
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

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

Titles:

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

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

Affiliations:

[1]Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas, Campinas, SP, Brazil.
[2]Department of Chemistry, University of York, Heslington, York, United Kingdom.
[3]Programa de Processos Tecnológicos e Ambientais, Universidade de Sorocaba - UNISO, Sorocaba, SP, Brazil.
[4]Interfaculty School of Bioengineers, Université Libre de Bruxelles, Brussels, Belgium.
[5]Department of Genetic, Evolution and Bioagents, Institute of Biology, University of Campinas, Campinas, SP, Brazil.
[6]Centro de Ciências Naturais e Humanas, Universidade Federal do ABC (UFABC), Santo André, SP, Brasil.
[7]Department of Geosciences and Natural Resource Management, Faculty of Science, University of Copenhagen, Frederiksberg C, Denmark.

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

Abstract (200 words)

Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent enzymes involved in the oxidative cleavage of polysaccharides. LPMOs belonging to the family AA15 are found widespread across the Tree of Life, including viruses, algae, oomycetes and animals. Recently, two AA15s from the insect *Thermobia domestica* were reported to have oxidative activity, the first against cellulose and chitin and the latter only on chitin. Herein, we report the identification and characterization of two novel family AA15 members from the termite *Coptotermes gestroi*. Structural models indicate that the LPMOs from *C. gestroi* (*Cg*AA15a/*Cg*AA15b) and *T. domestica* (*Td*AA15a) have a similar fold but with differences in the catalytic site residues, as well as the electrostatic potential of the binding surfaces. According to biochemical and biophysical studies, the copper ion had a thermostabilizing effect on both *C. gestroi* AA15 structures and ascorbate is an electron donor for H_2O_2 production. The *Cg*AA15a and *Cg*AA15b were able to only oxidize chitin from squid pen, with no activity on cellulose, xylan, xyloglucan and starch. To mimic the catalytic pocket of the cellulose and chitin-active *Td*AA15a, the Cu proximal axial and non-coordinating phenylalanine was substituted with tyrosine in the active site of *Cg*AA15a, but the mutation did not change the enzymatic specificity. Our data indicated that these enzymes are not involved in lignocellulose digestion and might play a role in developmental processes. Furthermore, this study illuminates the role of the Cu proximal axial aromatic residue on catalytic specificity of AA15 family members.

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 β-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 (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 (*Td*AA15a and *Td*AA15b) were biochemically and structurally characterized (16). *Td*AA15a was one of the most abundant LPMOs from *T. domestica*'s digestome and can oxidize both cellulose

and chitin substrates. *Td*AA15b was only able to cleave chitin, suggesting a role in the development and cuticle/exoskeleton remodelling of this insect (16). Both *Td*AA15a and *Td*AA15b share the conserved immunoglobulin-like protein fold as well as the His-brace, which is ubiquitous for all LPMOs. However, *Td*AA15a and *Td*AA15b differ by the nature of the aromatic residue found in the axial position of the Cu ion, being a Tyr or a Phe residue for the two enzymes, respectively (16). Because *T. domestica* is a primitive cellulolytic insect, we decided to investigate whether termites, which are among the major lignocellulolytic organisms on Earth, have LPMOs in their genomes and what role these enzymes might play.

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

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

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

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

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

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 *Td*AA15a and *Td*AA15b (16). These analyses permitted for classifying the *C. gestroi* LPMOs into the AA15 family. Thus, the *C. gestroi* full-length CDSs were named *Cg*AA15a and *Cg*AA15b, whereas the other CDSs are named *Cg*AA15c_partial and *Cg*AA15d_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 β-tongue-like protrusion region identified in the AA15s from *T. domestica*, which extends the surface surrounding the active site, is also conserved in *Cg*AA15a and *Cg*AA15b (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 Eurarthropoda. LPMO sequences from other families were used as outgroups. The LPMOs from *C. gestro*i, *Z. nevadensis, C. secundus, M. natalensis* and *Cortaritermes* sp. were clustered together along with the sequences of AA15s from *T.*

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

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

Transcripts encoding for *C. gestroi* AA15s were identified in the transcriptomic database of *C. gestroi* castes (based on 454-pyrosequences) (34), indicating their expression in both workers and soldiers. *CGAA15a* was found expressed in both termite castes as well as *CGAA15d*_partial, while *CGAA15b* was expressed only in soldiers. *CGAA15c*_partial had its transcripts identified only in the worker caste (Fig. S2A). Moreover, recent genomic and RNA-Seq analyses using Illumina technology were able to identify transcripts of *CGAA15a* as well as *CGAA15b* in the worker caste of *C. gestroi* (in preparation), confirming the results from the pyrosequencing analysis. The full-length CDSs for *CGAA15a* and *CGAA15b* have 633 and 1020 bp, respectively, including the signal peptide sequences for both enzymes and the transmembrane and cytosolic domains for *CGAA15b*. Based on these results, oligonucleotides were designed and only the catalytic domains of *CGAA15a* and *CGAA15b* from *C. gestroi* worker caste were amplified. After RNA extraction from the whole *C. gestroi* bodies and its transcription to cDNA, the sequences were amplified without the sequence encoding for the signal peptide, producing two fragments with 573 bp for C*GAA15A* and 591 bp for *CgAA15B*_cat, which are in accord with our predictions (Fig. S2B).

