Characterization of a bacterial copper-dependent lytic polysaccharide monooxygenase with an unusual second coordination sphere

Alessia Munzone, Bilal El Kerdi, Mathieu Fanuel, Hélène Rogniaux, David Ropartz, Marius Réglier, Antoine Royant, A. Jalila Simaan, Christophe Decroos

Characterization of a bacterial copper-dependent lytic polysaccharide monooxygenase with an unusual second coordination sphere

Alessia Munzone,[a] Bilal El Kerdi,[a] Mathieu Fanuel,[b] Hélène Rogniaux,[b] David Ropartz,[b] Marius Réglier,[a] Antoine Royant,[c,d] A. Jalila Simaan,[a] and Christophe Decroos[a] *

[a] Aix Marseille Univ, CNRS, Centrale Marseille, iSm2, Marseille, France

[b] INRA, UR1268 Biopolymers Interactions Assemblies, La Géraudière B.P. 71627, F-44316 Nantes, France

[c] Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), F-38000 Grenoble, France

[d] European Synchrotron Radiation Facility, F-38043 Grenoble, France

* Corresponding author:

Aix Marseille Univ, CNRS, Centrale Marseille, iSm2, Marseille, France

E-mail address: christophe.decroos@univ-amu.fr

Running title

Characterization of a LPMO from *P. luminescens*

Keywords

Lytic polysaccharide monooxygenase, copper metalloenzyme, chitin oxidation, *Photorhabdus luminescens*, X-ray crystallography

Abbreviations

AA, auxiliary activity; CAZyme, carbohydrate-active enzyme; CID, collision-induced dissociation; DP, degree of polymerization; EPR, electron paramagnetic resonance, ESI, electrospray ionization; IPTG, isopropyl-β-D-thiogalactopyranoside; LB; lysogeny broth; LPMO, lytic polysaccharide monooxygenase; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; MS, mass spectrometry, MWCO, molecular weight cut-off; *Pl*AA10, LPMO from *Photorhabdus luminescens*; rmsd, root-mean-square deviation; TSB, tryptic soy broth

Enzymes

*Pl*AA10, 1.14.99.53

Database

Structural data (atomic coordinates and structure factors) reported for *Pl*AA10 are available in the Protein Data Bank under accession number 6T5Z.

Abstract

Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent enzymes involved in the degradation of recalcitrant polysaccharides such as cellulose or chitin. LPMOs act in synergy with glycoside hydrolases such as cellulases and chitinases by oxidatively cleaving a number of glycosidic bonds at the surface of their crystalline substrate(s). Besides their role in biomass degradation, some bacterial LPMOs have been found to be virulence factors in some human and insect pathogens. *Photorhabdus luminescens* is a nematode symbiont bacterium that is pathogenic to a wide range of insects. A single gene encoding a LPMO is found in its genome. In this work, we report the characterization of this LPMO, referred to as *Pl*AA10. Surprisingly, *Pl*AA10 lacks the conserved alanine residue (substituted by an isoleucine) found in the second coordination sphere of the copper active site in bacterial LPMOs. *Pl*AA10 was found to be catalytically active on both α- and β-chitin, and exhibits a C1-oxidation regiospecificity, similarly to other chitin-active LPMOs. The 1.6 Å X-ray crystal structure confirmed that *Pl*AA10 adopts the canonical immunoglobulin-like fold typical for LPMOs. The geometry of the copper active site is not affected by the nearby isoleucine, as also supported by electron paramagnetic resonance. Nevertheless, the bulkier side chain of isoleucine protrudes from the substrate-binding surface. A bioinformatic study on putative bacterial LPMOs unveiled that they exhibit some variability at the conserved active site alanine position with a substitution in about 15% of all sequences analyzed.

Introduction

Cellulose and chitin are two most abundant biopolymers in nature, constituted of tightly packed polysaccharide chains of β-(1,4)-linked glucose and *N*-acetylglucosamine units, respectively. Thanks to their crystallinity conferring strength and recalcitrance, these biopolymers are found as major structural components of plant cell wall for cellulose [1], and of the exoskeleton of arthropods such as crustaceans and insects, as well as of the fungal and yeast cell walls for chitin [2]. Synergistic endo- and exo-acting glycoside hydrolases – including cellulases and chitinases – secreted by cellulose- and chitin degrading microorganisms (bacteria and fungi) were long considered to be the only enzymatic partners involved in recalcitrant polysaccharide breakdown. However, Reese *et al*. suggested that an additional non-hydrolytic factor could reduce cellulose recalcitrance prior to enzymatic hydrolysis by fungal cellulases [3]. This extra component remained elusive until a decade ago with the discovery of new proteins able to enhance the hydrolytic activity of chitinases and cellulases [4,5]. Notably, these enzymes called lytic polysaccharide monooxygenases (LPMOs) were later found to introduce chain breaks into the polysaccharide *via* an oxidative mechanism [6] relying on a mononuclear copper ion in their active site [7,8]. This disrupts locally the substrate crystallinity on its surface, which facilitates the action of glycoside hydrolases working in concert with LPMOs and thus boosts the degradation of recalcitrant polysaccharides. LPMOs are now classified as auxiliary activities (AA) in the carbohydrate-active enzymes database (CAZy database, http://www.cazy.org) due to their redox function [9,10], and are divided into seven families: AA9 (formerly GH61) [7], AA10 (formerly CBM33) [6], AA11 [11], AA13 [12,13], AA14 [14], AA15 [15], and AA16 [16].

LPMOs catalyze the oxidative cleavage of glycosidic bonds in recalcitrant polysaccharides (Figure 1a). Notably, they hydroxylate the strong C-H bond either at position C1 or C4. The resulting unstable intermediate further decomposes leading to different products. C1 oxidation

produces aldonolactones that are hydrolyzed to their corresponding aldonic acids. C4 oxidation generates ketoaldose products that are in equilibrium with their hydrated *gem*-diol forms. Generally, chitin-active LPMOs regiospecifically oxidize their substrate at position C1 [6,11,15,17-19]. On the contrary, some LPMOs active on cellulose were found to exclusively oxidize at position C1 or C4, while others can generate oxidation product originating from both C1 and C4 oxidation [8,16,17,20-26]. As previously mentioned, LPMOs reactivity depends on a catalytically obligatory copper ion at their active site. The Cu(II) resting state of LPMOs needs to be reduced to Cu(I) to initiate catalysis. Electrons are supplied by either proteic partners or by small molecule reductants [27]. Upon Cu(II) reduction to Cu(I), LPMOs can activate their co-substrate that can be either dioxygen [6,7,21] or hydrogen peroxide [27,28]. While the active species remains to be characterized, theoretical calculations identified a Cu(II)-oxyl as the intermediate responsible for substrate hydroxylation *via* hydrogen abstraction and oxygen-rebound mechanism [29,30].

