

This is a repository copy of *On the roles of AA15 lytic polysaccharide monooxygenases derived from the termite Coptotermes gestroi*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/185535/>

Version: Accepted Version

---

**Article:**

Walton, Paul Howard orcid.org/0000-0002-1152-1480, Lourenco Franco Cairo, Joao Paulo orcid.org/0000-0002-4744-3111, Cannella, David et al. (11 more authors) (2021) On the roles of AA15 lytic polysaccharide monooxygenases derived from the termite *Coptotermes gestroi*. JOURNAL OF INORGANIC BIOCHEMISTRY. 111316. ISSN 0162-0134

<https://doi.org/10.1016/j.jinorgbio.2020.111316>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

1 ~~**Journal:** Proceedings of the Royal Society B, Biochemistry, International Journal of Biological~~  
2 ~~Macromolecules, Biochemical Journal, New Biotechnology, JBC, FEBS.~~

3  
4 **Titles:**

5  
6 The discovery of AA15 lytic polysaccharide monooxygenases with specific chitinase activity from  
7 the lower termite *Coptotermes gestroi*

8  
9 **Authors:** João Paulo L. Franco Cairo<sup>1,2,3</sup>, Alessandro Paradisi<sup>2</sup>, David Cannella<sup>4</sup>, Thiago A.  
10 Gonçalves<sup>1,3</sup>, Marcelo V. Rúbio<sup>1</sup>, César R. F. Terrasan<sup>1</sup>, Robson Tramontina<sup>3</sup>, Luciana S Mofatto<sup>5</sup>,  
11 Marcelo F. Carrazolle<sup>5</sup>, Wanius Garcia<sup>6</sup>, Gideon J. Davies<sup>2</sup>, Claus Felby<sup>7</sup>, André R. L. Damásio<sup>1</sup>,  
12 Paul H. Walton<sup>2\*</sup>, Fabio Squina<sup>3\*</sup>.

13  
14 **Affiliations:**

15  
16 <sup>1</sup>Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas,  
17 Campinas, SP, Brazil.

18 <sup>2</sup>Department of Chemistry, University of York, Heslington, York, United Kingdom.

19 <sup>3</sup>Programa de Processos Tecnológicos e Ambientais, Universidade de Sorocaba - UNISO,  
20 Sorocaba, SP, Brazil.

21 <sup>4</sup>Interfaculty School of Bioengineers, Université Libre de Bruxelles, Brussels, Belgium.

22 <sup>5</sup>Department of Genetic, Evolution and Bioagents, Institute of Biology, University of Campinas,  
23 Campinas, SP, Brazil.

24 <sup>6</sup>Centro de Ciências Naturais e Humanas, Universidade Federal do ABC (UFABC), Santo André,  
25 SP, Brasil.

26 <sup>7</sup>Department of Geosciences and Natural Resource Management, Faculty of Science, University  
27 of Copenhagen, Frederiksberg C, Denmark.

28  
29 **Keywords:** LPMO, Termite, AA15, *Coptotermes*, CAZy, chitinases.

30  
31  
32

33 **Abstract (200 words)**

34

35 Lytic polysaccharide monoxygenases (LPMOs) are copper-dependent enzymes involved in the  
36 oxidative cleavage of polysaccharides. LPMOs belonging to the family AA15 are found widespread  
37 across the Tree of Life, including viruses, algae, oomycetes and animals. Recently, two AA15s  
38 from the insect *Thermobia domestica* were reported to have oxidative activity, the first against  
39 cellulose and chitin and the latter only on chitin. Herein, we report the identification and  
40 characterization of two novel family AA15 members from the termite *Coptotermes gestroi*.  
41 Structural models indicate that the LPMOs from *C. gestroi* (*CgAA15a/CgAA15b*) and *T. domestica*  
42 (*TdAA15a*) have a similar fold but with differences in the catalytic site residues, as well as the  
43 electrostatic potential of the binding surfaces. According to biochemical and biophysical studies,  
44 the copper ion had a thermostabilizing effect on both *C. gestroi* AA15 structures and ascorbate is  
45 an electron donor for H<sub>2</sub>O<sub>2</sub> production. The *CgAA15a* and *CgAA15b* were able to only oxidize chitin  
46 from squid pen, with no activity on cellulose, xylan, xyloglucan and starch. To mimic the catalytic  
47 pocket of the cellulose and chitin-active *TdAA15a*, the Cu proximal axial and non-coordinating  
48 phenylalanine was substituted with tyrosine in the active site of *CgAA15a*, but the mutation did not  
49 change the enzymatic specificity. Our data indicated that these enzymes are not involved in  
50 lignocellulose digestion and might play a role in developmental processes. Furthermore, this study  
51 illuminates the role of the Cu proximal axial aromatic residue on catalytic specificity of AA15 family  
52 members.

53

54

55

56

57

58

59

60

61

62

63

64

65

66

## 1. Introduction.

Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent enzymes, which oxidize recalcitrant polysaccharides such as chitin (EC. 1.14.99.53), cellulose (EC. 1.14.99.54; EC. 1.14.99.56) and hemicelluloses, diverging from the classical hydrolytic mechanism applied by glycoside hydrolases (GHs) (1). The discovery of LPMOs a decade ago (2) has stimulated the use of the recalcitrant plant cell wall (PCW) as a sugar source for the production of second-generation biofuels (3) and changed the classical view of cellulose degradation performed by the action of cellulases, exocellulases and  $\beta$ -glucosidases (4). Nowadays, LPMOs are included in the carbohydrate-active enzyme (CAZy) database and are classified as Auxiliary Activity (AA) class (5). Currently, there are seven LPMOs families reported (AA9, AA10, AA11, AA13, AA14, AA15 and AA16) according to their amino acid sequence, found mainly in lignocellulose-degrading microorganisms such as bacteria and fungi, which exhibit different substrate specificities and regioselectivity (6).

The oxidative mechanism that underlies the LPMOs activity is based on the copper active site, commonly referred to as the histidine-brace (His-brace) (7), in which two His residues, one being the N-terminus of the protein, coordinate a copper ion ( $\text{Cu}^{2+}$ ) in a T-shape arrangement (8). One of the axial positions of the Cu coordination sphere is occupied by an aromatic residue (tyrosine or phenylalanine), which plays a poorly understood role (9, 10). In some families, such as AA10, AA11 and AA15, an Alanine residue is also found highly conserved on the opposite side of the aromatic residue (1), however exceptions with Glycine and Isoleucine have been recently reported (11, 12). Using external electron donors and molecular oxygen as co-substrates, LPMOs can perform the hydroxylation of C1 and/or C4 carbons from sugar molecules, cleaving the glycosidic bonds of the polysaccharide (13). This overall enzymatic mechanism is shared among all LPMO families; however, the occurrence of C6 hydroxylation has been recently described for some AA9s, suggesting a different mode of operation of LPMOs (14, 15).

Among the LPMO families, the recently reported AA15s are widespread in the Tree of Life, including viruses, algae and oomycetes (16). Furthermore, it is in the Animal Kingdom, mainly in the Euarthropoda phylum, that the AA15s have a vast abundance of orthologous sequences, suggesting that these enzymes could play a pivotal role not only in catabolism but also in development (16). The first AA15s reported were identified in the cellulolytic firebrat *Thermobia domestica*. This insect has 23 LPMO coding sequences (CDS) in its genome, of which two (*TdAA15a* and *TdAA15b*) were biochemically and structurally characterized (16). *TdAA15a* was one of the most abundant LPMOs from *T. domestica*'s digestome and can oxidize both cellulose

104 and chitin substrates. *TdAA15b* was only able to cleave chitin, suggesting a role in the  
105 development and cuticle/exoskeleton remodelling of this insect (16). Both *TdAA15a* and  
106 *TdAA15b* share the conserved immunoglobulin-like protein fold as well as the His-brace, which is  
107 ubiquitous for all LPMOs. However, *TdAA15a* and *TdAA15b* differ by the nature of the aromatic  
108 residue found in the axial position of the Cu ion, being a Tyr or a Phe residue for the two  
109 enzymes, respectively (16). Because *T. domestica* is a primitive cellulolytic insect, we decided to  
110 investigate whether termites, which are among the major lignocellulolytic organisms on Earth,  
111 have LPMOs in their genomes and what role these enzymes might play.

112  
113 Termites are detritivore social insects that play a central role in the carbon and nitrogen  
114 cycles in forests and savannas worldwide (17, 18). For example, termites can decompose around  
115 58-64% (19) of deadwood in tropical rainforest, playing a crucial role in mitigating the effects of  
116 drought in the forest ecosystem (20). These insects digest lignocellulosic biomass with up to 90%  
117 efficiency (21), and this is achieved thanks to the termite's dual digestive system (22), which  
118 combines endogenous and symbiotic CAZymes that acts independently inside the termite gut to  
119 breakdown the PCW (23, 24). The termite endogenous cellulolytic system consists of few  
120 cellulases (GH9) and  $\beta$ -glucosidases (GH1), secreted in the fore and midguts, which deconstruct  
121 the woody material previously chopped by the mandibles (25, 26). CDSs for cellobiohydrolases  
122 (GH7 or GH6) are not found in the genome of termites (27–29), being present only in symbiotic  
123 flagellates from the hindgut of lower termites (30).

124  
125 In this context, the lack of endogenous GH7, alongside with reports from literature,  
126 suggest that host (25, 26) and symbiotic cellulases (30, 31) have low activity on crystalline  
127 celluloses, mainly when compared with fungal and microbial enzymes (32, 33), thus hinting to a  
128 possible “missing piece” in the termite cellulose breakdown mechanism. Furthermore, even when  
129 considering the arsenal of GHs from the host and symbiotic systems, it is difficult to rationalise  
130 the high efficiency of lignocellulose digestion by termites. Recently, reports from our group (34,  
131 35) and others (36, 37) reported that a collection of redox enzymes (related to Pro-oxidant,  
132 Antioxidant and Detoxification processes - PADs) may play an important role in the termite  
133 digestome, helping to overcome the lignocellulose recalcitrance and therefore explain the  
134 termite's high efficiency of PCW degradation (38).