2.3. Comparative and structural insights of *Cg*AA15a and *Cg*AA15b_cat reveals conserved hallmarks of AA15 family.

To obtain structural information regarding *Cg*AA15a and *Cg*AA15b_cat, both enzymes were modelled by employing the I-TASSER server and using the *Td*AA15a structure (PDB code 5MSZ) as a template. *Td*AA15a has 47% and 44% sequence identity with *Cg*A115a and *Cg*AA15b_cat, respectively. Thus, the three-dimensional homology models of *Cg*AA15a and *Cg*AA15b_cat were built with a high C-Scores (-5 to 2) of 0.77 and 0.91 with estimated RMSDs of 3.8±2.6Å and 3.6±2.5Å, respectively.

The AA15 homology models of *C. gestroi* showed a typical central and antiparallel β-sandwich fold of LPMOs, with several loops that are stabilized by five disulphide bonds (Fig. 2A I-III), in agreement with the structure of *Td*AA15a (PDB code 5MSZ) (Fig. 2A III). *Cg*AA15a and *Cg*AA15b_cat active sites were predicted with the presence of a copper ion coordinated by the His-brace motif (7). For both enzymes, the conserved residues His1 and His91 are predicted to coordinate the copper ion with a T-shaped geometry (Fig. 2B I-III) as observed for all characterized LPMOs (8). The *Cg*AA15a and *Cg*AA15b_cat were predicted to have a Phe (Phe182 and Phe191, respectively) (Fig. 2B I and II), diverging from *Td*AA15a which presents a Tyr (Fig. 2B III) occupying the axial site of the Cu on the His Brace coordination plane. The apical position of the Cu ion is occupied by a non-coordinating Ala residue (Ala89), which is common to all three structures (Fig. 2A I-III). Moreover, the catalytic site of the AA15s from *C. gestroi* presents many similarities to the majority of chitin-active LPMOs from the AA10 family (46) and, looking in detail to the AA15 LPMOs alignment (Fig. S1), all termite LPMOs share the same active site.

The electrostatic potential molecular surface was generated for both *C. gestroi* AA15s (Fig. 2C I-II) and compared with that of *Td*AA15a (Fig. 2C III). Comparison shows that the overall surface charge for *Cg*AA15a, *Cg*AA15b_cat and *Td*AA15a are divergent, mainly regarding the negative potential (red) around the His-brace for *Cg*AA15a and *Cg*AA15b_cat (Fig 2C I-II – see the ellipses). However, similar potentials are found in the protrusion region (neutral) and around the His-brace (slightly positive potential) (blue). The putative docking site for a protein partner predicted for *Td*AA15a (PDB code 5MSZ) (16) is absent in both *C. gestroi* AA15s.

Aromatic and polar charged residues surrounding the active site are solvent-exposed residues on the flat catalytic surface and were previously suggested to be involved in substrate binding for other LPMO families (10, 46, 47). The AA15s from *C. gestroi* also have these solvent-exposed residues at the surface (Fig. S3A I-II) and surrounding the His-brace as observed for *Td*AA15a (Fig. S3A III). Concerning aromatic residues, the protrusion region (not found in all AA15s) has the aromatic residues Trp164 (for *Cg*AA15a) and Trp173 (for *Cg*AA15b_cat) conserved at the same positions. On the opposite side of the protrusion, the residues Phe24 and Tyr24, for *Cg*AA15a and *Cg*AA15b_cat, respectively, are also exposed on the surface and these aromatics are present in all termite AA15s and highly positional conserved among AA15 members (70%) (16).

Other aromatic residues are also present on the flat catalytic surface of the three structures analyzed, highlighting the Phe34 for all of them. Moreover, *Cg*AA15a and

*Cg*AA15b_cat have Phe30 and Tyr30 exposed near the His-brace, respectively, as well as a Trp38 a little further from the catalytic pocket (Fig. S3 A I-II). The other aromatic residues exposed in *Cg*AA15a are Tyr39 and Tyr166, contrasting with the positions of others aromatic residues in *Cg*AA15b_cat and *Td*AA15a. The alignment of *C. gestroi* LPMOs with the LPMOs from *Z. nevadensis* and *C. secundus* also showed that the Tyr24 is conserved for all of them, with an exception for *Cg*AA15a, whereas the Trp residue around the positions 164-173 (protrusion) is fully conserved among all termite sequences (Fig. S1).