LPMOs all adopt a common distorted immunoglobulin-like β-sandwich fold, although they share very low sequence identity between different families [31]. They exhibit a relatively flat surface where the crystalline substrate binds [32]. The copper active site of LPMOs is located on this surface so that it is exposed to the substrate when bound. Two strictly conserved histidine residues coordinate the Cu(II) ion in a unique motif referred to as the histidine brace (Figures 1b and 1c), in which the *N*-terminal histidine serves as bidentate ligand (coordination by both the *N*-terminal amine and the imidazole side chain Nδ1) [7,31]. Two water molecules complete the Cu(II) first coordination sphere (Figure 1b). Due to X-ray photoreduction, most of the LPMOs structures are in the reduced Cu(I) state featuring a T-shaped geometry (no water molecule coordinated, Figure 1c) [33,34]. In fungal LPMOs, the Nε2 atom of the *N*-terminal histidine is posttranslationally methylated [7,13,31]. The histidine brace motif is fairly rare as it has only been observed so far in a handful of copper-metalloproteins besides LPMOs,

including the PmoB subunit of the copper-dependent particulate methane monooxygenase [35,36], and the bacterial copper-transport proteins CopC [37,38] and PmoF [39]. In LPMOs, a conserved tyrosine or phenylalanine residue is found in axial position of the copper coordination sphere [31]. Interestingly, a highly conserved second coordination sphere alanine residue in AA10 LPMOs (mostly bacterial) was tentatively hypothesized to provide some steric hindrance close to the solvent-exposed axial position of the metal center and therefore orient the binding and reactivity of dioxygen or hydrogen peroxide in the vacant equatorial position of the Cu(I) [33].

Apart from their role in polysaccharidic biomass breakdown, AA10 LPMOs are also involved in several physiological processes including endosymbiosis (host-microbe interactions) and pathogenicity (virulence factors) [40]. *Photorhabdus luminescens* is a Gram-negative broad-spectrum insect pathogenic bacterium that lives in symbiosis with the *Heterorhabditis bacteriophora* nematode. *Photorhabdus luminescens* is released in the insect larva hemocoel upon nematode invasion. The insect is rapidly killed after the production of toxins and virulence factors by the entomopathogenic bacteria. The insect cadaver is further bioconverted by enzymes secreted by the bacteria, supporting the growth of both the nematodes and *Photorhabdus luminescens*. Genes encoding for secreted enzymes involved in host chitin degradation (several chitinases and a single LPMO) were identified in *Photorhabdus luminescens* [41], as expected for insect pathogens [40,42,43]. In this work, we characterized the single LPMO from *Photorhabdus luminescens*, referred to as *Pl*AA10. Sequence analysis revealed that *Pl*AA10 lacks the conserved alanine residue found in the active site of AA10 LPMOs (Figure 1b). Instead, this second coordination sphere residue is substituted by a bulkier isoleucine in *Pl*AA10. This study reports the biophysical, biochemical, and structural characterization of this AA10 LPMO featuring an unusual second coordination sphere. Overall, this study provides new insights into the diversity observed among bacterial LPMOs.

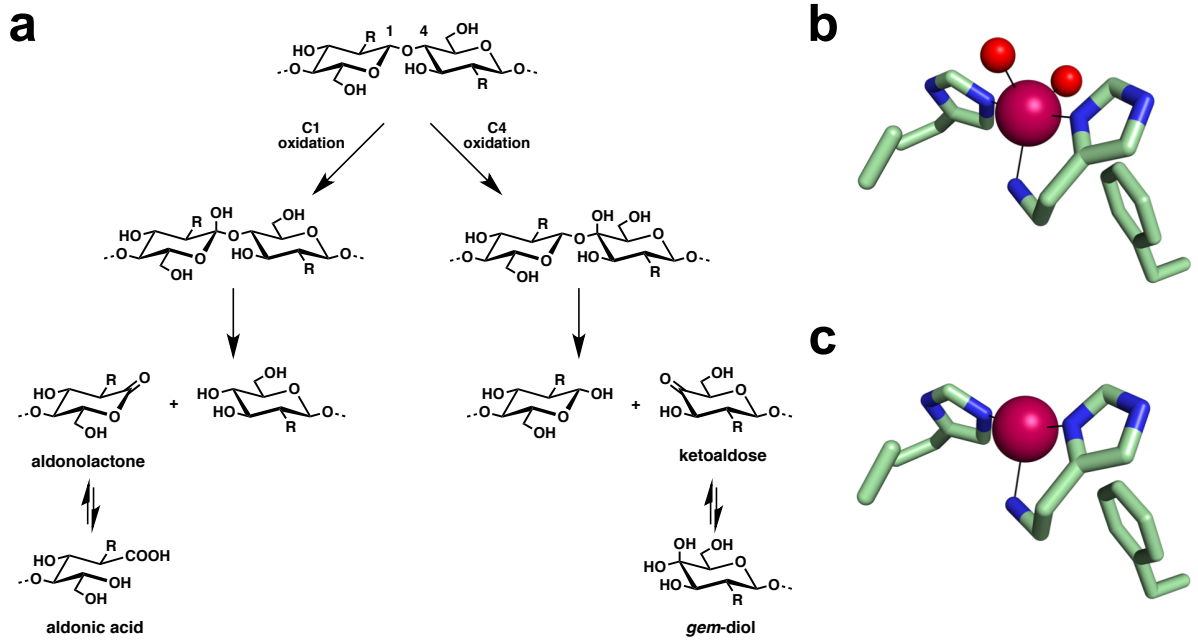

Figure 1. Reaction catalyzed by LPMOs and active site of bacterial AA10 chitin-active LPMOs. (a) Glycosidic bond oxidation by LPMOs either at position C1 or C4 leading to aldonolactone or ketoaldose products that are in equilibrium in aqueous medium with the corresponding aldonic acid or *gem*-diol, respectively. C1 and / or C4 oxidation by LPMOs has been reported for cellulose (R = OH), whereas chitin (R = NHAc) is generally oxidized at position C1. (b) Resting state Cu(II)- and (c) reduced Cu(I)-active site of the chitin-active *Ba*AA10 from *Bacillus amyloliquefaciens* (PDB accession codes 5IJU and 2YOX). Atomic color codes as follows: C = light green, N = blue, O = red, Cu = magenta sphere.

Results and Discussion

The putative chitin-active LPMO *Pl*AA10 harbors an isoleucine in place of the conserved active-site alanine residue

The genome from the entomopathogenic bacteria *Photorhabdus luminescens* contains over 70 genes encoding carbohydrate-active enzymes (CAZymes). Among this CAZymes, several chitinases and a single LPMO are predicted to be essential for insect larva chitin degradation during invasion, and later during bioconversion of its cadaver [41-43]. Accordingly, chitinase activity has already been reported in culture supernatants of *Photorhabdus luminescens* and other related insect pathogenic bacteria [44]. Nevertheless, no chitinolytic enzyme (chitinase or LPMO) from *Photorhabdus luminescens* has been characterized so far. The single LPMO from *Photorhabdus luminescens*, referred to as *Pl*AA10, is predicted to be secreted as a single-domain enzyme (catalytic domain: residues 26-199) after cleavage of its signal peptide (residues 1-25). From the phylogenetic analysis we performed (Figure 2a), *Pl*AA10 is implied to catalyze the oxidative cleavage of the glycosidic bond of chitin as it clusters with other characterized chitin-active LPMOs. In particular, *Pl*AA10 exhibits 52% sequence identity with the well-characterized chitin-active LPMO *Sm*AA10 (also known as CBP21) from the chitinolytic bacterium *Serratia marcescens* [6,32,45,46]. The two strictly conserved histidine residues required for copper binding are present in *Pl*AA10 – H26 and H115 defining the histidine brace – as well as the conserved second coordination sphere axial phenylalanine F188 (Figure 1b). Interestingly, multiple sequence alignment revealed that no conserved alanine is found within the active site of *Pl*AA10, but an isoleucine residue instead (I113) (Figure 2b). To preclude any possible previous sequencing error, a 1 kb fragment including the gene coding for *Pl*AA10 (*plu2352*) was amplified from *Photorhabdus luminescens* genomic DNA and further sequenced, which confirmed the presence of an unusual alanine-to-isoleucine substitution in *Pl*AA10. To date, all characterized chitin-active AA10 LPMOs possess the conserved active-

site alanine residue. We were therefore interested in studying the structural and functional properties of this enzyme.