135  
136 Based on this information, the study of termite digestion biology has become of industrial  
137 interest since its digestome serves as a reservoir of lignocellulolytic enzymes useful for plant  
138 biomass-to-bioproduct applications (39). Additionally, because the termite is considered one of  
139 the major urban pests on Earth, causing damage to buildings and houses around the world with  
140 an estimated economic losses of \$ 20 billion annually (40), the genes encoding for CAZy and

141 PAD enzymes have recently become targets for pest control management via interference RNA  
142 techniques (41, 42).

143

144 In the present study we report, for the first time, the identification of LPMOs belonging to  
145 the family AA15 from the genome of the lower termite *C. gestroi*, the major and ruthless urban  
146 pest in Brazil (43–45), followed by the functional and structural characterization for two of these  
147 enzymes. We also investigate the role of the Cu proximal axial aromatic residue in the catalytic  
148 site of the AA15s and its implications on *C. gestroi* LPMO functions.

149

150

## 2. Results

### 2.1. The genome of *C. gestroi* encodes four LPMOs belonging to the family AA15 with orthologous sequences in Termitoidea.

The analysis of the genome of *C. gestroi* identified four CDSs containing the Pfam domain PF03067, suggesting that functional LPMOs occur in this insect. This specific Pfam hosts entries for different LPMOs families from CAZy database such as AA10, AA13, AA15 and AA16. Two of the CDSs (*CgLPMOa* and *CgLPMOb*) could be mapped as complete open reading frames (ORFs), while the other sequences (*CgLPMOc* and *CgLPMOd*) were identified as partial ORFs.

A phylogenetic analysis was performed to classify the LPMOs from the *C. gestroi* genome into the AA families. First, a Blastp search using the *CgLPMOs* as a query was executed to identify orthologous sequences in other termite genomes, retrieving a total of thirteen CDSs from the termites *Zootermopsis nevadensis*, *Cryptotermes secundus*, *Macrotermes natalensis* and *Cortaritermes* sp.. Around 300 LPMO sequences from the Animal Kingdom and Euarthropoda phylum were retrieved from the EggNOG database, these sets of full-length amino acid sequences also included the *TdAA15a* and *TdAA15b* (16). These analyses permitted for classifying the *C. gestroi* LPMOs into the AA15 family. Thus, the *C. gestroi* full-length CDSs were named *CgAA15a* and *CgAA15b*, whereas the other CDSs are named *CgAA15c\_partial* and *CgAA15d\_partial*.

The analysis of signal peptide prediction was performed for all the sequences containing the Pfam domain 03067, confirming that the first residue after the cleavage site was a His (data not shown). Thus, a sequence alignment was performed using Muscle software with the termite LPMOs amino acid sequences and the AA15s from *T. domestica* (Fig. S1). Alignment showed that the His-brace motif is conserved in all termite LPMOs, as well as the Phe and Ala residues in the Cu proximal and distal axial positions, respectively. Moreover, the  $\beta$ -tongue-like protrusion region identified in the AA15s from *T. domestica*, which extends the surface surrounding the active site, is also conserved in *CgAA15a* and *CgAA15b* (Fig. S1) as well as for the other termite LPMOs. However, the protrusion is not conserved in all AA15 members as previously reported (16).

Finally, a maximum likelihood analysis was performed using the orthologous amino acid sequences from termites and also from Euarthropoda. LPMO sequences from other families were used as outgroups. The LPMOs from *C. gestroi*, *Z. nevadensis*, *C. secundus*, *M. natalensis* and *Cortaritermes* sp. were clustered together along with the sequences of AA15s from *T.*

188 *domestica*, forming a distinct group which was separated from the other LPMOs families with  
189 high bootstrap values (0.94), confirming that the *C. gestroi* enzymes belong to the AA15 family  
190 (Fig. 1A). These termite LPMOs were grouped in three separated subclades of the AA15 branch  
191 (Fig. 1B). Each subclade shared different protein architectures, as previously reported for AA15s  
192 (16). Subclade 1 has LPMOs bearing transmembrane and cytosolic domains downstream of the  
193 LPMO domain, subclade 2 displays sequences of unknown function downstream of the LMPO  
194 domain and subclade 3 has sequences containing only the LMPO domain (Fig. 1B).

## 197 **2.2. The *C. gestroi* AA15s are expressed in worker and soldier castes of *C. gestroi*.**

199 Transcripts encoding for *C. gestroi* AA15s were identified in the transcriptomic database of  
200 *C. gestroi* castes (based on 454-pyrosequences) (34), indicating their expression in both workers  
201 and soldiers. *CGAA15a* was found expressed in both termite castes as well as *CGAA15d\_partial*,  
202 while *CGAA15b* was expressed only in soldiers. *CGAA15c\_partial* had its transcripts identified  
203 only in the worker caste (Fig. S2A). Moreover, recent genomic and RNA-Seq analyses using  
204 Illumina technology were able to identify transcripts of *CGAA15a* as well as *CGAA15b* in the  
205 worker caste of *C. gestroi* (in preparation), confirming the results from the pyrosequencing  
206 analysis. The full-length CDSs for *CGAA15a* and *CGAA15b* have 633 and 1020 bp, respectively,  
207 including the signal peptide sequences for both enzymes and the transmembrane and cytosolic  
208 domains for *CGAA15b*. Based on these results, oligonucleotides were designed and only the  
209 catalytic domains of *CGAA15a* and *CGAA15b* from *C. gestroi* worker caste were amplified. After  
210 RNA extraction from the whole *C. gestroi* bodies and its transcription to cDNA, the sequences  
211 were amplified without the sequence encoding for the signal peptide, producing two fragments  
212 with 573 bp for *CGAA15A* and 591 bp for *CgAA15B\_cat*, which are in accord with our predictions  
213 (Fig. S2B).

## 216 **2.3. Comparative and structural insights of *CgAA15a* and *CgAA15b\_cat* reveals conserved hallmarks of AA15 family.**

219 To obtain structural information regarding *CgAA15a* and *CgAA15b\_cat*, both enzymes  
220 were modelled by employing the I-TASSER server and using the *TdAA15a* structure (PDB code  
221 5MSZ) as a template. *TdAA15a* has 47% and 44% sequence identity with *CgAA15a* and  
222 *CgAA15b\_cat*, respectively. Thus, the three-dimensional homology models of *CgAA15a* and  
223 *CgAA15b\_cat* were built with a high C-Scores (-5 to 2) of 0.77 and 0.91 with estimated RMSDs of  
224  $3.8\pm 2.6\text{\AA}$  and  $3.6\pm 2.5\text{\AA}$ , respectively.



225

226 The AA15 homology models of *C. gestroi* showed a typical central and antiparallel  $\beta$ -  
227 sandwich fold of LPMOs, with several loops that are stabilized by five disulphide bonds (Fig. 2A I-  
228 III), in agreement with the structure of *TdAA15a* (PDB code 5MSZ) (Fig. 2A III). *CgAA15a* and  
229 *CgAA15b\_cat* active sites were predicted with the presence of a copper ion coordinated by the  
230 His-brace motif (7). For both enzymes, the conserved residues His1 and His91 are predicted to  
231 coordinate the copper ion with a T-shaped geometry (Fig. 2B I-III) as observed for all  
232 characterized LPMOs (8). The *CgAA15a* and *CgAA15b\_cat* were predicted to have a Phe  
233 (Phe182 and Phe191, respectively) (Fig. 2B I and II), diverging from *TdAA15a* which presents a  
234 Tyr (Fig. 2B III) occupying the axial site of the Cu on the His Brace coordination plane. The apical  
235 position of the Cu ion is occupied by a non-coordinating Ala residue (Ala89), which is common to  
236 all three structures (Fig. 2A I-III). Moreover, the catalytic site of the AA15s from *C. gestroi*  
237 presents many similarities to the majority of chitin-active LPMOs from the AA10 family (46) and,  
238 looking in detail to the AA15 LPMOs alignment (Fig. S1), all termite LPMOs share the same  
239 active site.

240

241 The electrostatic potential molecular surface was generated for both *C. gestroi* AA15s  
242 (Fig. 2C I-II) and compared with that of *TdAA15a* (Fig. 2C III). Comparison shows that the overall  
243 surface charge for *CgAA15a*, *CgAA15b\_cat* and *TdAA15a* are divergent, mainly regarding the  
244 negative potential (red) around the His-brace for *CgAA15a* and *CgAA15b\_cat* (Fig 2C I-II – see  
245 the ellipses). However, similar potentials are found in the protrusion region (neutral) and around  
246 the His-brace (slightly positive potential) (blue). The putative docking site for a protein partner  
247 predicted for *TdAA15a* (PDB code 5MSZ) (16) is absent in both *C. gestroi* AA15s.

248

249 Aromatic and polar charged residues surrounding the active site are solvent-exposed  
250 residues on the flat catalytic surface and were previously suggested to be involved in substrate  
251 binding for other LPMO families (10, 46, 47). The AA15s from *C. gestroi* also have these solvent-  
252 exposed residues at the surface (Fig. S3A I-II) and surrounding the His-brace as observed for  
253 *TdAA15a* (Fig. S3A III). Concerning aromatic residues, the protrusion region (not found in all  
254 AA15s) has the aromatic residues Trp164 (for *CgAA15a*) and Trp173 (for *CgAA15b\_cat*)  
255 conserved at the same positions. On the opposite side of the protrusion, the residues Phe24 and  
256 Tyr24, for *CgAA15a* and *CgAA15b\_cat*, respectively, are also exposed on the surface and these  
257 aromatics are present in all termite AA15s and highly positional conserved among AA15  
258 members (70%) (16).