Polar residues exposed on the flat catalytic surface of LPMOs are also reported to bind the substrates, and among them the negatively charged Glu (around the positions 180 – 189) is exposed around the His-brace and conserved for *Cg*AA15a, *Cg*AA15b_cat and *Td*AA15a structures (Fig. S3B I-III), as well as for the other termite AA15s (Fig. S1). Moreover, the residues Glu25 and Glu28 are conserved in both *C. gestroi* LPMOs (Fig. S3B I-II) and the alignment analysis suggests that these residues are also conserved in all other termites CDSs as well as in the specific chitin-active *Td*AA15b (Fig. S1) from *T. domestica*. However, these residues are absent on the surface of the cellulose/chitin-active *Td*AA15a. In the order hand, the Asp residues exposed on the flat surface at positions Asp27 and Asp35 are only found in the *Td*AA15a structure (Fig. S3B III).

Together the phylogenetic and structural analyses described in this work indicate that an intriguing difference between the catalytic sites among the *Cg*AA15a and *Cg*AA15b with *Td*AA15a is the non-coordinating and Cu proximal axial aromatic residue. *Td*AA15a, with Tyr in the axial position, was reported to oxidize cellulose and chitin substrates. On the other hand, *Td*AA15b, which has a Phe residue in the same position, can oxidize only chitin. A similar pattern is also observed for AA10s, where enzymes with Tyr in the same position can oxidize at least cellulose and the majority of AA10 enzymes with Phe can oxidize only chitin (1). Therefore, it seems that the aromatic axial residue may play a role in substrate specificity in AA15s and also in other LPMO families. Thus, to investigate the role of the non-coordinating and Cu proximal axial aromatic residue in *C. gestroi* AA15s, we functionally characterized *Cg*AA15a and *Cg*AA15b_cat as well as a mutated version of *Cg*AA15a, where the Phe182 was changed to a Tyr, herein named *Cg*AA15a^{F182Y}.

2.4. Functional characterization of *Cg*AA15a, *Cg*AA15b_cat and *Cg*AA15a^{F182Y}.

The CDS regions coding for the LPMO domain in *Cg*AA15a and *Cg*AA15b_cat were cloned in pET-26b(+) without the signal peptide and with a C-terminal strep-tag. Next, the

recombinant proteins were successfully produced in the periplasmic space of *Escherichia coli* (Fig. S4A). Protein identity was confirmed by LC-MS/MS analysis of the protein bands isolated from the SDS polyacrylamide denaturing gels (data not shown). The production of soluble recombinant proteins could only be achieved using M9 minimal medium without copper ion supplementation, followed by periplasmic extraction using osmotic shock and affinity chromatography (Fig. S4B and C). The mutant *Cg*AA15a^{F182Y} was generated using site direct mutagenesis, and protein production and purification were performed following the same methodology described above (Fig. S4D). The proteins were copper loaded to maintain their stability, and the unbound copper ions were washed out using size-exclusion chromatography.

After purification the LPMO activities were first evaluated using the 2,6-dimethoxyphenol (2,6-DMP) assay at pH 6.0 and 30 ºC (Fig. 3A) as described previously (48). The specific activities were 6.19 ± 0.71 U/g and 7.55 ± 0.36 U/g for *Cg*AA15a and *Cg*AA15b_cat, respectively. The specific activity for the mutant *Cg*AA15a^{F182Y} was lower than the native enzyme, resulting in a value of 4.75 ± 0.39 U/g. To determine the melting temperature (T_m) for *C. gestroi* AA15 LPMOs, a thermofluor analysis using SYPRO Orange dye was carried out. *Cg*AA15a showed a T_m of 52.7 ºC and after treatment with 5 mM EDTA for copper ion removal, the T_m value decreased to 49.3 ºC (ΔT_m of 3.4 ºC). *Cg*AA15b_cat showed a T_m of 59.2 ºC, and similarly to *Cg*AA15a, the T_m value decreased to 53.0 ºC (ΔT_m of 6.2 ºC) after treatment with EDTA. For *Cg*AA15a^{F182Y}, the T_m determined was 60.3 ºC, and after Cu removal the T_m value decreased to 46 ºC (ΔT_m of 14.3 ºC). The mutant *Cg*AA15a^{F182Y} showed a ΔT_m of 7.6 ºC compared to the wild type *Cg*AA15a, suggesting that the mutation increases the thermal stability of the enzyme. However, the T_m after stripping the copper ion from the protein was similar between both enzymes. These results indicate that the copper ion has a thermostabilizing effect on the protein structures as reported previously for other LPMOs from different families (11, 49).