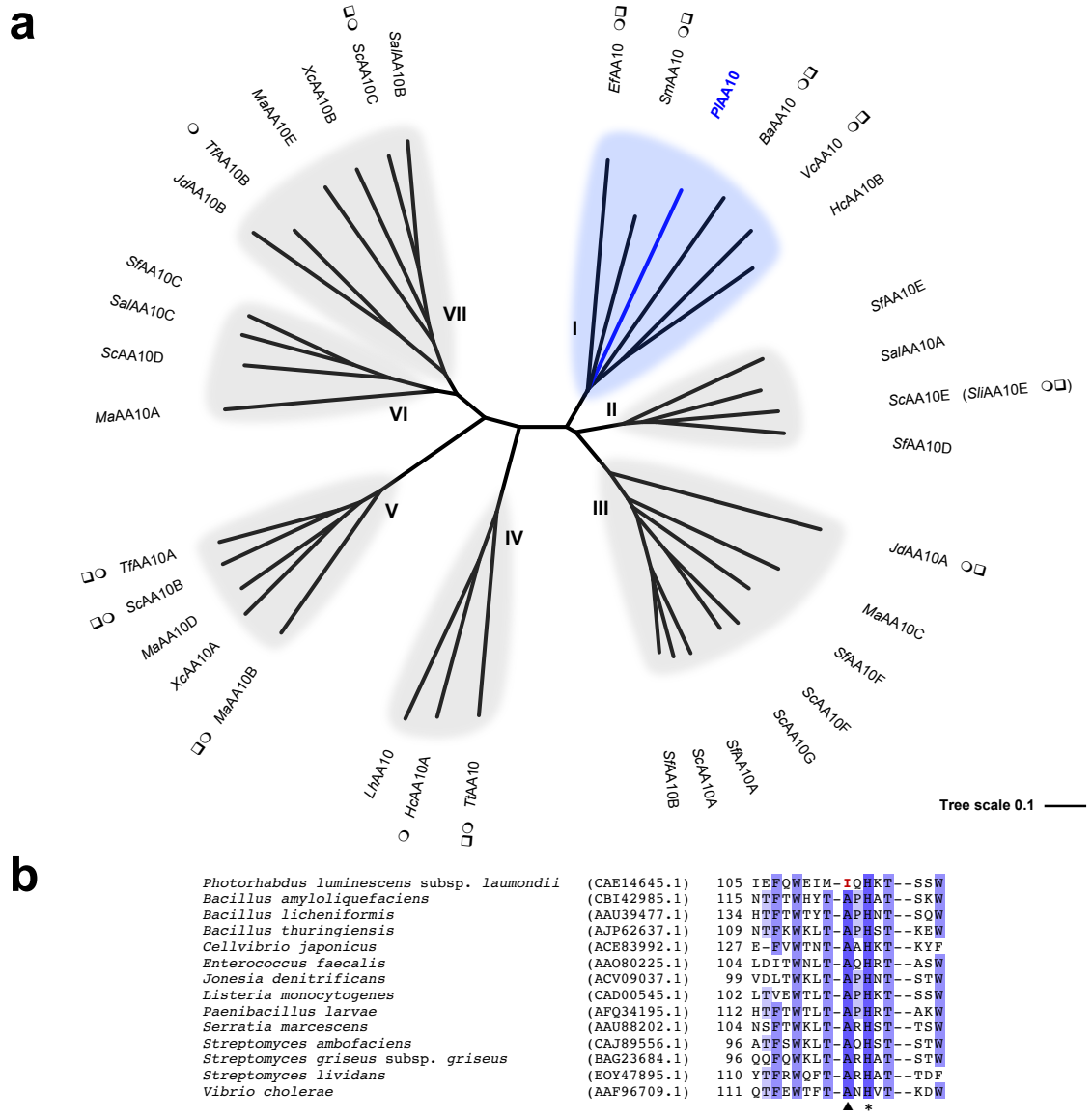

Figure 2. Sequence analysis of *Pl*AA10 and other bacterial AA10 LPMOs. (a) Unrooted phylogenetic tree of bacterial AA10 LPMOs. Characterized members of subclades I, II, and III are active on chitin. Chitin-active LPMOs from subclade III differ from others by the size of their catalytic domain that is smaller (~ 15 kDa *vs* ~ 20 kDa for members of subclades I and II) [18]. Characterized members of subclades IV and VII are active on cellulose with a C1-regiospecificity of oxidation. Cellulose oxidation at both position C1 and C4 were reported for

members of subclade V. Interestingly, some members of this subclade are also active on chitin. Phylogenetic analyses have shown that cellulose-active AA10 LPMOs derive from chitin-active AA10 LPMOs. It has been proposed that the chitin activity of subclade V members could be a residual ancestral activity [47]. So far, no member of subclade VI has been biochemically characterized. Nevertheless, the membrane-bound *Sc*AA10D from *Streptomyces coelicolor* A3(2) was shown to bind crystalline cellulose [48]. Some biochemically (❍) and / or structurally (❑) characterized AA10 LPMOs are highlighted. The origin and GenBank accession number of putative and characterized bacterial LPMOs included in this unrooted phylogenetic tree are given in Table S1. The multiple sequence alignment was generated with Clustal Omega and the unrooted phylogenetic tree drawn with Interactive Tree of Life (iTOL). (b) Multiple sequence alignment of *Pl*AA10 with previously characterized chitin-active AA10 LPMOs featuring the strictly conserved second histidine active site (∗) and the conserved second coordination sphere alanine residue observed for all characterized chitin-active AA10 (▲). The multiple sequence alignment was generated with Clustal Omega and visualized with Jalview.

***Pl*AA10 binds copper similarly to other chitin-active LPMOs**

In their native hosts, LPMOs are produced as secreted enzymes after cleavage of their signal peptide. This leads to a mature protein with a *N*-terminal histidine required for Cu(II)-binding (histidine brace) and therefore catalysis. Bacterial LPMOs are usually overexpressed in the periplasm of *Escherichia coli* [33,45,49]. However, other strategies for AA10 LPMO production in the cytoplasm of *E.* coli with a cleavable *N*-terminal tag [17,19,20] or as secreted protein using Gram-positive bacteria (*Brevibacillus choshinensis* SP3 or *Bacillus subtilis*) [18,50], cyanobacteria (*Synechococcus elongatus* UTEX 2973) [51], or yeast (*Pichia pastoris*) [52] as expression host have also been successfully developed. Using *E.* coli as recombinant expression system, we chose the *N*-terminal pelB signal peptide to direct *Pl*AA10 production to the periplasm. Accordingly, the codon-optimized gene coding for *Pl*AA10 was cloned directly in frame with the pelB signal peptide in the vector pET-22b(+) to ensure proper *N*-terminal processing. *Pl*AA10 was overexpressed in the periplasm of *E. coli* BL21(DE3) after an overnight induction with IPTG at low temperature (16°C). The periplasmic protein fraction was released by cold osmotic shock and apo-*Pl*AA10 was purified by affinity chromatography over chitin resin. Overall, the average yield of pure protein (as determined SDS-PAGE) was about 8 mg per liter of culture.