259

260 Other aromatic residues are also present on the flat catalytic surface of the three  
261 structures analyzed, highlighting the Phe34 for all of them. Moreover, *CgAA15a* and

262 *CgAA15b\_cat* have Phe30 and Tyr30 exposed near the His-brace, respectively, as well as a  
263 Trp38 a little further from the catalytic pocket (Fig. S3 A I-II). The other aromatic residues  
264 exposed in *CgAA15a* are Tyr39 and Tyr166, contrasting with the positions of others aromatic  
265 residues in *CgAA15b\_cat* and *TdAA15a*. The alignment of *C. gestroi* LPMOs with the LPMOs  
266 from *Z. nevadensis* and *C. secundus* also showed that the Tyr24 is conserved for all of them,  
267 with an exception for *CgAA15a*, whereas the Trp residue around the positions 164-173  
268 (protrusion) is fully conserved among all termite sequences (Fig. S1).

270 Polar residues exposed on the flat catalytic surface of LPMOs are also reported to bind the  
271 substrates, and among them the negatively charged Glu (around the positions 180 – 189) is  
272 exposed around the His-brace and conserved for *CgAA15a*, *CgAA15b\_cat* and *TdAA15a*  
273 structures (Fig. S3B I-III), as well as for the other termite AA15s (Fig. S1). Moreover, the residues  
274 Glu25 and Glu28 are conserved in both *C. gestroi* LPMOs (Fig. S3B I-II) and the alignment  
275 analysis suggests that these residues are also conserved in all other termites CDSs as well as in  
276 the specific chitin-active *TdAA15b* (Fig. S1) from *T. domestica*. However, these residues are  
277 absent on the surface of the cellulose/chitin-active *TdAA15a*. In the order hand, the Asp residues  
278 exposed on the flat surface at positions Asp27 and Asp35 are only found in the *TdAA15a*  
279 structure (Fig. S3B III).

281 Together the phylogenetic and structural analyses described in this work indicate that an  
282 intriguing difference between the catalytic sites among the *CgAA15a* and *CgAA15b* with  
283 *TdAA15a* is the non-coordinating and Cu proximal axial aromatic residue. *TdAA15a*, with Tyr in  
284 the axial position, was reported to oxidize cellulose and chitin substrates. On the other hand,  
285 *TdAA15b*, which has a Phe residue in the same position, can oxidize only chitin. A similar pattern  
286 is also observed for AA10s, where enzymes with Tyr in the same position can oxidize at least  
287 cellulose and the majority of AA10 enzymes with Phe can oxidize only chitin (1). Therefore, it  
288 seems that the aromatic axial residue may play a role in substrate specificity in AA15s and also in  
289 other LPMO families. Thus, to investigate the role of the non-coordinating and Cu proximal axial  
290 aromatic residue in *C. gestroi* AA15s, we functionally characterized *CgAA15a* and *CgAA15b\_cat*  
291 as well as a mutated version of *CgAA15a*, where the Phe182 was changed to a Tyr, herein  
292 named *CgAA15a*<sup>F182Y</sup>.

#### 295 **2.4. Functional characterization of *CgAA15a*, *CgAA15b\_cat* and *CgAA15a*<sup>F182Y</sup>.**

297 The CDS regions coding for the LPMO domain in *CgAA15a* and *CgAA15b\_cat* were  
298 cloned in pET-26b(+) without the signal peptide and with a C-terminal strep-tag. Next, the

299 recombinant proteins were successfully produced in the periplasmic space of *Escherichia coli*  
300 (Fig. S4A). Protein identity was confirmed by LC-MS/MS analysis of the protein bands isolated  
301 from the SDS polyacrylamide denaturing gels (data not shown). The production of soluble  
302 recombinant proteins could only be achieved using M9 minimal medium without copper ion  
303 supplementation, followed by periplasmic extraction using osmotic shock and affinity  
304 chromatography (Fig. S4B and C). The mutant *CgAA15a*<sup>F182Y</sup> was generated using site direct  
305 mutagenesis, and protein production and purification were performed following the same  
306 methodology described above (Fig. S4D). The proteins were copper loaded to maintain their  
307 stability, and the unbound copper ions were washed out using size-exclusion chromatography.  
308

309 After purification the LPMO activities were first evaluated using the 2,6-dimethoxyphenol  
310 (2,6-DMP) assay at pH 6.0 and 30 °C (Fig. 3A) as described previously (48). The specific  
311 activities were  $6.19 \pm 0.71$  U/g and  $7.55 \pm 0.36$  U/g for *CgAA15a* and *CgAA15b\_cat*, respectively.  
312 The specific activity for the mutant *CgAA15a*<sup>F182Y</sup> was lower than the native enzyme, resulting in  
313 a value of  $4.75 \pm 0.39$  U/g. To determine the melting temperature ( $T_m$ ) for *C. gestroi* AA15  
314 LPMOs, a thermofluor analysis using SYPRO Orange dye was carried out. *CgAA15a* showed a  
315  $T_m$  of 52.7 °C and after treatment with 5 mM EDTA for copper ion removal, the  $T_m$  value  
316 decreased to 49.3 °C ( $\Delta T_m$  of 3.4 °C). *CgAA15b\_cat* showed a  $T_m$  of 59.2 °C, and similarly to  
317 *CgAA15a*, the  $T_m$  value decreased to 53.0 °C ( $\Delta T_m$  of 6.2 °C) after treatment with EDTA. For  
318 *CgAA15a*<sup>F182Y</sup>, the  $T_m$  determined was 60.3 °C, and after Cu removal the  $T_m$  value decreased to  
319 46 °C ( $\Delta T_m$  of 14.3 °C). The mutant *CgAA15a*<sup>F182Y</sup> showed a  $\Delta T_m$  of 7.6 °C compared to the wild  
320 type *CgAA15a*, suggesting that the mutation increases the thermal stability of the enzyme.  
321 However, the  $T_m$  after stripping the copper ion from the protein was similar between both  
322 enzymes. These results indicate that the copper ion has a thermostabilizing effect on the protein  
323 structures as reported previously for other LPMOs from different families (11, 49).  
324

325 Previous studies have shown that LPMOs in the presence of molecular oxygen and an  
326 external electron donor (without the substrate) can generate hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) (50).  
327 Taking advantage of this feature, the termite LPMOs were tested with different electron donors  
328 using the Amplex<sup>®</sup>Red assay for H<sub>2</sub>O<sub>2</sub> quantification at pH 7.0 and 30 °C. The highest H<sub>2</sub>O<sub>2</sub>  
329 generation was achieved for *CgAA15a*, *CgAA15b\_cat* and *CgAA15a*<sup>F182Y</sup> when ascorbic acid  
330 was used as the donor, generating more than 2 μM of H<sub>2</sub>O<sub>2</sub> after 30 minutes of reaction for each  
331 enzyme (Fig. 4A-C). Interestingly, the mutant *CgAA15a*<sup>F182Y</sup> reached the same levels of H<sub>2</sub>O<sub>2</sub>  
332 production as *CgAA15a*, however at a lower velocity. The second highest production of H<sub>2</sub>O<sub>2</sub> was  
333 achieved with L-cysteine, but after a few minutes of the assay it appeared that resofurin, the  
334 Amplex<sup>®</sup>Red reaction product, reacted with the electron donor, exhibiting a distorted curve.

335 Generation of H<sub>2</sub>O<sub>2</sub> in the presence of gallic acid and pyrogallol, two phenolic compounds, was  
336 detected, but in low amounts for all three LPMOs tested.

### 338 **2.5 CgAA15a, CgAA15b\_cat and CgAA15a<sup>F162Y</sup> can oxidize and degrade chitin only.**

339  
340 To investigate the substrate specificities of the *C. gestroi* LPMOs, activity assays were  
341 carried out at pH 6 and 30 °C, with a range of different polysaccharides as substrates and  
342 ascorbic acid as the external electron donor. After the reactions, the supernatants were analysed  
343 by MALDI-TOF Mass Spectrometry and the peak masses from the oligosaccharides were  
344 compared to data previously reported (11, 16).

345  
346 *CgAA15a* and *CgAA15b\_cat* were only able to oxidize chitin from Squid Pen in the  
347 presence of an external electron donor, with a C1-oxidation pattern and generation of C1-almonic  
348 acids. Native peaks relative to the mono-sodiated unoxidized oligosaccharides as well as  
349 oxidized peaks corresponding to mono- or di-sodiated adducts of C1-almonic acids, imparting +16  
350 or +38 m/z respectively, were identified in the spectra. Moreover, peaks for the mono-sodiated  
351 lactone (-2) were also identified. The peak profiles in the mass spectra consisted of even-  
352 numbered oxidized oligosaccharides with different degrees of polymerization (DP), with the  
353 peaks corresponding to DP6 and DP8 at higher intensities than the odd-numbered DP5 and DP7.  
354 This pattern is in accordance with previous chitin-active LPMOs that exhibited the ability to  
355 breakdown crystalline structures (2), suggesting that *CgAA15a* and *CgAA15b\_cat* are correlated  
356 with chitin remodelling or exoskeleton moulting in *C. gestroi*.

357  
358 *CgAA15a<sup>F182Y</sup>* was also tested to oxidize a range of different polysaccharides; however, in  
359 its native form the mutant was only able to oxidize chitin from squid pen. The mutant showed the  
360 same peak masses as *CgAA15a* and *CgAA15b\_cat*, presenting the C1-oxidation pattern and the  
361 even-numbered oxidized oligosaccharides. Finally, the negative control assays (substrate and  
362 ascorbic acid) did not generate any oxidized products (Fig. S5 A). Moreover, the three termite  
363 LPMOs were unable to produce oxidized peaks in the presence of 5 mM EDTA (Fig. S5 B-D),  
364 showing that the copper ion is essential for the oxidative reaction.