Previous studies have shown that LPMOs in the presence of molecular oxygen and an external electron donor (without the substrate) can generate hydrogen peroxide (H_2O_2) (50). Taking advantage of this feature, the termite LPMOs were tested with different electron donors using the Amplex®Red assay for H_2O_2 quantification at pH 7.0 and 30 ºC. The highest H_2O_2 generation was achieved for *Cg*AA15a, *Cg*AA15b_cat and *Cg*AA15a^{F182Y} when ascorbic acid was used as the donor, generating more than 2 µM of H_2O_2 after 30 minutes of reaction for each enzyme (Fig. 4A-C). Interestingly, the mutant *Cg*AA15a^{F182Y} reached the same levels of H_2O_2 production as *Cg*AA15a, however at a lower velocity. The second highest production of H_2O_2 was achieved with L-cysteine, but after a few minutes of the assay it appeared that resofurin, the Amplex®Red reaction product, reacted with the electron donor, exhibiting a distorted curve.

Generation of H_2O_2 in the presence of gallic acid and pyrogallol, two phenolic compounds, was detected, but in low amounts for all three LPMOs tested.

2.5 *Cg*AA15a, *Cg*AA15b_cat and *Cg*AA15a^{F162Y} can oxidize and degrade chitin only.

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

*Cg*AA15a and *Cg*AA15b_cat were only able to oxidize chitin from Squid Pen in the presence of an external electron donor, with a C1-oxidation pattern and generation of C1-aldonic acids. Native peaks relative to the mono-sodiated unoxidized oligosaccharides as well as oxidized peaks corresponding to mono- or di-sodiated adducts of C1-aldonic acids, imparting +16 or +38 m/z respectively, were identified in the spectra. Moreover, peaks for the mono-sodiated lactone (−2) were also identified. The peak profiles in the mass spectra consisted of even-numbered oxidized oligosaccharides with different degrees of polymerization (DP), with the peaks corresponding to DP6 and DP8 at higher intensities than the odd-numbered DP5 and DP7. This pattern is in accordance with previous chitin-active LPMOs that exhibited the ability to breakdown crystalline structures (2), suggesting that *Cg*AA15a and *Cg*AA15b_cat are correlated with chitin remodelling or exoskeleton moulting in *C. gestroi*.

*Cg*AA15a^{F182Y} was also tested to oxidize a range of different polysaccharides; however, in its native form the mutant was only able to oxidize chitin from squid pen. The mutant showed the same peak masses as *Cg*AA15a and *Cg*AA15b_cat, presenting the C1-oxidation pattern and the even-numbered oxidized oligosaccharides. Finally, the negative control assays (substrate and ascorbic acid) did not generate any oxidized products (Fig. S5 A). Moreover, the three termite LPMOs were unable to produce oxidized peaks in the presence of 5 mM EDTA (Fig. S5 B-D), showing that the copper ion is essential for the oxidative reaction.

3. Discussion.

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

not involved in lignocelluloses degradation. The presence of a Tyr instead of a Phe is the major divergence among the catalytic sites of *Td*AA15a (cellulose/chitin-active) and *Td*AA15b (chitin-active) from the firebrat *T. domestica* (16), as well as for all termite AA15 LPMOs. According to our data, the mutant *Cg*AA15a^{F182Y} was unable to oxidize cellulose (PASC and Avicel).

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

For example, the change of Y213F in the axial position of the specific cellulose-active *Tf*AA10a from *Thermobifida fusca* only decreased its oxidative activity against cellulose and increased substrate binding (53), however the potential for chitin oxidation by *Tf*AA10a^{Y213F} was not tested. On the other hand, the change of F219Y in *Sc*LPMO10c from *Streptomyces coelicolor* abolished the oxidative activity on cellulose while the change to F219A only reduced it (51), although the authors also did not test the mutants for chitin oxidation. Despite the efforts to understand the role of aromatic residues in the axial position, it is found that native AA10s with Tyr in axial position, as well as fungal AA9s, are naturally cellulose-active enzymes at the least, while fungal AA11s, with Tyr in the active-site, are specific chitin-active LPMOs.

Protein engineering applied to specific cellulose-active *Sc*LPMO10C (with Phe in axial position) has recently shown that a broader set of aromatic and polar residues, which are located on the flat catalytic surface of the enzyme, can modulate substrate binding and specificity (47). According to the study (47), the authors performed the substitution of aromatic residues from the catalytic surface (previously reported to interact with cellulose (54)) with polar residues, which gave *Sc*LPMO10C the ability to oxidize chitin. Moreover, Zhou and co-authors presented an interesting analysis in the electrostatic potential binding surfaces of several AA10 enzymes, indicating that chitin-active AA10s have negatively charged binding surfaces whereas cellulose-active AA10s have uncharged or positively charged surfaces (52). This finding corroborates our data, where binding surfaces of *Cg*AA15a, *Cg*AA15b and *Td*AA15b (Figure S6) are negatively charged although the cellulose/chitin-active *Td*AA15a 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 *Cg*AA15a and
414 *Cg*AA15b 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 *Csp*AA15a, 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 *Csp*AA15a. 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, *Dm*AA15a
427 (*CG42749*) and *Dm*AA15b (*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 *Dm*AA15a and *Dm*AA15b, 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 *Cg*AA15a and *Cg*AA15b_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 *Cg*AA15a and *Cg*AA15b_cat from termites and *Td*AA15b 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