The intrinsic fluorescence of tryptophan residues was used to probe the Cu(II) binding to *Pl*AA10. Upon titration of the apo-enzyme with Cu(II), tryptophan fluorescence of *Pl*AA10 is linearly quenched (Figure S1a) as already observed for *Sli*AA10E, a chitin-active LPMO from *Streptomyces lividans* [53]. A break point is reached at about 1.0-1.2 equivalent of Cu(II) after which the fluorescence remains constant (Figure S1b). These results are indicative of a 1:1 stoichiometry for the formation of the Cu(II)-*Pl*AA10 complex, and also reveal a high affinity of the enzyme for its Cu(II) cofactor. Since the titration was carried out with 1 μM of apo-enzyme, it implies a dissociation constant for Cu(II) at least ten times less than the protein

concentration *i.e.* below 100 nM as expected for a bacterial LPMO. For instance, K_d values for Cu(II) in the low- to mid-nanomolar range were determined for only three chitin-active AA10 LPMOs: 55 ± 8 nM, 62.5 ± 0.2 nM, and from 6 to 80 nM (from pH 5 to 7) for *Sm*AA10, *Sli*AA10E, and *Ba*AA10, respectively [19,32,33,53]. Although we did not measure the exact dissociation constant for Cu(II) in the case of *Pl*AA10, the presence of a bulkier isoleucine in the second coordination sphere does not drastically impair copper binding. When needed, the holoenzyme was prepared by incubation of the apoenzyme with Cu(II) (5 equivalents) and thorough removal of the excess by two consecutive buffer exchanges.

The environment of the Cu(II) ion in *Pl*AA10 was further studied by continuous wave EPR. The X-band EPR spectrum of Cu(II)-*Pl*AA10 recorded at 120K in the presence of 10% (v/v) glycerol (Figure 3) is similar in shape to that of other chitin-active LPMOs [17,19,33,53]. In particular, the rhombicity of the spectrum ($g_z \neq g_y \neq g_x$) is indicative of a distorted geometry intermediate between square pyramidal and trigonal bipyramidal coordination geometries. In the parallel region, the spin Hamiltonian parameters were determined: $g_z = 2.260$ and $|A_z| = 355$ MHz (112 G). *Pl*AA10 clusters with other chitin-active AA10 LPMOs and more closely to *Sm*AA10 using the determined g_z and $|A_z|$ parameters on a Peisach-Blumberg plot [54,55]. As previously noted, the EPR signal of LPMOs can be affected by the presence or absence of glycerol as glassing agent [15,55]. Accordingly, the EPR signal of *Pl*AA10 is more complex in the absence of glycerol and clearly shows the contribution of two species (Figure S2). This complex EPR spectrum could be fitted with a major species (~ 80%) corresponding to that observed in the presence of glycerol (as in Figure 3) and a minor species (~ 20%). The latter exhibits a slightly lower g_z value (2.235) but a significantly higher $|A_z|$ parameter (547 MHz, 175G) than the major species. However, further spectroscopic analysis such as Q-band EPR would be required to unambiguously assign the full spin Hamiltonian parameters of the Cu(II) ion for both species. The presence of several EPR signals was already reported for another

chitin-active AA10 LPMO, *Sli*AA10E. The second species also exhibits a significantly higher $|A_z|$ parameter and was proposed to originate from the discoordination of the *N*-terminal amine [53]. Heterogeneity in the histidine brace by ligand discoordination could occur during freezing of the sample in the absence of glassing agent [55].

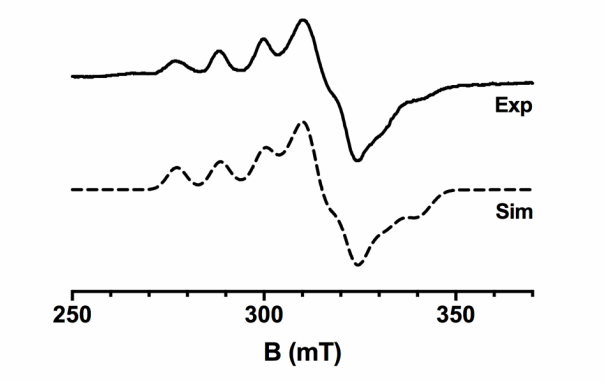

Figure 3. EPR spectroscopy of *Pl*AA10. X-band EPR spectrum (plain line) of Cu(II)-*Pl*AA10 (50 mM MES set at pH 6.5) in the presence of 10% (v/v) glycerol) recorded at 120 K. The spin Hamiltonian parameters used for the simulated spectrum (dashed line) are as follows: 2.025, 2.103, and 2.260 for g_x, g_y, and g_z respectively; and 220 MHz (78 G), 95 MHz (32 G), and 355 MHz (112 G) for $|A_x|$, $|A_y|$, and $|A_z|$, respectively.

*Pl*AA10 is a C1-oxidizing LPMO active on both α- and β-chitin

*Pl*AA10 is a putative chitin-active LPMO according to multiple sequence alignment with other characterized AA10 LPMOs and corresponding phylogenetic analysis (Figure 2). Additionally, chitin-binding by *Pl*AA10 was verified as it was purified by affinity chromatography over chitin resin. Therefore, we evaluated *Pl*AA10 activity towards both α- and β-chitin. These two chitin allomorphs differ in the relative arrangement of polymeric chains: anti-parallel for α-chitin and parallel for β -chitin. *Pl*AA10 was incubated with either α- or β-chitin in the presence of ascorbic acid as reductant under aerobic conditions *i.e.* with dioxygen as co-substrate. *Pl*AA10 exhibits activity towards both α- and β-chitin. Accordingly, MALDI-TOF MS analysis of the reaction supernatant led to the detection of soluble oxidized chitooligosaccharides from degree of polymerization (DP) 2 to 6 and from 3 to 10 for α- and β-chitin, respectively (Figure 4a). A similar product distribution is obtained whether the reaction was run in an unbuffered solution (Figure 4a) or at different pH values (5.0, 6.5, and 8.0, Figure S3 and S4). For each DP, three peaks are observed in the MS spectrum (Figure 4b). The first isotopic pattern corresponds to the aldonolactone C1-oxidization product (as Na^+ adduct), and the second and third to the aldonic acid (as Na^+ adduct or deprotonated with two Na^+, respectively) resulting from the hydrolysis of the aldonolactone (Figure 4c). Further MS/MS confirmed the identification of the aldonolactone and its related aldonic acid (Figure S5), excluding any product resulting from C4 oxidation. Therefore, *Pl*AA10 is a C1-oxidizing LPMO active on both α- and β-chitin. The C1-oxidation regiospecificity is a hallmark of chitin-active LPMOs [6,11,15,17-19]. According to MS data, oxidized products with even DP are preferred over odd DP for each substrate (Figure 4a) [6]. This preference derives from the structure of chitin with a two-fold screw axis along the polymer chain, leaving one every two glycosidic bonds of a surface polysaccharide chain buried inside the crystalline chitin and consequently inaccessible to the enzyme.

