (M/S). Ions of interest were further isolated and fragmented by collision-
562 induced dissociation in the transfer cell of the instrument (MS/MS). In these experiments, ion
563 mobility (IM) was activated to reduce interference from sample impurities. IM was performed in
564 a travelling-wave ion mobility (TWIM) cell. The gas flows were held at 180 mL.min⁻¹ He in the
565 helium cell and at 90 mL.min⁻¹ N₂ in the mobility cell. The IM traveling wave height was set to
566 40 V and its wave velocity was set to 480 m.s⁻¹ for positive ionization mode and 500 m.s⁻¹ for
567 negative ionization mode. Samples were diluted 10-fold in MeOH/H₂O (1:1, v/v) and infused at
568 a flow rate of 5 μL.min⁻¹ in the instrument. The instrument was operated in positive or negative
569 polarity, and in “sensitivity” mode.

570 **Crystallization, data collection, structure determination and refinement**

571 All crystallization experiments were carried out at 20°C by the sitting-drop vapour-diffusion
572 method using 96-well crystallization plates (Swissci) and a Mosquito[®] Crystal (TTP labtech)
573 crystallization robot. Reservoirs consisted of 40 μL of commercial screens and crystallization
574 drops were prepared by mixing 100 nL reservoir solution with 100, 200 and 300 nL of protein
575 solution. An initial hit was obtained after 1 week from a condition of the AmSO₄ screen (Qiagen)
576 consisting of 2.4 M (NH₄)₂SO₄ and 0.1 M citric acid pH 4.0. This condition was further

577 optimized to obtain diffraction-grade crystals by mixing protein solution at 28 mg mL⁻¹ with
578 precipitant solution consisting of 2.4 M (NH₄)₂SO₄ and 0.1 M citric acid pH 4.4 at a volume ratio
579 of 3:1. *PcAA14B* crystals grew to dimensions of 0.15×0.15×0.05 mm in one week. Crystals
580 belong to space group P4₁2₁2 with cell axes 204×204×110 Å and two molecules *per* asymmetric
581 unit.

582 Crystals of *PcAA14B* were soaked for 5 min in a solution where 2.4 M (NH₄)₂SO₄ of the mother
583 liquor was replaced by 2.4 M Li₂SO₄ for the sake of cryoprotection prior to flash-cooling in
584 liquid nitrogen. As X-ray fluorescence scans on native crystals did not reveal a significant
585 presence of copper within the crystals, a heavy atom derivative was prepared by soaking the
586 crystals in reservoir solution supplemented with 55 mM of the gadolinium complex gadoteridol
587 prior to cryo-cooling. Native diffraction data were collected on beamline ID23-1, while a MAD
588 dataset at wavelengths of 1.711 and 1.698 Å for peak/inflection and remote energies, was
589 collected on beamline ID30B at the European Synchrotron Radiation Facility (ESRF), Grenoble,
590 France. Data were indexed and integrated in space group P4₁2₁2 using XDS⁴¹ and subsequent
591 processing steps were performed with the CCP4 software suite⁴². Determination of the Gd³⁺
592 substructure and subsequent phasing combined with solvent flattening were carried out with
593 SHELXC/D/E⁴², leading to a pseudo-free correlation coefficient of 71.8%. Starting from
594 experimental phases, an initial model comprising 526 residues (out of 584), was automatically
595 built with Buccaneer⁴³ and manually completed with Coot⁴⁴. This initial model was used for
596 rigid body refinement followed by restrained refinement against native data with the program
597 Refmac⁴⁵. A random set of 5% of reflections was set aside for cross-validation purposes. Model
598 quality was assessed with internal modules of Coot⁴⁴ and using the Molprobit server⁴⁶. Figures
599 representing structural renderings were generated with the PyMOL Molecular Graphics System

600 (DeLano, W.L. The PyMOL Molecular Graphics on <http://www.pymol.org/>). Atomic
601 coordinates and structure factors have been deposited within the Protein Data Bank
602 <http://www.rcsb.org>⁴⁷. Data collection and refinement statistics are summarized in

603 **Supplementary Table 4.**

604 **EPR**

605 Continuous wave (cw) X-band frozen solution EPR spectra of a 0.2 to 0.3 mM solution of
606 Cu(II)-*PcAA14A* and *PcAA14B*, prepared and copper loaded as described above, in 10% v/v
607 glycerol at pH 5.2 (50 mM sodium acetate buffer) and 165 K were acquired on a Bruker EMX
608 spectrometer operating at ~9.30 GHz, with modulation amplitude of 4 G, modulation frequency
609 100 kHz and microwave power of 10.02 mW (4 scans). Both enzymes showed identical EPR
610 spectra. Cw Q-band frozen solution spectra of 1.0 mM solution of Cu(II)- *PcAA14A* at pH 5.2
611 (50 mM sodium acetate buffer) and 113 K were acquired on a Jeol JES-X320 spectrometer
612 operating at ~34.7 GHz, with modulation width 1 mT and microwave power of 1.0 mW (8
613 scans).