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

4. Material and Methods

4.1 Genome and transcriptome analyses.

The genome and transcriptome from worker and soldier castes of the lower termite *C. gestroi* were used as a database (34) to search for CDSs containing the Pfam domain 03067 (*LPMO_10 Superfamily*) using default parameters (e-value 1). Afterwards, each CDS had their number of transcripts in worker and soldier libraries counted and normalized (57).

4.2 Phylogenetic analyses.

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

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

4.3 In silico molecular modelling.

The three-dimensional homology models of *Cg*AA15a, *Cg*AA15b_cat and *Td*AA15b (without the signal peptide and Strep-tag sequences) were built using the I-TASSER (66) software with the LPMO *Td*AA15a from *Thermobia domestica* (PDB id: 5MSZ_A) as a reference structure (template). The C-scores were calculated for all enzyme models. The PDB visualizations were performed using the PyMOL™ software (1.7.4.5 Edu) and the electrostatic surface potential was calculated using the APBS plug-in (67) for PyMOL.

4.4 RNA extraction, cDNA transcription and PCR reactions.

The total RNA extraction from *C. gestroi* as well as the cDNA transcription was performed as described previously by Franco Cairo et. al., 2013 (26). PCR reactions for amplification of the full-length CDS of *CGAA15A* and the CDS for the LPMO domain of *CGAA15B* were performed in two rounds using the following oligonucleotides: 1st round PCRs - *Cg*AA15a_F_cDNA 5'-CATGCCCAG CTTATGGACCCTGTG3-'; *Cg*AA15a_R_cDNA 5'-GACGAGAGAAATGTCTGAGCAGCC-3'; *Cg*AA15b-cat_F_cDNA 5'CATGGACGTCTTATTGAACCTCCCTCC3'; *Cg*AA15b-cat_R_cDNA 5'-AACATCAGCACATGCTCGGAATTC-3'. For the 2nd round PCRs - *Cg*AA15a_F_pET26b_S 5'-**GCCCAGCCGGCGATGGCG**CATGCCCAGCTTATG-3'; *Cg*AA15a_R_pET26b_S 5'-CGGGTG GCTCCACGCGCTGACGAGAGAAATGTC-3'; *Cg*AA15b-cat_F_pET26b_S 5'-**GCCCAGCCGGCGATGGCG**CATGGACGTCTTATTG; *Cg*AA15b-cat_R_pET26b_S 5'-CGGGTGGCTCCACGCGCTAACATCAGCACATGC-3', where the bold letters are complementary to the pelB signal peptide from the pET-26b(+) vector and the underlined letters are complementary to the Strep-tag II sequence from previously modified pET-26b(+) (16). For linearization of the pET-26b-Strep-tag II vector, the following oligonucleotides were used: PF_Lin_26b_pelB 5'-AGCGCGTGGAGCCACCCGCAG-3 and PR_Lin_26b_Strep 5'-CGCCATCGCCGGCTGGGC-3'. For site-directed mutagenesis in *Cg*AA15a-F182Y, the following primers were used: *Cg*AA15a_mut1_F 5'-GCAAGAAACGtatCGTGGCTGTAGCGACATCTC-3'and *Cg*AA15a_mut1_R 5'-GGCCCACACCCTACGGCA-3', with the mutation as lower-case letters in the former primer. The PCR reactions for amplification of target CDSs were set up as reported by Franco Cairo et al., (26).

4.5 Cloning, recombinant protein production and purification.

For recombinant protein production and purification, the *C. gestroi* LPMOs were cloned in the pET-26b(+) vector, seeking periplasmic expression. Each CDS was inserted between the pelB leader signal peptide and the *Bam*HI restriction site, with a strep-tag inserted in the LPMO C-terminal followed by a stop codon, thus removing the 6xHis-tag. Cloning of *Cg*AA15a and *Cg*AA15b_cat was performed using the Gibson Assembly Master Mix according to manufacture instructions. The mutant *Cg*AA15a-F182Y was generated using the Q5 Mutagenesis kit from NEB according to manufacture instructions with one modification: the use of Phusion *Taq* DNA polymerase instead of Q5 DNA polymerase. Afterwards, the two constructions, as well as the mutated version of *Cg*AA15a-F182Y, were transformed in DH5-α for vector replication. Finally,

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

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

After protein expression, the LPMOs were extracted from the periplasmic space using the Osmotic Shock protocol (68), followed by affinity chromatography using Streptacitin resin. Next, the enzymes were concentrated and copper ion ($CuCl_2$) was loaded in excess of five-times, followed by size-exclusion chromatography in 25 mM MES buffer adjusted to pH 6 and containing 150 mM NaCl (69).