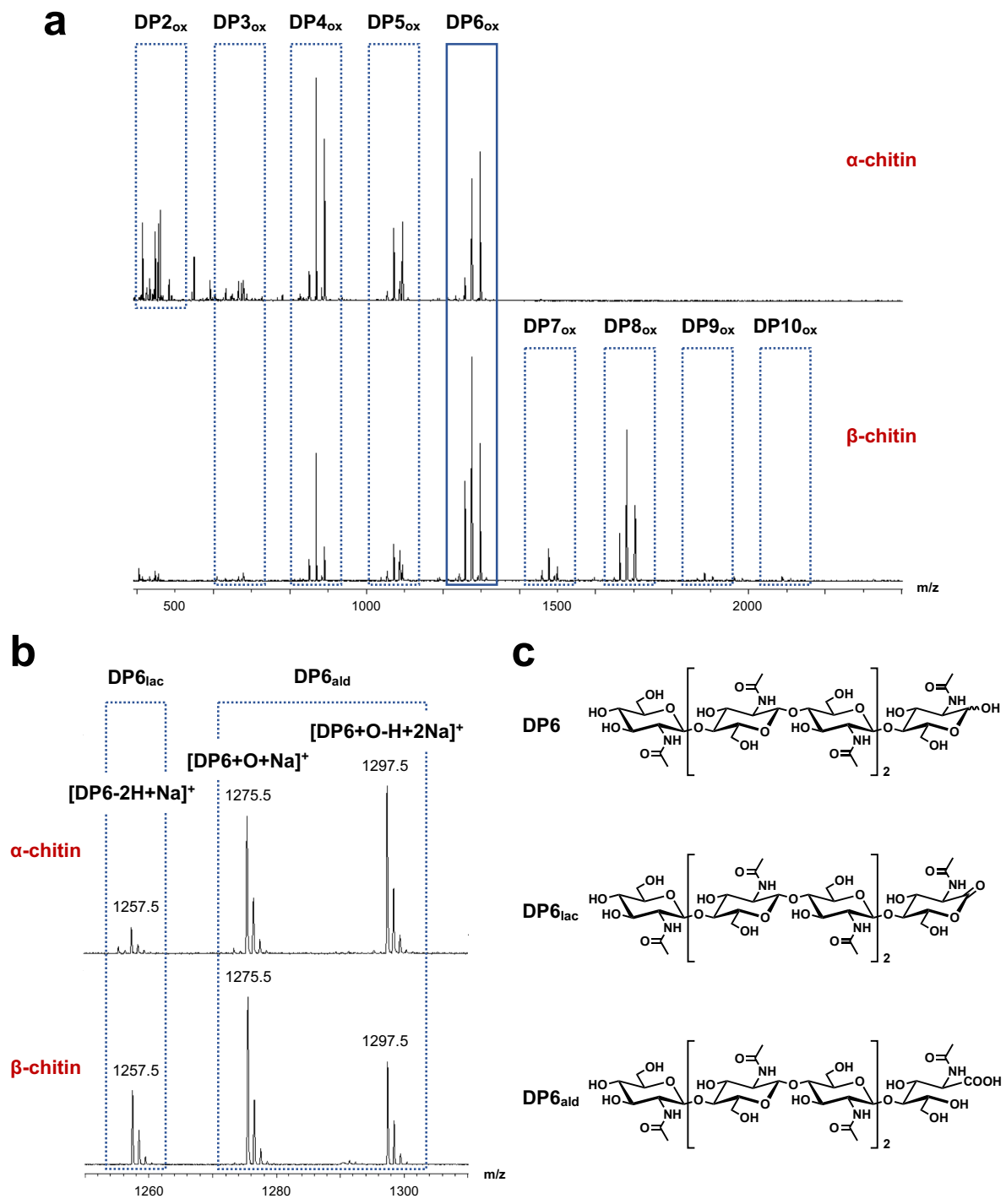

Figure 4. Activity of *Pl*AA10 towards α- and β-chitin. (a) MALDI-TOF mass spectrometry analysis of reaction supernatant upon reaction of *Pl*AA10 on α- and β-chitin in the presence of dioxygen as co-substrate and ascorbic acid as reductant in an unbuffered solution (deionized water). This analysis reveals the formation of oxidized products with varying degree of polymerization (DP) up to DP10 depending on the substrate. (b) Zoom on

the mass peaks of oxidized $DP6_{ox}$ products. (c) Chemical structure of non-oxidized hexa-*N*-acetylchitohexaose (DP6) and the corresponding oxidized aldonolactone ($DP6_{lac}$) and aldonic acid ($DP6_{ald}$).

***Pl*AA10 adopts the canonical immunoglobulin-like β-sandwich fold**

To gain further insights into the influence of I113, we determined the X-ray crystal structure of *Pl*AA10 at 1.6 Å resolution. *Pl*AA10 crystallized in space group $P2_12_12_1$ with three protein molecules in the asymmetric unit (Table 1). The crystal structure of *Pl*AA10 reveals a three-stranded and a four-stranded β-sheet forming a β-sandwich and displaying an overall immunoglobulin-like fold typical for LPMOs. Notably, a protrusion domain consisting of mainly loops, small helices, and two short β-strands is appended to the core β-sandwich (Figure 5a). This adjacent region exhibits the most variability among LPMOs [31]. A first disulfide bridge (C39-C50) is found within this protrusion domain, while the second one (C146-C163) interconnects the two β-sheets of the core β-sandwich (Figure 5a). The overall structure of *Pl*AA10 is similar to that of *Sm*AA10 (PDB accession code 2BEM) with root-mean-square deviation (rmsd) values ranging from 0.73 to 0.76 Å for 169 aligned Cα atoms for monomers A, B, and C (monomer A as reference for *Sm*AA10). In addition to the active site copper ion coordinated by H26 and H115 forming the so-called histidine brace motif (geometry described below), a second copper site is found between monomers A and B of two adjacent asymmetric units (Figure S6a). This copper-mediated crystal contact involves surface residues H83 (monomer A), D30 (monomer B), E130 (monomer B), and an alternative conformation of Q108 (monomer B) (Figure S6b). This second metal-binding site is probably due to the presence of excess Cu(II) ions in the protein buffer used for crystallization.