### 367 **3. Discussion.**

368 In this study we report for the first time the identification and characterization of two  
369 LPMOs belonging to the family AA15 from a lower termite. The two AA15s from *C. gestroi* were  
370 expressed in worker and soldier castes and they were able to oxidize only chitin from squid pen,  
371 suggesting that both enzymes are related to chitin remodelling or exoskeleton moulting, and are

372 not involved in lignocelluloses degradation. The presence of a Tyr instead of a Phe is the major  
373 divergence among the catalytic sites of *TdAA15a* (cellulose/chitin-active) and *TdAA15b* (chitin-  
374 active) from the firebrat *T. domestica* (16), as well as for all termite AA15 LPMOs. According to  
375 our data, the mutant *CgAA15a*<sup>F182Y</sup> was unable to oxidize cellulose (PASC and Avicel).  
376

377 The role of Tyr in the catalytic site of AA9s, which was described recently, is related to  
378 protecting the protein from inactivation during uncoupled turnover (9) and it does not seem to be  
379 involved in substrate specificity. Among the AA10 family, LPMOs with Tyr at the same axial  
380 position are at least cellulose-active enzymes, with some members showing cellulose/chitin  
381 oxidation (1, 46, 51); moreover to the best of our knowledge no specific chitin-active AA10 are  
382 reported with a Tyr in the active site. On the other hand, most chitin-active AA10 enzymes  
383 contain Phe in the axial position (1, 52), with some enzyme exemplars exhibiting mixed activities  
384 towards cellulose/chitin or specific activity for cellulose (51). Moreover, change of the proximal  
385 axial and non-coordinating Phe to Tyr in the catalytic site of AA10s dramatically decreased  
386 enzyme activity or eliminated it, which corroborates with our data (10).  
387

388 For example, the change of Y213F in the axial position of the specific cellulose-active  
389 *TfAA10a* from *Thermobifida fusca* only decreased its oxidative activity against cellulose and  
390 increased substrate binding (53), however the potential for chitin oxidation by *TfAA10a*<sup>Y213F</sup> was  
391 not tested. On the other hand, the change of F219Y in *ScLPMO10c* from *Streptomyces coelicolor*  
392 abolished the oxidative activity on cellulose while the change to F219A only reduced it (51),  
393 although the authors also did not test the mutants for chitin oxidation. Despite the efforts to  
394 understand the role of aromatic residues in the axial position, it is found that native AA10s with  
395 Tyr in axial position, as well as fungal AA9s, are naturally cellulose-active enzymes at the least,  
396 while fungal AA11s, with Tyr in the active-site, are specific chitin-active LPMOs.  
397

398 Protein engineering applied to specific cellulose-active *ScLPMO10C* (with Phe in axial  
399 position) has recently shown that a broader set of aromatic and polar residues, which are located  
400 on the flat catalytic surface of the enzyme, can modulate substrate binding and specificity (47).  
401 According to the study (47), the authors performed the substitution of aromatic residues from the  
402 catalytic surface (previously reported to interact with cellulose (54)) with polar residues, which  
403 gave *ScLPMO10C* the ability to oxidize chitin. Moreover, Zhou and co-authors presented an  
404 interesting analysis in the electrostatic potential binding surfaces of several AA10 enzymes,  
405 indicating that chitin-active AA10s have negatively charged binding surfaces whereas cellulose-  
406 active AA10s have uncharged or positively charged surfaces (52). This finding corroborates our  
407 data, where binding surfaces of *CgAA15a*, *CgAA15b* and *TdAA15b* (Figure S6) are negatively  
408 charged although the cellulose/chitin-active *TdAA15a* has an uncharged binding surface. Thus,

409 the role of the axial and non-coordinating aromatic residues in the catalytic pocket of LPMOs  
410 allied with the study of binding residues on the catalytic surface of the enzymes deserves to be  
411 continually investigated.

412  
413 Collectively, our biochemical data and structural insights indicate that *CgAA15a* and  
414 *CgAA15b* are not involved in lignocellulose digestion in *C. gestroi* and they might play other roles  
415 in termite biology. The recent transcriptome analysis in the gut epithelium of the higher termite  
416 *Cortaritermes sp.* also reported no differential expression of the gene encoding for LPMO  
417 *CspAA15a*, after changing the termite diet from wild (non-lignocellulosic) to lignocellulosic-based  
418 (raw *Miscanthus sp.*) (55). The same expression pattern was also reported in genes encoding for  
419 chitodextrinases and chitinases, suggesting that chitin utilisation by the host did not change  
420 significantly upon *Miscanthus* feeding. Interestingly, when considering the termite AA15s  
421 alignment (Figure S1), the Phe residue also occupies the proximal axial position in the catalytic  
422 site of *CspAA15a*. Another point worth mentioning is that these AA15s may also have a role in  
423 the digestion of chitin derived from termite corpses since cannibalism is a recurrent phenomenon  
424 among these insects, including *C. gestroi* (56, 57).

425  
426 Moreover, Sabbadin et al., (2018) (16) reported the presence of two AA15s, *DmAA15a*  
427 (*CG42749*) and *DmAA15b* (*CG4362*), in the tracheal system and peritrophic matrix (midgut),  
428 respectively, of the 3rd instar larvae of *Drosophila melanogaster*. Both organs are rich in chitin  
429 and chitin-binding proteins and these tissues are in constant chitin remodelling and deposition  
430 during moulting (larval stage) and metamorphosis (pupa). Considering the amino acid sequences  
431 of *DmAA15a* and *DmAA15b*, as well as the other three AA15 sequences from *D. melanogaster*  
432 (*CG4367*, *CG42598* and *CG41284*), it is interesting to highlight that all the enzymes exhibit Phe  
433 as the proximal residue from the catalytic site.

434  
435 In conclusion, the present study provides functional and structural characteristics of  
436 *CgAA15a* and *CgAA15b\_cat*, which can support future studies on the role of these enzymes in  
437 the developmental and digestive biology of termites. Because the specific chitin-active LPMOs  
438 *CgAA15a* and *CgAA15b\_cat* from termites and *TdAA15b* from firebrat share the same catalytic  
439 site, as well as similarities in the electrostatic potential binding surfaces of the enzymes, we can  
440 hypothesize that AA15s from the termites *Z. nevadensis*, *C. secundus*, *M. natalensis* and  
441 *Cortaritermes sp.* are also specific chitin-active, and most probably they are not related to plant  
442 cell wall digestion. However, only the full functional characterization of these enzymes can  
443 confirm this hypothesis. Finally, the role of AA15s in termites deserves investigation regarding  
444 their potential biotechnological applications (e.g. pest control) since the deletion or silencing of

445 AA15s in *Drosophila melanogaster*, as well as *Tribolium castaneum*, is lethal or partially lethal  
446 (16).

447

448

## 449 **4. Material and Methods**

### 450 **4.1 Genome and transcriptome analyses.**

451

452 The genome and transcriptome from worker and soldier castes of the lower termite *C.*  
453 *gestroi* were used as a database (34) to search for CDSs containing the Pfam domain 03067  
454 (*LPMO\_10 Superfamily*) using default parameters (e-value 1). Afterwards, each CDS had their  
455 number of transcripts in worker and soldier libraries counted and normalized (57).

456

### 457 **4.2 Phylogenetic analyses.**

458

459 For the identification of orthologous sequences in termites, the CDSs containing the Pfam  
460 domain 03067 were used as queries to execute a BLASTP (protein-protein BLAST) analysis  
461 using default parameters against Termitoidae (taxid:1912919). After retrieving only termite  
462 sequences, the CDSs were used as a query to search for orthologous sequences in the EggNOG  
463 database (58). Next, they were submitted to signal peptide sequence analysis in the SignalP 5.0  
464 (59) platform and sequences without a His residue after the cleavage site were discarded.

465 Maximum likelihood analyses were performed to classify the termite CDSs into an LPMO  
466 family. Thus, CDSs from other LPMOs families were retrieved from the CAZy database (60) and  
467 used as external groups. The analysis was executed at [www.phylogeny.fr](http://www.phylogeny.fr) in the “*a la carte*” mode  
468 (61). The multiple alignments were performed using ProbCons (62) with default parameters and  
469 without curation, followed by construction of the phylogenetic tree using PhyML (63) also with  
470 default metrics, except for not removing the gaps from alignment. Statistical tests for branch  
471 support were performed using the Approximate Likelihood-Ratio Test (aLRT) in the SH-like mode  
472 (64). The tree was visualized using the iTOL web-platform (65).

473

### 474 **4.3 In silico molecular modelling.**

475

476 The three-dimensional homology models of *CgAA15a*, *CgAA15b\_cat* and *TdAA15b*  
477 (without the signal peptide and Strep-tag sequences) were built using the I-TASSER (66)  
478 software with the LPMO *TdAA15a* from *Thermobia domestica* (PDB id: 5MSZ\_A) as a reference  
479 structure (template). The C-scores were calculated for all enzyme models. The PDB  
480 visualizations were performed using the PyMOL™ software (1.7.4.5 Edu) and the electrostatic  
481 surface potential was calculated using the APBS plug-in (67) for PyMOL.

482

#### 483 **4.4 RNA extraction, cDNA transcription and PCR reactions.**

484

485 The total RNA extraction from *C. gestroi* as well as the cDNA transcription was performed  
486 as described previously by Franco Cairo et. al., 2013 (26). PCR reactions for amplification of the  
487 full-length CDS of *CGAA15A* and the CDS for the LPMO domain of *CGAA15B* were performed in  
488 two rounds using the following oligonucleotides: 1<sup>st</sup> round PCRs - *CgAA15a\_F\_cDNA* 5'-  
489 CATGCCAG CTTATGGACCCTGTG3-'; *CgAA15a\_R\_cDNA* 5'-  
490 GACGAGAGAAATGTCTGAGCAGCC-3'; *CgAA15b-cat\_F\_cDNA*  
491 5'-CATGGACGTCTTATTGAACCTCCCTCC3'; *CgAA15b-cat\_R\_cDNA* 5'-  
492 AACATCAGCACATGCTCGGAATTC-3'. For the 2<sup>nd</sup> round PCRs - *CgAA15a\_F\_pET26b\_S* 5'-  
493 **GCCCAGCCGGCGATGGCG**CATGCCAGCTTATG-3'; *CgAA15a\_R\_pET26b\_S* 5'-CGGGTG  
494 GCTCCACGCGCTGACGAGAGAAATGTC-3'; *CgAA15b-cat\_F\_pET26b\_S* 5'-  
495 **GCCCAGCCGGCGATGGCG**CATGGACGTCTTATTG; *CgAA15b-cat\_R\_pET26b\_S* 5'-  
496 CGGGTGGCTCCACGCGCTAACATCAGCACATGC-3', where the bold letters are  
497 complementary to the pelB signal peptide from the pET-26b(+) vector and the underlined letters  
498 are complementary to the Strep-tag II sequence from previously modified pET-26b(+) (16). For  
499 linearization of the pET-26b-Strep-tag II vector, the following oligonucleotides were used:  
500 PF\_Lin\_26b\_pelB 5'-AGCGCGTGGAGCCACCCGCAG-3 and PR\_Lin\_26b\_Strep 5'-  
501 CGCCATCGCCGGCTGGGC-3'. For site-directed mutagenesis in *CgAA15a-F182Y*, the following  
502 primers were used: *CgAA15a\_mut1\_F* 5'-GCAAGAAACGtatCGTGGCTGTAGCGACATCTC-  
503 3'and *CgAA15a\_mut1\_R* 5'-GGCCCACACCCTACGGCA-3', with the mutation as lower-case  
504 letters in the former primer. The PCR reactions for amplification of target CDSs were set up as  
505 reported by Franco Cairo et al., (26).