4.6 Functional characterizations of *C. gestroi* LPMOs.

The LPMO activity assays using 2,6-dimethoxyphenol (2,6-DMP) and hydrogen peroxide (H_2O_2) as co-substrates were performed as reported by Breslmayr et al. (2018) (48). The reactions were set up in total volumes of 200 µL, containing 100 µM H_2O_2, 10 mM 2,6-DMP and 1 µM of *C. gestroi* LPMOs in 100 mM ammonium acetate buffer pH 6 at 30 ºC. The reaction was initiated by adding the LPMO after 15 min of pre-incubation of the reagents and absorbance was monitored at 469 nm every 15 s during 5 min in the Epoch 2 Microplate Reader (BioTek). The LPMOs had their melting temperature (T_m) determined according to the protein thermal shift assay (ThermoFisher Scientific) (70). The LPMOs at 0.5 mg/mL (20 mM MES pH 6, 150 mM NaCl) were mixed with 1X SYPRO Orange in 30 µL reaction volumes and the fluorescence emission was monitored using a CFX384 Touch Real-Time PCR machine (BioRad) in a 384-well clear plate from 20 °C to 99 °C, in triplicate, where the proteins were incubated at each temperature for 2 min before taking the reading.

The Amplex®Red assay for H_2O_2 quantification was performed to determine molecules with potential to donate electrons for *C. gestroi* LPMOs. The reactions were set up as reported by Kittl et al. (2012) (50) in a final volume of 100 µL in clear microplate and in triplicate. The assays contained 50 µM of each electron donor (ascorbic acid, pyrogallol, gallic acid and L-cysteine), 50 µM Amplex®Red, 7 U/mL Horseradish peroxidase (HRP) and 1 µM LPMO as final concentrations. The assays started by adding the LPMOs and the resofurin absorbance was measured at 590 nm. The reactions were monitored in kinetic mode in the Epoch 2 Microplate

Reader (BioTek) during 35 min at 30 ºC with 3 s shaking before the readings. A standard curve was constructed using H_2O_2 concentrations in the range of 0.1 – 5 µM for its quantification.