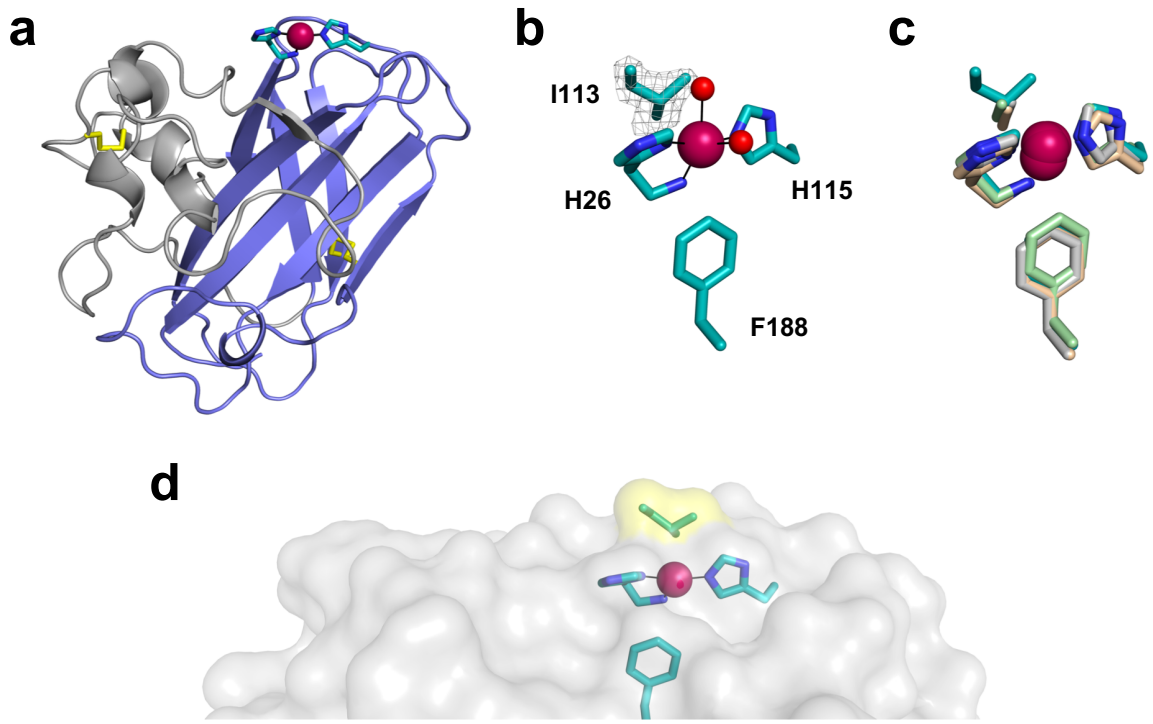

Figure 5. X-ray crystal structure of *Pl*AA10. (a) Overall structure of *Pl*AA10. The core beta sandwich is shown in blue, the protrusion domain in gray, and the disulfide bridges in yellow. As in other LPMOs structures, the histidine brace motif (C in teal and N in blue) coordinates the copper ion (magenta sphere). (b) Active site of *Pl*AA10 including a simulated annealing omit map (contoured at 2.8σ) of I113 side chain (color code as in (a)). The two copper-bound water molecules (red spheres) are partially occupied (~ 40%), which is indicative of a mixed Cu(I)/Cu(II) redox state of the active site metal center. (c) Superimposition of different chitin-active AA10 LPMOs: *Pl*AA10 (teal), *Ba*AA10 from *Bacillus amyloliquefaciens* (light gray, PDB accession code 2YOX), *Ef*AA10 from *Enterococcus faecalis* (light green, PDB accession code 4ALR), and *Bt*AA10A from *Bacillus thuringiensis* subsp. *kurstaki* (wheat, PDB accession code 5WSZ). (d) Chitin-binding surface of *Pl*AA10 (gray) displaying a small protrusion close to the active site due to I113 (yellow highlight). Superimpositions were performed with COOT and all the figures prepared with PyMol.

I113 does not affect the active site copper coordination geometry but protrudes from the substrate-binding surface

Full occupancy copper ions are found within the active site of *Pl*AA10. The histidine brace coordination geometry around the copper is consistent with EPR data and similar to that of other chitin-active bacterial AA10 LPMOs (Table S2) [56]. Namely, the Cu-NH_2(H26) distance (2.2-2.3 Å) is indicative of a mixed Cu(II)/Cu(I) redox state. This is further supported by the presence of partially occupied (~ 40%) water molecules coordinating the active site copper ion (monomer A). Partial reduction of the native active Cu(II) ion of *Pl*AA10 occurred during X-ray diffraction data collection. Indeed, X-ray-induced photoreduction of the five-coordinate Cu(II) exhibiting a distorted geometry intermediate between a square pyramid and a trigonal bipyramid leads to a reduced three-coordinate T-shaped Cu(I) center after discoordination of the two water molecules as previously reported [33,34]. The absorbed dose of 178 kGy lies in the beginning of the dose scale probed by Gundmundsson *et al*. [83 kGy – 1.07 MGy] to visualize the complete reduction of the Cu(II) ion [34]. Interestingly, the copper first coordination sphere is different in monomers B and C with the coordination of an extra ligand from a monomer of an adjacent asymmetric unit. Due to crystal contacts, the *C*-terminal carboxyl group of A199 (in monomers A and B) binds to the active site copper ion (in monomers C and B, respectively) (Figure S7), as observed in the X-ray crystal structure of the bacterial chitin- and cellulose-active AA10 LPMO (*Ma*AA10B) from *Micromonospora auriantica* [25]. A similar coordination mode is also noticed in the structure of the fungal chitin-active AA11 LPMO from *Aspergillus oryzae* or that of the viral AA10 fusolin dimer from *Melolontha melolontha* entomopoxvirus, in which the active site copper ion is coordinated by the side chain of a neighboring glutamate residue in the crystal lattice or in the fusolin dimer, respectively [11,57].

The side chain of I113 adopts a single conformation as shown by a clear electron density in each monomer (Figure 5b). Carbon Cβ of I113 is positioned almost identically to carbon Cβ of the conserved alanine in some chitin-active AA10 LPMOs (Figure 5c), pointing towards the copper and the exterior of the substrate-binding surface. Specifically, carbon Cβ of I113 is 3.9 Å away from the copper ion similarly to that of the alanine residue in other chitin-active AA10 LPMOs (3.8-4.2 Å). A superimposition of all the structures from bacterial chitin-active AA10 LPMOs determined to date shows that there is some slight variation regarding the relative position of the alanine side chain between enzymes (1.0 and 1.1 Å between the corresponding Cα carbons and Cβ carbons, respectively, of the two most dispersed alanine residues between aligned structures, Figure S8a). More variability is observed for bacterial cellulose-active AA10 LPMOs, although fewer X-ray crystal structures have been determined so far. Accordingly, the loop containing the conserved alanine can adopt significantly different conformations between enzymes [22,25]. Consequently, the position of side chain methyl group for this residue can be quite different (2.4 and 2.1 Å between the corresponding Cα carbons and Cβ carbons, respectively, of the two most dispersed alanine residues between aligned structures, Figure S8b) even though its distance to the active site copper ion stays within 3.5-4.0 Å. It was further shown that the alanine residue in bacterial cellulose-active LPMOs is not a determinant of oxidation regioselectivity [25].

The topology of the surface surrounding the active site is essential as it allows substrate binding. For chitin-active LPMOs, interactions with the substrate are mostly mediated by polar residues and a single aromatic amino acid (tyrosine, and sometimes tryptophan) from the substrate-binding surface [32,45,49,53]. Most amino acids important for chitin-binding by *Sm*AA10, as determined by site-directed mutagenesis and NMR, are conserved in *Pl*AA10 (*Pl*AA10 residue numbering: Y55, E56, S59, E61, H115, T117, D183, and N186) [32,45]. However, some residues from the surface that were identified in *Sm*AA10 to interact with chitin

differ in *Pl*AA10: K54, I111, M112, and I113, instead of Q53, L110, T111, and A112 in *Sm*AA10 [32]. T111 and A112 in *Sm*AA10 are in close contact with chitin as probed by molecular dynamics [46]. Notably, the side chain hydroxyl group of T111 in *Sm*AA10 hydrogen bonds with the acetyl group of an adjacent *N*-acetylglucosamine unit of chitin. This interaction will be abolished in *Pl*AA10 with the methionine substitution. However, given the side chain orientation of M112 in the crystal structure of *Pl*AA10 relative to T111 in *Sm*AA10, it might not affect substrate binding. On the contrary, the side chain of A112 in *Sm*AA10 is directly oriented towards chitin and its substitution by a bulkier isoleucine in *Pl*AA10 is likely to cause a steric clash. For instance, the side chain of I113 protrudes from the chitin-binding surface in the crystal structure (Figure 5d) while it is relatively flat in other bacterial chitin-active AA10 LPMOs (Figure S9). The side chain of I113 will most likely adopt a conformation different from that observed in the crystal structure upon substrate binding but still remains bulkier than an alanine. Accordingly, the chitin-*Pl*AA10 interaction is expected to be locally affected close to the active site while it is globally preserved thanks to all other conserved residues important for substrate binding distributed over the chitin-binding surface.