506

#### 507 **4.5 Cloning, recombinant protein production and purification.**

508

509 For recombinant protein production and purification, the *C. gestroi* LPMOs were cloned in  
510 the pET-26b(+) vector, seeking periplasmic expression. Each CDS was inserted between the  
511 pelB leader signal peptide and the *Bam*HI restriction site, with a strep-tag inserted in the LPMO  
512 C-terminal followed by a stop codon, thus removing the 6xHis-tag. Cloning of *CgAA15a* and  
513 *CgAA15b\_cat* was performed using the Gibson Assembly Master Mix according to manufacture  
514 instructions. The mutant *CgAA15a-F182Y* was generated using the Q5 Mutagenesis kit from  
515 NEB according to manufacture instructions with one modification: the use of Phusion *Taq* DNA  
516 polymerase instead of Q5 DNA polymerase. Afterwards, the two constructions, as well as the  
517 mutated version of *CgAA15a-F182Y*, were transformed in DH5- $\alpha$  for vector replication. Finally,



518 the plasmids were extracted, sequenced and then transformed in BL21(DE3) Rosetta 2 pLys S  
519 for recombinant protein production.

520 Cells containing the targeted constructions were pre-inoculated in LB (Lysogenic Broth)  
521 overnight at 37 °C and 200 rpm. Next, 20 mL of culture was inoculated in 1 L of 1X M9 minimal  
522 medium with glucose as the sole carbon source as reported in Fowler et al., 2019, and the cells  
523 were grown at 37 °C and 210 rpm until the optical density reached 0.7 (8 to 10 h). The cells were  
524 then cooled to 18 °C for 1 h and the protein expression was later induced by adding 1 mM IPTG  
525 to the final concentration. Recombinant protein production was performed for 16 h at 18 °C and  
526 180 rpm.

527 After protein expression, the LPMOs were extracted from the periplasmic space using the  
528 Osmotic Shock protocol (68), followed by affinity chromatography using Streptacitin resin. Next,  
529 the enzymes were concentrated and copper ion (CuCl<sub>2</sub>) was loaded in excess of five-times,  
530 followed by size-exclusion chromatography in 25 mM MES buffer adjusted to pH 6 and containing  
531 150 mM NaCl (69).

532

#### 533 **4.6 Functional characterizations of *C. gestroi* LPMOs.**

534

535 The LPMO activity assays using 2,6-dimethoxyphenol (2,6-DMP) and hydrogen peroxide  
536 (H<sub>2</sub>O<sub>2</sub>) as co-substrates were performed as reported by Breslmayr et al. (2018) (48). The  
537 reactions were set up in total volumes of 200 µL, containing 100 µM H<sub>2</sub>O<sub>2</sub>, 10 mM 2,6-DMP and 1  
538 µM of *C. gestroi* LPMOs in 100 mM ammonium acetate buffer pH 6 at 30 °C. The reaction was  
539 initiated by adding the LPMO after 15 min of pre-incubation of the reagents and absorbance was  
540 monitored at 469 nm every 15 s during 5 min in the Epoch 2 Microplate Reader (BioTek). The  
541 LPMOs had their melting temperature ( $T_m$ ) determined according to the protein thermal shift  
542 assay (ThermoFisher Scientific) (70). The LPMOs at 0.5 mg/mL (20 mM MES pH 6, 150 mM  
543 NaCl) were mixed with 1X SYPRO Orange in 30 µL reaction volumes and the fluorescence  
544 emission was monitored using a CFX384 Touch Real-Time PCR machine (BioRad) in a 384-well  
545 clear plate from 20 °C to 99 °C, in triplicate, where the proteins were incubated at each  
546 temperature for 2 min before taking the reading.

547 The Amplex®Red assay for H<sub>2</sub>O<sub>2</sub> quantification was performed to determine molecules  
548 with potential to donate electrons for *C. gestroi* LPMOs. The reactions were set up as reported by  
549 Kittl et al. (2012) (50) in a final volume of 100 µL in clear microplate and in triplicate. The assays  
550 contained 50 µM of each electron donor (ascorbic acid, pyrogallol, gallic acid and L-cysteine), 50  
551 µM Amplex®Red, 7 U/mL Horseradish peroxidase (HRP) and 1 µM LPMO as final  
552 concentrations. The assays started by adding the LPMOs and the resofurin absorbance was  
553 measured at 590 nm. The reactions were monitored in kinetic mode in the Epoch 2 Microplate

554 Reader (BioTek) during 35 min at 30 °C with 3 s shaking before the readings. A standard curve  
555 was constructed using H<sub>2</sub>O<sub>2</sub> concentrations in the range of 0.1 – 5 µM for its quantification.

#### 557 **4.7 Activity assays for substrate specificity determinations.**

558  
559 The activity assays aiming to determine the substrate specificities of *C. gestroi* LPMOs  
560 were performed on a range of different polysaccharides as follows: microcrystalline cellulose  
561 (Avicel), Phosphoric Acid Swollen Cellulose (PASC), squid pen chitin, xyloglucan (from tamarind,  
562 Megazyme), corn starch (Sigma) and xylan from oat spelt (Sigma). The reactions were carried  
563 out for 16 h, in 20 mM ammonium acetate (pH 6.0) with 25 mM NaCl, using 1 mM ascorbic acid  
564 as reductant and 1 µM of LPMO, in a final volume of 100 µL at 30 °C and 950 rpm using the  
565 Eppendorf thermomixer. All assays were performed in triplicate. After centrifugation, 2 µL of the  
566 reaction supernatants were applied to a MTP 384 ground steel target plate TF (Bruker Daltonics)  
567 and mixed with 2 µL of 10 mg/mL 2,5-dihydroxybenzoic acid (DHB) dissolved in 30% acetonitrile,  
568 0.1% trifluoroacetic acid (TFA), followed by air-drying. Data collection was carried out according  
569 to Abdul Rahman et al., (2014) (71), using a matrix-assisted laser desorption/ionization time-of-  
570 flight (MALDI-ToF/TOF-MS) UltrafleXtreme® or Ultraflex III (Bruker Daltonics GmbH, Bremen,  
571 Germany). The data was collected using a 2-kHz smart beam-II laser and acquired in reflector  
572 mode (mass range 500–3500 Da) for MS analysis. Data analyses were performed in the Bruker  
573 FlexControl or FlexAnalysis software, using the peptide standard as a calibrant and supplied by  
574 Bruker Daltonics. The peak series was compared with previous reports (16, 69).

## 576 **5. References**

- 578 1. Ciano, L., Davies, G. J., Tolman, W. B., and Walton, P. H. (2018) Bracing copper for the  
579 catalytic oxidation of C–H bonds. *Nature Catalysis*. **1**, 571
- 580 2. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M., and Eijsink, V.  
581 G. H. (2010) An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant  
582 Polysaccharides. *Science*. **330**, 219–222
- 583 3. Hemsworth, G. R., Johnston, E. M., Davies, G. J., and Walton, P. H. (2015) Lytic  
584 Polysaccharide Monooxygenases in Biomass Conversion. *Trends in Biotechnology*. **33**, 747–761
- 585 4. Horn, S., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V. G. (2012) Novel enzymes for  
586 the degradation of cellulose. *Biotechnology for Biofuels*. **5**, 45–45
- 587 5. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013)  
588 Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox  
589 enzymes. *Biotechnol Biofuels*. **6**, 41
- 590 6. Garron, M.-L., and Henrissat, B. (2019) The continuing expansion of CAZymes and their  
591 families. *Current Opinion in Chemical Biology*. **53**, 82–87
- 592 7. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J.-C. N., Johansen, K.  
593 S., Krogh, K. B. R. M., Jorgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P.,  
594 Dupree, P., Xu, F., Davies, G. J., and Walton, P. H. (2011) Insights into the oxidative degradation  
595 of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the*  
596 *National Academy of Sciences*. **108**, 15079–15084