614 Spectral simulations were carried out using EasySpin 5.0.3⁴⁸. Simulation parameters are given in
615 **Supplementary Table 3.** g_z and $|A_z|$ values were determined accurately from the absorptions at
616 low field. It was assumed that g and A tensors were axially coincident. Accurate determination
617 of the g_x , g_y , $|A_x|$ and $|A_y|$ was obtained by simultaneous fitting of both X and Q band spectra.
618 The superhyperfine coupling values for the nitrogen atoms could not be determined accurately,
619 although it was noted that satisfactory simulation could only be achieved with the addition of
620 two nitrogen atoms with coupling in the range 30-36 MHz.

621 **Statistics**

622 For all statistics, $n = 3$ values were used to calculate the standard error of the mean. Values
623 resulted from independent experiments. For all representative results, experiments were repeated
624 at least two times and at least 20 images were collected for microscopy analyses.

625
626 **Accession codes**

627 *PcAA14A* and *PcAA14B* sequences were deposited in GenBank under accession numbers
628 KY769369 and KY769370, respectively. The X-ray structure of *PcAA14B* was deposited in the
629 Protein Data Bank with accession 5NO7. Raw EPR data are available on request through the
630 Research Data York (DOI: 10.15124/8758d712-1e67-467e-b0f0-f0dd99f0232a).

631 **Data Availability Statement**

632 All data generated or analysed during this study are included in this published article (and its
633 supplementary information files).

634
635 **Methods-only references**

- 636 28. Miyauchi, S. *et al.* Visual Comparative Omics of Fungi for Plant Biomass Deconstruction.
637 *Front. Microbiol.* **7**, 1335 (2016)
- 638 29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment
639 search tool. *J. Mol. Biol.* **215**, 403-410 (1990)
- 640 30. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
641 throughput. *Nucleic Acids Res.* **32**, 1792-1797 (2004)
- 642 31. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
643 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780. (2013)
- 644 32. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and
645 ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564-577
646 (2007)
- 647 33. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood
648 trees for large alignments. *PLoS One* **5**, e9490 (2010)

- 649 34. Huson, D. H. & Scornavacca, C. Dendroscope 3: an interactive tool for rooted phylogenetic
650 trees and networks. *Syst. Biol.* **61**, 1061-1067 (2012)
- 651 35. Vos, R. A., Caravas, J., Hartmann, K., Jensen, M.A. & Miller, C. BIO:Phylo-
652 phyloinformatic analysis using perl. *BMC Bioinformatics* **12**, 632011 (2011)
- 653 36. Haon, M. *et al.* Recombinant protein production facility for fungal biomass-degrading
654 enzymes using the yeast *Pichia pastoris*. *Front. Microbiol.* **6**, 1002 (2015)
- 655 37. Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. & Ludwig, R. Production of four
656 *Neurospora crassa* lytic polysaccharide monoxygenases in *Pichia pastoris* monitored by a
657 fluorimetric assay. *Biotechnol. Biofuels* **5**, 79 (2012)
- 658 38. Englyst, H.N. & Cummings, J.H. Improved method for measurement of dietary fiber as non-
659 starch polysaccharides in plant foods. *J. Assoc. Off. Anal. Chem.* **71**, 808-814 (1988)
- 660 39. Westbye, P., Svanberg, C. & Gatenholm, P. The effect of molecular composition of xylan
661 extracted from birch on its assembly onto bleached softwood kraft pulp. *Holzforschung* **60**,
662 143-148 (2006)
- 663 40. Larsson, P. T., Wickholm, K. & Iversen, T. A CP/MAS ¹³C NMR investigation of molecular
664 ordering in celluloses. *Carbohydrate Research* **302**, 19-25 (1997)
- 665 41. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125-132 (2010)
- 666 42. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr.*
667 *D Biol. Crystallogr.* **67**, 235-242 (2011)
- 668 43. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein
669 chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002-1011 (2006)
- 670 44. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot.
671 *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501 (2010)
- 672 45. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures
673 by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240-255
674 (1997)
- 675 46. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular
676 crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21 (2010)
- 677 47. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank.
678 *Nat. Struct. Biol.* **10**, 980 (2003)
- 679 48. Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral
680 simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42-55 (2006)

681