Bacterial AA10 LPMOs exhibit some variability at the active-site alanine position

To date, all the previously reported characterized chitin-active bacterial AA10 LPMOs display the conserved alanine residue in the second coordination sphere of the active site Cu(II) ion. The cellulose-active LPMOs *Cj*AA10B from *Cellvibrio japonicus*, *Hc*AA10A from *Hahella chejuensis*, and *Tt*AA10 from *Teredinibacter turnerae* (all belonging to subclade IV, Figure 2a) are the only characterized bacterial AA10 LPMOs without the active site alanine, but substituted by a glycine instead [23,24,26]. Additionally, a crystal structure of a putative AA10 LPMO from *Burkholderia pseudomallei* with a methionine in place of the alanine was released in the Protein Data Bank (PDB accession code 3UAM). Nevertheless, no functional data are available and sequence analysis indicates that the first 18 *N*-terminal residues (*i.e.*

before the conserved *N*-terminal histidine) are not predicted as a signal-peptide by SignalP 5.0 [58] and Phobius [59]. Accordingly, the full protein was crystallized and no metal-coordinating histidine brace motif is observed in the crystal structure. *Pl*AA10 is the first functional bacterial AA10 LPMO with a bulkier amino acid in place of the alanine to be biochemically and structurally characterized.

We explored potential substitutions at the conserved active-site alanine position among bacterial AA10 LPMOs annotated in the CAZy database. Multiple sequence alignment revealed some variability at this position. As expected, most of the putative bacterial AA10 LPMOs (~ 85%, 3762 out of the 4419 analyzed sequences) display the conserved alanine residue (Table S3). However, conservation at this position is not as high as previously thought. Different types of amino acids were found to replace the second coordination sphere alanine in putative bacterial AA10 LPMO including the less bulky glycine residue as in *Cj*AA10B, *Hc*AA10A, and *Tt*AA10 (24 sequences); hydrophobic residues such as isoleucine (78 sequences) and valine (59 sequences); polar residues such as threonine (278 sequences), glutamine (58 sequences), lysine (10 sequences), serine (2 sequence), and arginine (1 sequence); and metal-coordinating residues such as methionine (143 sequences) and glutamate (4 sequences).

Interestingly, most of the putative isoleucine-containing AA10 LPMOs belong to the membrane-associated subclade (Figure 2a, subclade VI), the activity of which remains to be established although one member was shown to bind crystalline cellulose [48]. Only few members are found within the chitin-active subgroup (Figure 2a, subclade I). Notably, these putative chitin-active AA10 LPMOs share 89-100% sequence identity with *Pl*AA10 (mature protein) and all come from the genus of Gram-negative bacteria *Photorhabdus*: the DSPV002N strain of *Photorhabdus luminescens*, different strains of *Photorhabdus luminescens* subsp. *laumondii* (TT01 and an associated rifampicin-resistant laboratory strain, and DJC), *Photorhabdus asymbiotica* subsp. *asymbiotica*, and *Photorhdabdus thracensis*. Overall, this

analysis unveils additional unusual second coordination sphere residues in place of the conserved alanine in other putative bacterial AA10 LPMOs.

Conclusions

Herein we biochemically and structurally characterized the single LPMO – *Pl*AA10 – from the insect pathogenic bacterium *Photorhabdus luminescens*. *Pl*AA10 was predicted to be active on chitin based on a phylogenetic analysis but sequence alignment revealed it lacks the conserved alanine residue in the close vicinity of the active site copper ion and on the substrate binding surface. Instead, the alanine is substituted by a bulkier isoleucine. We evidenced that *Pl*AA10 binds Cu(II) similarly to other bacterial chitin-active LPMOs and that it is catalytically active on chitin (α- and β-forms) with a C1-oxidation regiospecificity. Its 1.6 Å resolution X-ray crystal structure confirmed that *Pl*AA10 adopts the immunoglobulin-like β-sandwich fold typical for LPMOs. Its overall structure is similar to that of other chitin-active AA10 LPMOs and includes a patch of conserved residues on the surface surrounding its active site that are essential for chitin binding. The copper coordination geometry is not altered by the nearby-isoleucine side chain, although it potentially remodels the substrate-binding surface close to the active site.

In closing, *Pl*AA10 is the first AA10 LPMOs with a bulkier substitution of the conserved active site alanine residue to be characterized. It is interesting to consider the relative variability at the conserved alanine position in the active site of AA10 LPMOs as revealed by a bioinformatic analysis. Substitution by hydrophobic, polar, or even potentially metal coordinating residues provides some diversity close to the copper active site and on the substrate-binding surface. It is expected that these substitutions will alter in some extent enzyme activity and / or enzyme binding to its solid substrate. Future studies will address these effects on AA10 LPMOs catalytic activity and substrate binding and describe the characterization of other bacterial LPMOs with unusual second coordination sphere.

Materials and Methods

Protein sequence analysis

The signal peptide of the protein (GenBank accession number CAE14645.1, UNIPROT accession number Q7N4I5) encoded by *plu2352* was determined and confirmed with web-servers SignalP 5.0 [58] and Phobius [59]. Multiple sequence alignment of *Pl*AA10 (residues 26 to 199) along with the catalytic domain of 32 other characterized and putative bacterial AA10 LPMOs (Table S1) was performed with Clustal Omega [60]. An unrooted phylogenetic tree was built from this alignment and visualized using the web-based tool Interactive Tree Of Life [61].

GenBank accession numbers for all bacterial AA10 LPMOs referenced in the CAZy database (4549 as of 2019, September 1st) were used to retrieve the corresponding amino acid sequences in NCBI. A total of 4532 sequences (16 were found as duplicates and one removed from GenBank) were subsequently aligned with Clustal Omega [60]. 113 sequences lacking either one of the copper(II)-coordinating histidine residues or the conserved axial phenylalanine / tyrosine residue were excluded from the analysis. The remaining 4419 sequences were classified according to the nature of the amino acid at the conserved active site alanine position (Table S3).

Bacterial DNA sequence analysis

Tryptic soy broth (TSB, 5 mL) was inoculated with a *Photorhabdus luminescens* subsp. *laumondii* TT01 (DSM 15139, DSMZ, Germany) glycerol stock (in TSB, stored at -80°C) and was cultured overnight at 28°C under agitation (200 rpm). After centrifugation at 5000 rpm at 4°C for 10 min, genomic DNA was extracted from the cell pellets using the DNeasy Blood and Tissue kit (QIAGEN) according to the manufacturer's protocol. A 1 kb fragment containing *plu2352* was amplified from *Photorhabdus luminescens* genomic DNA by pcr using 5'-CTAGAATTGATAGTTAAGTCCTGGGGCTTTATTTGTCG-3' and 5'-

CAAAGATCCAATGTTAGATTCTATACGCTTGCAGCAGCC-3' as forward and reverse primers, respectively. The pcr product was purified by electrophoresis on a 1% (w/v) agarose gel and extracted using the QIAquick Gel Extraction kit (QIAGEN) according to the manufacturer's protocol. The pcr product was partially sequenced (Genewiz, UK) using 5'-CTAGAATTGATAGTTAAGTCCTGGG-3' as sequencing primer and the *plu2352* nucleotide sequence and corresponding amino acid sequence are reported in Supplementary Information Table S4.