- 597 8. Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J.  
598 P., Parkin, A., Davies, G. J., and Walton, P. H. (2013) The copper active site of CBM33  
599 polysaccharide oxygenases. *Journal of the American Chemical Society*. **135**, 6069–6077
- 600 9. Paradisi, A., Johnston, E. M., Tovborg, M., Nicoll, C. R., Ciano, L., Dowle, A., McMaster,  
601 J., Hancock, Y., Davies, G. J., and Walton, P. H. (2019) Formation of a Copper(II)–Tyrosyl  
602 Complex at the Active Site of Lytic Polysaccharide Monooxygenases Following Oxidation by  
603 H<sub>2</sub>O<sub>2</sub>. *J. Am. Chem. Soc.* **141**, 18585–18599
- 604 10. Loose, J. S. M., Arntzen, M. Ø., Bissaro, B., Ludwig, R., Eijsink, V. G. H., and Vaaje-  
605 Kolstad, G. (2018) Multipoint Precision Binding of Substrate Protects Lytic Polysaccharide  
606 Monooxygenases from Self-Destructive Off-Pathway Processes. *Biochemistry*. **57**, 4114–4124
- 607 11. Fowler, Claire. A., Sabbadin, F., Ciano, L., Hemsworth, G. R., Elias, L., Bruce, N.,  
608 McQueen-Mason, S., Davies, G. J., and Walton, P. H. (2019) Discovery, activity and  
609 characterisation of an AA10 lytic polysaccharide oxygenase from the shipworm symbiont  
610 *Teredinibacter turnerae*. *Biotechnology for Biofuels*. **12**, 232
- 611 12. Munzone, A., El Kerdi, B., Fanuel, M., Rogniaux, H., Ropartz, D., Réglie, M., Royant, A.,  
612 Simaan, A. J., and Decroos, C. (2020) Characterization of a bacterial copper-dependent lytic  
613 polysaccharide monooxygenase with an unusual second coordination sphere. *FEBS J.*  
614 10.1111/febs.15203
- 615 13. Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D., and Marletta, M. A. (2014)  
616 Determinants of Regioselective Hydroxylation in the Fungal Polysaccharide Monooxygenases. *J.*  
617 *Am. Chem. Soc.* **136**, 562–565
- 618 14. Chen, J., Guo, X., Zhu, M., Chen, C., and Li, D. (2019) Polysaccharide monooxygenase-  
619 catalyzed oxidation of cellulose to glucuronic acid-containing cello-oligosaccharides. *Biotechnol*  
620 *Biofuels*. **12**, 42
- 621 15. Chen, C., Chen, J., Geng, Z., Wang, M., Liu, N., and Li, D. (2018) Regioselectivity of  
622 oxidation by a polysaccharide monooxygenase from *Chaetomium thermophilum*. *Biotechnology*  
623 *for Biofuels*. **11**, 155
- 624 16. Sabbadin, F., Hemsworth, G. R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T.,  
625 Marques, R. D. S., Sweeney, S. T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A. A., Bates,  
626 R., Gomez, L. D., Simister, R., Davies, G. J., Walton, P. H., Bruce, N. C., and McQueen-Mason,  
627 S. J. (2018) An ancient family of lytic polysaccharide monooxygenases with roles in arthropod  
628 development and biomass digestion. *Nat Commun.* **9**, 756
- 629 17. Schyra, J., and Korb, J. (2019) Termite Communities along A Disturbance Gradient in a  
630 West African Savanna. *Insects*. **10**, 17
- 631 18. Casalla, R., and Korb, J. (2019) Termite diversity in Neotropical dry forests of Colombia  
632 and the potential role of rainfall in structuring termite diversity. *Biotropica*. **51**, 165–177
- 633 19. Griffiths, H. M., Ashton, L. A., Evans, T. A., Parr, C. L., and Eggleton, P. (2019) Termites  
634 can decompose more than half of deadwood in tropical rainforest. *Current Biology*. **29**, R118–  
635 R119
- 636 20. Ashton, L. A., Griffiths, H. M., Parr, C. L., Evans, T. A., Didham, R. K., Hasan, F., Teh, Y.  
637 A., Tin, H. S., Vairappan, C. S., and Eggleton, P. (2019) Termites mitigate the effects of drought  
638 in tropical rainforest. *Science*. **363**, 174–177
- 639 21. Fujita, A., Hojo, M., Aoyagi, T., Hayashi, Y., Arakawa, G., Tokuda, G., and Watanabe, H.  
640 (2010) Details of the digestive system in the midgut of *Coptotermes formosanus* Shiraki. *Journal*  
641 *of Wood Science*. **56**, 222–226
- 642 22. Tokuda, G., Watanabe, H., and Lo, N. (2007) Does correlation of cellulase gene  
643 expression and cellulolytic activity in the gut of termite suggest synergistic collaboration of  
644 cellulases? *Gene*. **401**, 131–134
- 645 23. Brune, A., and Ohkuma, M. (2011) Role of the Termite Gut Microbiota in Symbiotic  
646 Digestion. in *Biology of Termites: a Modern Synthesis* (Bignell, D. E., Roisin, Y., and Lo, N. eds),  
647 pp. 439–475, Springer Netherlands
- 648 24. Ni, J., and Tokuda, G. (2013) Lignocellulose-degrading enzymes from termites and their  
649 symbiotic microbiota. *Biotechnology Advances*. **31**, 838–850
- 650 25. Scharf, M. E., Karl, Z. J., Sethi, A., and Boucias, D. G. (2011) Multiple levels of synergistic  
651 collaboration in termite lignocellulose digestion. *PLoS One*. **6**, e21709

- 652 26. Franco Cairo, J. P. L., Oliveira, L. C., Uchima, C. a., Alvarez, T. M., Citadini, A. P. D. S.,  
653 Cota, J., Leonardo, F. C., Costa-Leonardo, A. M., Carazzolle, M. F., Costa, F. F., Pereira, G. a  
654 G., and Squina, F. M. (2013) Deciphering the synergism of endogenous glycoside hydrolase  
655 families 1 and 9 from *Coptotermes gestroi*. *Insect Biochemistry and Molecular Biology*. **43**, 970–  
656 981
- 657 27. Terrapon, N., Li, C., Robertson, H. M., Ji, L., Meng, X., Booth, W., Chen, Z., Childers, C.  
658 P., Glastad, K. M., Gokhale, K., Gowin, J., Gronenberg, W., Hermansen, R. A., Hu, H., Hunt, B.  
659 G., Huylmans, A. K., Khalil, S. M. S., Mitchell, R. D., Munoz-Torres, M. C., Mustard, J. A., Pan,  
660 H., Reese, J. T., Scharf, M. E., Sun, F., Vogel, H., Xiao, J., Yang, W., Yang, Z., Yang, Z., Zhou,  
661 J., Zhu, J., Brent, C. S., Elsik, C. G., Goodisman, M. A. D., Liberles, D. A., Roe, R. M., Vargo, E.  
662 L., Vilcinskis, A., Wang, J., Bornberg-Bauer, E., Korb, J., Zhang, G., and Liebig, J. (2014)  
663 Molecular traces of alternative social organization in a termite genome. *Nature Communications*.  
664 10.1038/ncomms4636
- 665 28. Poulsen, M., Hu, H., Li, C., Chen, Z., Xu, L., Otani, S., Nygaard, S., Nobre, T., Klaubauf,  
666 S., Schindler, P. M., Hauser, F., Pan, H., Yang, Z., Sonnenberg, A. S., de Beer, Z. W., Zhang, Y.,  
667 Wingfield, M. J., Grimmelikhuijzen, C. J., de Vries, R. P., Korb, J., Aanen, D. K., Wang, J.,  
668 Boomsma, J. J., and Zhang, G. (2014) Complementary symbiont contributions to plant  
669 decomposition in a fungus-farming termite. *Proc Natl Acad Sci U S A*. **111**, 14500–14505
- 670 29. Harrison, M. C., Jongepier, E., Robertson, H. M., Arning, N., Bitard-Feildel, T., Chao, H.,  
671 Childers, C. P., Dinh, H., Doddapaneni, H., Dugan, S., Gowin, J., Greiner, C., Han, Y., Hu, H.,  
672 Hughes, D. S. T., Huylmans, A.-K., Kemena, C., Kremer, L. P. M., Lee, S. L., Lopez-Ezquerria,  
673 A., Mallet, L., Monroy-Kuhn, J. M., Moser, A., Murali, S. C., Muzny, D. M., Otani, S., Piulachs, M.-  
674 D., Poelchau, M., Qu, J., Schaub, F., Wada-Katsumata, A., Worley, K. C., Xie, Q., Ylla, G.,  
675 Poulsen, M., Gibbs, R. A., Schal, C., Richards, S., Belles, X., Korb, J., and Bornberg-Bauer, E.  
676 (2018) Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nature Ecology &*  
677 *Evolution*. **2**, 557–566
- 678 30. Sethi, A., Kovaleva, E. S., Slack, J. M., Brown, S., Buchman, G. W., and Scharf, M. E.  
679 (2013) A GHF7 cellulase from the protist symbiont community of *Reticulitermes flavipes* enables  
680 more efficient lignocellulose processing by host enzymes. *Arch Insect Biochem Physiol*. **84**, 175–  
681 193
- 682 31. Woon, J. S.-K., King, P. J. H., Mackeen, M. M., Mahadi, N. M., Wan Seman, W. M. K.,  
683 Broughton, W. J., Abdul Murad, A. M., and Abu Bakar, F. D. (2017) Cloning, Production and  
684 Characterization of a Glycoside Hydrolase Family 7 Enzyme from the Gut Microbiota of the  
685 Termite *Coptotermes curvignathus*. *Mol Biotechnol*. **59**, 271–283
- 686 32. Cannella, D., Hsieh, C. C., Felby, C., and Jørgensen, H. (2012) Production and effect of  
687 aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content.  
688 *Biotechnology for Biofuels*. **5**, 26–26
- 689 33. Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J.-C., Brown, K.,  
690 Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., and Lo Leggio, L.  
691 (2010) Stimulation of Lignocellulosic Biomass Hydrolysis by Proteins of Glycoside Hydrolase  
692 Family 61: Structure and Function of a Large, Enigmatic Family. *Biochemistry*. **49**, 3305–3316
- 693 34. Franco Cairo, João. P. L., Carazzolle, M. F., Leonardo, F. C., Mofatto, L. S., Brenelli, L. B.,  
694 Gonçalves, T. A., Uchima, C. A., Domingues, R. R., Alvarez, T. M., Tramontina, R., Vidal, R. O.,  
695 Costa, F. F., Costa-Leonardo, A. M., Leme, P., F, A., Pereira, G. A. G., and Squina, F. M. (2016)  
696 Expanding the Knowledge on Lignocellulolytic and Redox Enzymes of Worker and Soldier  
697 Castes from the Lower Termite *Coptotermes gestroi*. *Front. Microbiol*. 10.3389/fmicb.2016.01518
- 698 35. Tramontina, R., Franco Cairo, J. P. L., Liberato, M. V., Mandelli, F., Sousa, A., Santos, S.,  
699 Rabelo, S. C., Campos, B., Ienczak, J., Ruller, R., Damásio, A. R. L., and Squina, F. M. (2017)  
700 The *Coptotermes gestroi* aldo–keto reductase: a multipurpose enzyme for biorefinery  
701 applications. *Biotechnology for Biofuels*. **10**, 4
- 702 36. Bissaro, B., Várnai, A., Røhr, Å. K., and Eijsink, V. G. H. (2018) Oxidoreductases and  
703 Reactive Oxygen Species in Conversion of Lignocellulosic Biomass. *Microbiol. Mol. Biol. Rev.* **82**,  
704 e00029-18
- 705 37. Geng, A., Cheng, Y., Wang, Y., Zhu, D., Le, Y., Wu, J., Xie, R., Yuan, J. S., and Sun, J.  
706 (2018) Transcriptome analysis of the digestive system of a wood-feeding termite (*Coptotermes*