4.7 Activity assays for substrate specificity determinations.

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

5. References

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

8. Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J., and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. *Journal of the American Chemical Society*. **135**, 6069–6077
9. Paradisi, A., Johnston, E. M., Tovborg, M., Nicoll, C. R., Ciano, L., Dowle, A., McMaster, J., Hancock, Y., Davies, G. J., and Walton, P. H. (2019) Formation of a Copper(II)–Tyrosyl Complex at the Active Site of Lytic Polysaccharide Monooxygenases Following Oxidation by H2O2. *J. Am. Chem. Soc.* **141**, 18585–18599
10. Loose, J. S. M., Arntzen, M. Ø., Bissaro, B., Ludwig, R., Eijsink, V. G. H., and Vaaje-Kolstad, G. (2018) Multipoint Precision Binding of Substrate Protects Lytic Polysaccharide Monooxygenases from Self-Destructive Off-Pathway Processes. *Biochemistry*. **57**, 4114–4124
11. Fowler, Claire. A., Sabbadin, F., Ciano, L., Hemsworth, G. R., Elias, L., Bruce, N., McQueen-Mason, S., Davies, G. J., and Walton, P. H. (2019) Discovery, activity and characterisation of an AA10 lytic polysaccharide oxygenase from the shipworm symbiont Teredinibacter turnerae. *Biotechnology for Biofuels*. **12**, 232
12. Munzone, A., El Kerdi, B., Fanuel, M., Rogniaux, H., Ropartz, D., Réglier, M., Royant, A., Simaan, A. J., and Decroos, C. (2020) Characterization of a bacterial copper-dependent lytic polysaccharide monooxygenase with an unusual second coordination sphere. *FEBS J.* 10.1111/febs.15203
13. Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D., and Marletta, M. A. (2014) Determinants of Regioselective Hydroxylation in the Fungal Polysaccharide Monooxygenases. *J. Am. Chem. Soc.* **136**, 562–565
14. Chen, J., Guo, X., Zhu, M., Chen, C., and Li, D. (2019) Polysaccharide monooxygenase-catalyzed oxidation of cellulose to glucuronic acid-containing cello-oligosaccharides. *Biotechnol Biofuels*. **12**, 42
15. Chen, C., Chen, J., Geng, Z., Wang, M., Liu, N., and Li, D. (2018) Regioselectivity of oxidation by a polysaccharide monooxygenase from Chaetomium thermophilum. *Biotechnology for Biofuels*. **11**, 155
16. Sabbadin, F., Hemsworth, G. R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R. D. S., Sweeney, S. T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A. A., Bates, R., Gomez, L. D., Simister, R., Davies, G. J., Walton, P. H., Bruce, N. C., and McQueen-Mason, S. J. (2018) An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun*. **9**, 756
17. Schyra, J., and Korb, J. (2019) Termite Communities along A Disturbance Gradient in a West African Savanna. *Insects*. **10**, 17
18. Casalla, R., and Korb, J. (2019) Termite diversity in Neotropical dry forests of Colombia and the potential role of rainfall in structuring termite diversity. *Biotropica*. **51**, 165–177
19. Griffiths, H. M., Ashton, L. A., Evans, T. A., Parr, C. L., and Eggleton, P. (2019) Termites can decompose more than half of deadwood in tropical rainforest. *Current Biology*. **29**, R118–R119
20. Ashton, L. A., Griffiths, H. M., Parr, C. L., Evans, T. A., Didham, R. K., Hasan, F., Teh, Y. A., Tin, H. S., Vairappan, C. S., and Eggleton, P. (2019) Termites mitigate the effects of drought in tropical rainforest. *Science*. **363**, 174–177
21. Fujita, A., Hojo, M., Aoyagi, T., Hayashi, Y., Arakawa, G., Tokuda, G., and Watanabe, H. (2010) Details of the digestive system in the midgut of Coptotermes formosanus Shiraki. *Journal of Wood Science*. **56**, 222–226
22. Tokuda, G., Watanabe, H., and Lo, N. (2007) Does correlation of cellulase gene expression and cellulolytic activity in the gut of termite suggest synergistic collaboration of cellulases? *Gene*. **401**, 131–134
23. Brune, A., and Ohkuma, M. (2011) Role of the Termite Gut Microbiota in Symbiotic Digestion. in *Biology of Termites: a Modern Synthesis* (Bignell, D. E., Roisin, Y., and Lo, N. eds), pp. 439–475, Springer Netherlands
24. Ni, J., and Tokuda, G. (2013) Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnology Advances*. **31**, 838–850
25. Scharf, M. E., Karl, Z. J., Sethi, A., and Boucias, D. G. (2011) Multiple levels of synergistic collaboration in termite lignocellulose digestion. *PLoS One*. **6**, e21709

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

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

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

6. Figure legends

Figure 1. Phylogenetic analysis with termites LPMOs. A) Radial phylogram of LPMO families for the classification of termite enzymes. All AA families were clustered with high bootstraps (> 0.94). B) Detail of the AA15 branch from radial phylogram, showing all termite LPMOs. *C. gestroi* enzymes are highlighted with red circles and *T. domestica* with orange asterisks. The LPMOs AA15a from *Z. nevadensis, C. secundus, M. natalensis* and *Cortaritermes* sp. are highlighted with blue rhombus, squares, asterisks and circles, respectively. The termite AA15s are grouped in three different clades (light grey boxes), each with a different protein architecture: 1, 2 and 3 in blue boxes. SP= Signal peptide, TM=transmembrane. CT= cytosolic. The sequences from *Mn*AA15b, *Cg*AA15c, *Cg*AA15d, *Td*AA15a and *Td*AA15b are partial sequences from mRNA sequencing.

Figure 2. The structural hallmarks of *Cg*AA15s LPMOs. A) Overall homology models for *Cg*AA15a (I) and *Cg*AA15b_cat (II) in comparison with *Td*AA15a (PDB code 5MSZ) (III). The copper ion is in gold. **B)** Detail of the catalytic site showing the coordinated copper-His-brace and the non-coordinated Ala and Phe residues for *Cg*AA15a (I), *Cg*AA15b_cat (II) and *Td*AA15a (III). **C)** The electrostatic surface potential of *Cg*AA15a (I), *Cg*AA15b_cat (II) and *Td*AA15a (III), the putative binding site in yellow ellipses and the copper in mocha circles.

Figure 3. The specific activities and the apparent melting temperature of *C. gestroi* AA15s. A) Determination of specific activity of *Cg*AA15a, *Cg*AA15b_cat and *Cg*AA15a ^{F182Y} using the 2,6-DMP assay. Thermal shift curves of *Cg*AA15a **(B)**, *Cg*AA15b_cat **(C)** and *Cg*AA15a^{F182Y} **(D)** and the melting temperature (T_m) of the copper ion loaded version of the enzymes (right curves) and the apo form after EDTA treatment (left curves).

Figure 4. Determination *in-vitro* of the optimal electron donor for *C. gestroi* AA15s. The Amplex®Red assay was applied for the quantification of H_2O_2 and to determine the optimal electron donor for the LPMOs. Four molecules were used as donors: L-cysteine (blue rhombus), Ascorbic acid (yellow times), Pyrogallol (grey triangles) and Gallic Acid (orange boxes). Kinetic measurements for **(A)** *Cg*AA15a, **(B)** *Cg*AA15b_cat and **(C)** *Cg*AA15a^{F182Y} were performed for 35 min.