- 707 formosanus) revealed a unique mechanism for effective biomass degradation. *Biotechnology for*  
708 *Biofuels*. **11**, 24
- 709 38. Brune, A. (2014) Symbiotic digestion of lignocellulose in termite guts. *Nature Reviews*  
710 *Microbiology*. **12**, 168–180
- 711 39. Tokuda, G. (2019) Plant cell wall degradation in insects: Recent progress on endogenous  
712 enzymes revealed by multi-omics technologies. in *Advances in Insect Physiology* (Jurenka, R.  
713 ed), pp. 97–136, Academic Press, **57**, 97–136
- 714 40. Govorushko, S. (2019) Economic and ecological importance of termites: A global review.  
715 *Entomological Science*. **22**, 21–35
- 716 41. Kolliopoulou, A., Taning, C. N. T., Smagghe, G., and Swevers, L. (2017) Viral Delivery of  
717 dsRNA for Control of Insect Agricultural Pests and Vectors of Human Disease: Prospects and  
718 Challenges. *Front Physiol*. 10.3389/fphys.2017.00399
- 719 42. Wu, W., Gu, D., Yan, S., and Li, Z. (2019) RNA interference of endoglucanases in the  
720 formosan subterranean termite *Coptotermes formosanus shiraki* (Blattodea: Rhinotermitidae) by  
721 dsRNA injection or ingestion. *Journal of Insect Physiology*. **112**, 15–22
- 722 43. Chouvinc, T., Li, H. F., Austin, J., Bordereau, C., Bourguignon, T., Cameron, S. L.,  
723 Canello, E. M., Constantino, R., Costa-Leonardo, A. M., Eggleton, P., Evans, T. A., Forschler,  
724 B., Grace, J. K., Husseneder, C., Křeček, J., Lee, C. Y., Lee, T., Lo, N., Messenger, M., Mullins,  
725 A., Robert, A., Roisin, Y., Scheffrahn, R. H., Sillam-Dussès, D., Šobotník, J., Szalanski, A.,  
726 Takematsu, Y., Vargo, E. L., Yamada, A., Yoshimura, T., and Su, N. Y. (2015) Revisiting  
727 *Coptotermes* (Isoptera: Rhinotermitidae): A global taxonomic road map for species validity and  
728 distribution of an economically important subterranean termite genus. *Systematic Entomology*.  
729 10.1111/syen.12157
- 730 44. Hapukotuwa, N. K., and Grace, J. K. (2011) Preferences of *Coptotermes formosanus*  
731 *Shiraki* and *Coptotermes gestroi* (Wasmann) (Blattodea: Rhinotermitidae) among Three  
732 Commercial Wood Species. *Insects*. **2**, 499–508
- 733 45. Kirton, L., and Brown, V. (2003) The Taxonomic Status of Pest Species of *Coptotermes* in  
734 Southeast Asia: Resolving the Paradox in the Pest Status of the Termites, *Coptotermes gestroi*,  
735 *C. havilandi* and *C. travians* (Isoptera: Rhinotermitidae). *Sociobiology*
- 736 46. Vaaje-Kolstad, G., Forsberg, Z., Loose, J. S., Bissaro, B., and Eijsink, V. G. (2017)  
737 Structural diversity of lytic polysaccharide monoxygenases. *Curr. Opin. Struct. Biol.* **44**, 67–76
- 738 47. Jensen, M. S., Klinkenberg, G., Bissaro, B., Chylenski, P., Vaaje-Kolstad, G., Kvitvang, H.  
739 F., Nærdal, G. K., Sletta, H., Forsberg, Z., and Eijsink, V. G. H. (2019) Engineering chitinolytic  
740 activity into a cellulose-active lytic polysaccharide monoxygenase provides insights into  
741 substrate specificity. *J. Biol. Chem.* 10.1074/jbc.RA119.010056
- 742 48. Breslmayr, E., Hanžek, M., Hanrahan, A., Leitner, C., Kittl, R., Šantek, B., Oostenbrink, C.,  
743 and Ludwig, R. (2018) A fast and sensitive activity assay for lytic polysaccharide  
744 monoxygenase. *Biotechnol Biofuels*. **11**, 79
- 745 49. Gregory, R. C., Hemsworth, G. R., Turkenburg, J. P., Hart, S. J., Walton, P. H., and  
746 Davies, G. J. (2016) Activity, stability and 3-D structure of the Cu(ii) form of a chitin-active lytic  
747 polysaccharide monoxygenase from *Bacillus amyloliquefaciens*. *Dalton Trans.* **45**, 16904–  
748 16912
- 749 50. Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., and Ludwig, R. (2012) Production of four  
750 *Neurospora crassa* lytic polysaccharide monoxygenases in *Pichia pastoris* monitored by a  
751 fluorimetric assay. *Biotechnology for Biofuels*. **5**, 79
- 752 51. Forsberg, Z., Røhr, Å. K., Mekasha, S., Andersson, K. K., Eijsink, V. G. H., Vaaje-Kolstad,  
753 G., and Sørli, M. (2014) Comparative study of two chitin-active and two cellulose-active AA10-  
754 type lytic polysaccharide monoxygenases. *Biochemistry*. **53**, 1647–1656
- 755 52. Zhou, X., Qi, X., Huang, H., and Zhu, H. (2019) Sequence and Structural Analysis of AA9  
756 and AA10 LPMOs: An Insight into the Basis of Substrate Specificity and Regioselectivity. *Int J*  
757 *Mol Sci*. 10.3390/ijms20184594
- 758 53. Kruer-Zerhusen, N., Alahuhta, M., Lunin, V. V., Himmel, M. E., Bomble, Y. J., and Wilson,  
759 D. B. (2017) Structure of a *Thermobifida fusca* lytic polysaccharide monoxygenase and  
760 mutagenesis of key residues. *Biotechnology for Biofuels*. **10**, 243
- 761 54. Forsberg, Z., Mackenzie, A. K., Sørli, M., Røhr, Å. K., Helland, R., Arvai, A. S., Vaaje-

- 762 Kolstad, G., and Eijsink, V. G. H. (2014) Structural and functional characterization of a conserved  
763 pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proceedings of the*  
764 *National Academy of Sciences of the United States of America*. **111**, 8446–51
- 765 55. Calusinska, M., Marynowska, M., Bertucci, M., Untereiner, B., Klimek, D., Goux, X.,  
766 Sillam-Dussès, D., Gawron, P., Halder, R., Wilmes, P., Ferrer, P., Gerin, P., Roisin, Y., and  
767 Delfosse, P. (2020) Integrative omics analysis of the termite gut system adaptation to *Miscanthus*  
768 diet identifies lignocellulose degradation enzymes. *Communications Biology*. **3**, 1–12
- 769 56. Chouvenc, T. (2020) Limited survival strategy in starving subterranean termite colonies.  
770 *Insectes Sociaux*. 10.1007/s00040-019-00729-5
- 771 57. Alba, R., Fei, Z., Payton, P., Liu, Y., Moore, S. L., Debbie, P., Cohn, J., D’Ascenzo, M.,  
772 Gordon, J. S., Rose, J. K. C., Martin, G., Tanksley, S. D., Bouzayen, M., Jahn, M. M., and  
773 Giovannoni, J. (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for  
774 dissecting plant physiology and development. *The Plant Journal*. **39**, 697–714
- 775 58. Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., Rattei,  
776 T., Mende, D. R., Sunagawa, S., Kuhn, M., Jensen, L. J., von Mering, C., and Bork, P. (2016)  
777 eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for  
778 eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res*. **44**, D286–D293
- 779 59. Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0:  
780 discriminating signal peptides from transmembrane regions. *Nat Methods*. **8**, 785–786
- 781 60. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014)  
782 The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. **42**, D490–5
- 783 61. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F.,  
784 Guindon, S., Lefort, V., Lescot, M., Claverie, J.-M., and Gascuel, O. (2008) Phylogeny.fr: robust  
785 phylogenetic analysis for the non-specialist. *Nucleic Acids Res*. **36**, W465–469
- 786 62. Do, C. B., Mahabhashyam, M. S. P., Brudno, M., and Batzoglou, S. (2005) ProbCons:  
787 Probabilistic consistency-based multiple sequence alignment. *Genome Res*. **15**, 330–340
- 788 63. Guindon, S., Delsuc, F., Dufayard, J.-F., and Gascuel, O. (2009) Estimating maximum  
789 likelihood phylogenies with PhyML. *Methods Mol. Biol*. **537**, 113–137
- 790 64. Anisimova, M., and Gascuel, O. (2006) Approximate likelihood-ratio test for branches: A  
791 fast, accurate, and powerful alternative. *Syst. Biol*. **55**, 539–552
- 792 65. Letunic, I., and Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the  
793 display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. **44**, W242–W245
- 794 66. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite:  
795 protein structure and function prediction. *Nat. Methods*. **12**, 7–8
- 796 67. Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H.,  
797 Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes,  
798 D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G.-W., Holst, M. J.,  
799 McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation  
800 software suite. *Protein Science*. **27**, 112–128
- 801 68. Neu, H. C., and Heppel, L. A. (1965) The release of enzymes from *Escherichia coli* by  
802 osmotic shock and during the formation of spheroplasts. *J. Biol. Chem*. **240**, 3685–3692
- 803 69. Hemsworth, G. R., Henrissat, B., Davies, G. J., and Walton, P. H. (2013) Discovery and  
804 characterization of a new family of lytic polysaccharide monooxygenases. *Nature Chemical*  
805 *Biology*. **10**, 122–126
- 806 70. Huynh, K., and Partch, C. L. (2015) Current Protocols in Protein Science. *Curr Protoc*  
807 *Protein Sci*. **79**, 28.9.1–28.9.14
- 808 71. Abdul Rahman, S., Bergström, E., Watson, C. J., Wilson, K. M., Ashford, D. A., Thomas, J.  
809 R., Ungar, D., and Thomas-Oates, J. E. (2014) Filter-Aided N-Glycan Separation (FANGS): A  
810 Convenient Sample Preparation Method for Mass Spectrometric N-Glycan Profiling. *J Proteome*  
811 *Res*. **13**, 1167–1176
- 812

## 813 **6. Figure legends**

814

815 **Figure 1. Phylogenetic analysis with termites LPMOs.** A) Radial phylogram of LPMO families  
816 for the classification of termite enzymes. All AA families were clustered with high bootstraps (>  
817 0.94). B) Detail of the AA15 branch from radial phylogram, showing all termite LPMOs. *C. gestroi*  
818 enzymes are highlighted with red circles and *T. domestica* with orange asterisks. The LPMOs  
819 AA15a from *Z. nevadensis*, *C. secundus*, *M. natalensis* and *Cortaritermes* sp. are highlighted with  
820 blue rhombus, squares, asterisks and circles, respectively. The termite AA15s are grouped in three  
821 different clades (light grey boxes), each with a different protein architecture: 1, 2 and 3 in blue  
822 boxes. SP= Signal peptide, TM=transmembrane. CT= cytosolic. The sequences from *MnAA15b*,  
823 *CgAA15c*, *CgAA15d*, *TdAA15a* and *TdAA15b* are partial sequences from mRNA sequencing.

824

825 **Figure 2. The structural hallmarks of CgAA15s LPMOs.** A) Overall homology models for  
826 *CgAA15a* (I) and *CgAA15b\_cat* (II) in comparison with *TdAA15a* (PDB code 5MSZ) (III). The  
827 copper ion is in gold. B) Detail of the catalytic site showing the coordinated copper-His-brace and  
828 the non-coordinated Ala and Phe residues for *CgAA15a* (I), *CgAA15b\_cat* (II) and *TdAA15a* (III).  
829 C) The electrostatic surface potential of *CgAA15a* (I), *CgAA15b\_cat* (II) and *TdAA15a* (III), the  
830 putative binding site in yellow ellipses and the copper in mocha circles.

831

832 **Figure 3. The specific activities and the apparent melting temperature of C. gestroi AA15s.**

833 A) Determination of specific activity of *CgAA15a*, *CgAA15b\_cat* and *CgAA15a*<sup>F182Y</sup> using the 2,6-  
834 DMP assay. Thermal shift curves of *CgAA15a* (B), *CgAA15b\_cat* (C) and *CgAA15a*<sup>F182Y</sup> (D) and  
835 the melting temperature ( $T_m$ ) of the copper ion loaded version of the enzymes (right curves) and  
836 the apo form after EDTA treatment (left curves).

837

838 **Figure 4. Determination in-vitro of the optimal electron donor for C. gestroi AA15s.** The  
839 Amplex®Red assay was applied for the quantification of H<sub>2</sub>O<sub>2</sub> and to determine the optimal  
840 electron donor for the LPMOs. Four molecules were used as donors: L-cysteine (blue rhombus),  
841 Ascorbic acid (yellow times), Pyrogallol (grey triangles) and Gallic Acid (orange boxes). Kinetic  
842 measurements for (A) *CgAA15a*, (B) *CgAA15b\_cat* and (C) *CgAA15a*<sup>F182Y</sup> were performed for 35  
843 min.

844

845 **Figure 5. Substrate specificity of C. gestroi AA15s.** MALDI-TOF MS spectrum of reaction  
846 products after the incubation of 1 μM *CgAA15a* (A), *CgAA15b\_cat* and (C) *CgAA15a*<sup>F182Y</sup> (E) with  
847 1% chitin from squid pen in the presence of 2 mM ascorbic acid as the electron donor. Detailed  
848 view of the expanded spectrum for DP6, DP8 and DP6 of the reactions applying *CgAA15a* (B),  
849 *CgAA15b\_cat* (D) and *CgAA15a*<sup>F182Y</sup> (F), respectively. Native species are marked in black and  
850 the oxidized species in red.

851

852 **Figure S1. Amino acid sequence alignment among termites and *T domestica* LPMOs from**  
853 **family AA15.** The alignment was generated using T-COFFEE (Homology Extension PSI-coffee  
854 mode). Blue Navy empty boxes for the conserved amino acids composing the catalytic site (his-  
855 brace and axial residues). Grey empty box for the conserved protrusion region. Red empty boxes  
856 for conserved aromatic residues and salmon for polar residues exposed in the flat catalytic  
857 surface of AA15s.

858

859 **Figure S2. AA15s transcript counts in castes of *C. gestroi* and CDS amplifications. A)**  
860 Normalized number of transcripts of *CGAA15s* between worker (red) and soldier (light red)  
861 castes of the lower termite *C. gestroi*. **B)** 1% agarose gels showing the amplification products for  
862 AA15 CDSs from cDNA of the *C. gestroi* work caste. L = Ladder, 1-3 = *CGAA15A*, 4 and 5 =  
863 *CGAA15b\_cat* and 6 = negative control.

864

865 **Figure S3. Models for the flat catalytic surface of *C. gestroi* AA15s. A)** Aromatic residues  
866 (red) exposed on the flat catalytic surface of *CgAA15a* (I), for *CgAA15b-cat* (II) and *TdAA15a*  
867 PDB id 5MSZ (III). **B)** The negatively charged residues (orange) for the same LPMOs and **C)**  
868 Superposition of aromatic and negatively charged residues of the three LPMOs.

869

870 **Figure S4. Recombinant protein production and purification. A)** Schematic view of the  
871 expression vector (pET-26b) and the protein construction. The protein was cloned in frame with  
872 the pelB (signal peptide) and also with a streptavidin tag (S-Tag) for protein purification. **B)** 12%  
873 SDS-PAGE gel of protein extraction and purification for *CgAA15a*: M = ladder, 1 = Osmotic shock  
874 fraction, 2 = column flowthrough, 3 = wash fraction, 4-7 = Elution fractions with 3 mM  
875 desthiobiotin. **C)** 12% SDS-PAGE gel of protein extraction and purification for *CgAA15b\_cat*: M =  
876 ladder, 1 = Osmotic shock fraction, 2 = column flowthrough, 3 = wash fraction, 4 and 5 = Elution  
877 fractions with 3 mM desthiobiotin. **D)** 12% SDS-PAGE gel of protein extraction and purification for  
878 *CgAA15a*<sup>F182Y</sup>: M = ladder, 1 = Osmotic shock fraction, 2 = column flowthrough, 3 = wash  
879 fraction, 4-6 = Elution fractions with 3 mM d-Desthiobiotin.

880

881 **Figure S5. Substrate specificity of *C. gestroi* LPMOs.** MALDI-TOF MS spectrum of control  
882 reaction products. **A)** Spectra for 1% chitin from squid pen and 2 mM ascorbic acid. **B)** Spectra  
883 for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1  $\mu$ M of *CgAA15a*. **C)** Spectra  
884 for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1  $\mu$ M of *CgAA15b\_cat*. **D)**  
885 Spectra for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1  $\mu$ M of  
886 *CgAA15a*<sup>F182Y</sup>. A range of different polysaccharides were tested within the same conditions and  
887 controls (data not shown).

888



889 **Figure S6. Structural hallmarks of *TdAA15b*.** **A)** Overall model for the structure based  
890 on *TdAA15a* PDB id 5MSZ (III). The calculated C-Score was 1.71 with estimated RMSD of  
891  $2.0 \pm 1.6 \text{ \AA}$ . The copper metal is in gold and the catalytic residues His1, His91 and Phe184 are in  
892 blue. **B)** The electrostatic surface potential of *TdAA15b*, showing the protrusion in black circles,  
893 the putative binding site in yellow ellipses and copper in moca circles.

894

## 895 **7. Acknowledgements**

896 This research was supported by grants from Fundação de Amparo à Pesquisa do Estado  
897 de São Paulo - FAPESP (FMS 2015/50590-4; JPLFC 2016/09950-0; 2017/11952-3; WG  
898 2017/17275-3; TAG 2017/16089-1) and Conselho Nacional de Desenvolvimento Científico e  
899 Tecnológico - CNPq. (FMS 428527/2018-3 and 305740/2017-2; WG 422132/2018-7). We would  
900 like also to acknowledge the MASS facility from Brazilian Biosciences National Laboratory (LNBio  
901 – CNPEM) and the Analytic Center from the Institute of Chemistry from University of Campinas.  
902 This research was supported by grants from BBSRC XXXX.

903

## 904 **8. Authors contributions**

905 JPLFC designed and performed the experiments and analyzed the data as well as draft  
906 the manuscript. AP performed experiments and analyzed the data. TAG designed experiments  
907 and discussed the data. MVR performed phylogenetic analysis and analyzed the data. LSM  
908 performed the bioinformatic analyses. DAC designed experiments, analyzed the data and draft  
909 the manuscript. LC performed experiments, analyzed the data and draft the manuscript. CRFT  
910 designed and performed experiments. WG performed experiments and wrote the draft  
911 manuscript. MFC designed and performed the bioinformatic analyses. GJD analyzed the data  
912 and revised the final version of the manuscript. CF designed, economically supported the  
913 experiments and revised the final version of the manuscript. PHW designed experiments,  
914 analyzed the data, economically supported the work and drafted the manuscript. FMS designed  
915 experiments, analyzed the data, economically supported the work and revised the manuscript.

916

## 917 **9. Conflict of interest**

918 The authors declare no conflict of interest.