Figure 5. Substrate specificity of *C. gestroi* AA15s. MALDI-TOF MS spectrum of reaction products after the incubation of 1 μM *Cg*AA15a **(A)**, *Cg*AA15b_cat and **(C)** *Cg*AA15a^{F182Y} **(E)** with 1% chitin from squid pen in the presence of 2 mM ascorbic acid as the electron donor. Detailed view of the expanded spectrum for DP6, DP8 and DP6 of the reactions applying *Cg*AA15a **(B)**, *Cg*AA15b_cat **(D)** and *Cg*AA15a^{F182Y} **(F),** respectively. Native species are marked in black and the oxidized species in red.

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

Figure S2. AA15s transcript counts in castes of *C. gestroi* and CDS amplifications. A) Normalized number of transcripts of *CGAA15s* between worker (red) and soldier (light red) castes of the lower termite *C. gestroi*. **B)** 1% agarose gels showing the amplification products for AA15 CDSs from cDNA of the *C. gestroi* work caste. L = Ladder, 1-3 = *CGAA15A*, 4 and 5 = *CGAA15b_cat* and 6 = negative control.

Figure S3. Models for the flat catalytic surface of *C. gestroi* AA15s. A) Aromatic residues (red) exposed on the flat catalytic surface of *Cg*AA15a (I), for *Cg*AA15b-cat (II) and *Td*AA15a PDB id 5MSZ (III). **B)** The negatively charged residues (orange) for the same LPMOs and **C)** Superposition of aromatic and negatively charged residues of the three LPMOs.

Figure S4. Recombinant protein production and purification. A) Schematic view of the expression vector (pET-26b) and the protein construction. The protein was cloned in frame with the pelB (signal peptide) and also with a streptavidin tag (S-Tag) for protein purification. **B)** 12% SDS-PAGE gel of protein extraction and purification for *Cg*AA15a: M = ladder, 1 = Osmotic shock fraction, 2 = column flowthrough, 3 = wash fraction, 4-7 = Elution fractions with 3 mM destiobiothin. **C)** 12% SDS-PAGE gel of protein extraction and purification for *Cg*AA15b_cat: M = ladder, 1 = Osmotic shock fraction, 2 = column flowthrough, 3 = wash fraction, 4 and 5 = Elution fractions with 3 mM destiobiothin. **D)** 12% SDS-PAGE gel of protein extraction and purification for *Cg*AA15a^{F182Y}: M = ladder, 1 = Osmotic shock fraction, 2 = column flowthrough, 3 = wash fraction, 4-6 = Elution fractions with 3 mM d-Desthiobiotin.

Figure S5. Substrate specificity of *C. gestroi* LPMOs. MALDI-TOF MS spectrum of control reaction products. **A)** Spectra for 1% chitin from squid pen and 2 mM ascorbic acid. **B)** Spectra for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1 μM of *Cg*AA15a. **C)** Spectra for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1 μM of *Cg*AA15b_cat. **D)** Spectra for 1% chitin from squid pen, 2 mM ascorbic acid, 1 mM EDTA and 1 μM of *Cg*AA15a^{F182Y}. A range of different polysaccharides were tested within the same conditions and controls (data not shown).

Figure S6. Structural hallmarks of *Td*AA15b*. A)* Overall model for the structure based on*Td*AA15a PDB id 5MSZ (III). The calculated C-Score was 1.71 with estimated RMSD of 2.0±1.6Å. The copper metal is in gold and the catalytic residues His1, His91 and Phe184 are in blue. **B)** The electrostatic surface potential of *Td*AA15b, showing the protrusion in black circles, the putative binding site in yellow ellipses and copper in moca circles.

7. Acknowledgements

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

8. Authors contributions

JPLFC designed and performed the experiments and analyzed the data as well as draft the manuscript. AP performed experiments and analyzed the data. TAG designed experiments and discussed the data. MVR performed phylogenetic analysis and analyzed the data. LSM performed the bioinformatic analyses. DAC designed experiments, analyzed the data and draft the manuscript. LC performed experiments, analyzed the data and draft the manuscript. CRFT designed and performed experiments. WG performed experiments and wrote the draft manuscript. MFC designed and performed the bioinformatic analyses. GJD analyzed the data and revised the final version of the manuscript. CF designed, economically supported the experiments and revised the final version of the manuscript. PHW designed experiments, analyzed the data, economically supported the work and drafted the manuscript. FMS designed experiments, analyzed the data, economically supported the work and revised the manuscript.

9. Conflict of interest

The authors declare no conflict of interest**.**