Cloning

The synthetic codon-optimized DNA fragment (Life Technologies) coding for the catalytic domain of the putative LPMO from *Photorhabdus luminescens* subsp. *laumondii* TT01 (GenBank ID CAE14645.1, residues 26 to 199) was amplified by pcr using 5'-CTCCTCGCTGCCCAGCCGGCGATGGCCCATGGCTATATTGATAGTCCGGGTAG-3' and 5'-GTGGTGCTCGAGTGCGGCCGCAAGCTTTATGCCGGACTAATGGTGGTATCAATCACC-3' as forward and reverse primers, respectively. The pcr product was purified by electrophoresis on a 1% (w/v) agarose gel and extracted using the QIAquick Gel Extraction kit (QIAGEN) according to the manufacturer's protocol. The insert was cloned in frame with the pelB signal peptide into the pET-22b(+) vector (Novagen) (double-digested with MscI and HindIII restriction enzymes (NEB)) using the sequence- and ligation-independent cloning method [62]. The corresponding pET-22b(+) SP_{pelB}-*Pl*AA10 plasmids from positive clones were sequenced (Genewiz, UK) to confirm the correct incorporation of the desired construct (nucleotide and amino acid sequences in Table S4).

Protein expression and purification

The pET-22b(+) SP_{pelB}-*Pl*AA10 plasmid was transformed into chemocompetent BL21(DE3) *Escherichia coli* cells. A preculture of a fresh transformant in lysogeny broth (LB,

50 mL) supplemented with ampicillin (100 µg/mL) was grown overnight at 37°C under agitation (200 rpm), and was used to inoculate 6 × 500 mL cultures (1% v/v inoculum in LB supplemented with 100 µg/mL ampicillin). Cells were grown at 37°C under agitation (200 rpm) until OD_{600} ~ 0.5 at which point the temperature was lowered to 16 °C. At OD_{600} ~ 0.7, cells were induced with isopropyl-β-D-thiogalactopyranoside (IPTG, 0.2 mM final concentration) and grown overnight (20 h) at 16 °C and 200 rpm. Cells were harvested by centrifugation at 4 °C (10 min at 5000 g). The periplasmic protein fraction was prepared by cold osmotic shock as previously described with slight modifications [53]. Briefly, the cell pellet was re-suspended in 100 ml (1/30 of the total culture volume) ice-cold sucrose buffer (50 mM Tris, 20% w/v sucrose, 1 mM EDTA, pH 8.0) and incubated at 4°C under gentle agitation (4 rpm on a tube rotator). After 1h, the cell suspension was supplemented with $MgSO_4$ (60 µL of 1M $MgSO_4$ solution for each gram of wet cell pellet) and incubated under the same conditions for 30 min. The supernatant was collected by centrifugation at 4°C (20 min at 20000 g) and stored at 4°C. The corresponding cell pellet was re-suspended in 100 mL (1/30 of the total culture volume) ice-cold milliQ water and further incubated under gentle agitation (4 rpm on a tube rotator) for 45 min at 4°C. The cell suspension was centrifuged (20 min at 20000 g, 4 °C) to recover the supernatant. The first supernatant fraction was dialyzed against 20 mM Tris (pH 8.0), 1 mM EDTA in 3.5 kDa molecular weight cut-off (MWCO) dialysis tubing (Spectra/Por, Spectrumlabs). The second supernatant fraction was supplemented with 20 mM Tris (pH 8.0). Both periplasmic fractions were adjusted to 1.0 M $(NH_4)_2SO_4$ by slow addition of solid ammonium sulfate at 4°C, and filtered over a 0.22 µm membrane. Periplasmic extracts were loaded at 4°C onto a chitin bead column (NEB) pre-equilibrated with 20 mM Tris (pH 8.0), 1.0 M $(NH_4)_2SO_4$ (running buffer). After washing with the running buffer, *Pl*AA10 was eluted with 20 mM acetic acid. Fractions containing purified *Pl*AA10 as determined by SDS-PAGE were adjusted with 70 mM Tris buffer (pH 8.0), pooled, and concentrated (10 kDa MWCO Vivaspin

20, Sartorius). Protein metalation was done at 4°C by the addition of a 5-fold molar excess of $CuSO_4$. Excess Cu(II) was removed by two successive buffer exchanges over a PD-10 column (GE Healthcare) with 50 mM MES (pH 6.5) as storage buffer. Apo-*Pl*AA10 was prepared similarly *i.e.* two successive buffer exchanges in the same storage buffer after the purification by affinity chromatography. After concentration, apo- and holo-*Pl*AA10 were stored at -80°C. Protein concentration was determined from the absorbance at 280 nm using a calculated extinction coefficient of 38180 $M^{-1}cm^{-1}$ [63].

Fluorescence

Fluorescence spectroscopy was performed on a Horiba Jobin-Yvon Fluoromax-4 fluorimeter. Emission spectra following an excitation at 295 nm (1 nm excitation slit widths) were recorded at 25°C between 300 and 450 nm (20 nm emission slit widths). Protein solution was prepared by dilution of a stock solution of apo-*Pl*AA10 into 50 mM MES (pH 6.5), with a final protein concentration of 1 µM. Fluorescence was recorded upon protein titration with a solution of $Cu(II)(NO_3)_2$ (successive addition of 0.2 equivalent) after an equilibration of 2 min at 25°C.

Electron paramagnetic resonance (EPR) spectroscopy and simulation

Protein samples (holo-*Pl*AA10 concentrations ranging from 150 to 200 µM) for continuous wave EPR were prepared in 50 mM MES buffer (pH 6.5) supplemented or not with 10% (v/v) glycerol. Buffers with and without glycerol were prepared independently and both adjusted to pH 6.5. EPR spectra were recorded on a Bruker Elexsys E500 spectrometer operating at X-band at 120 K (BVT 3000 digital temperature controller) with the following acquisition parameters: modulation frequency, 100 kHz; modulation amplitude, 5 G; conversion time, 120 ms; sweep time, 123 s; and microwave power, 20 mW. EPR spectra were simulated using the EasySpin toolbox developed for Matlab [64].

Chitin cleavage assay

The activity of *Pl*AA10 was assessed on both α-chitin from shrimp shell (practical grade, powder, Sigma Aldrich) and β-chitin from squid-pen (particle size < 500 µm, kind gift from Dr. D. Gillet, Gillet Chitosan SAS, France). 1 mL-reactions containing 5 µM *Pl*AA10, 2.5 mg chitin, 1 mM ascorbic acid in deionized water (no buffer), or in 10 mM sodium acetate (pH 5.0), or in 10 mM MES (pH 6.5), or in 10 mM Tris (pH 8.0) were set up in capped 15 mL round-bottom polypropylene tubes. Samples were incubated horizontally at 37°C for 24 hours whilst rotating to ensure proper substrate mixing. Reaction supernatants were collected after centrifugation at 13000 rpm and filtrated over a 10 kDa MWCO centrifugal concentrator to remove any remaining enzyme before analysis of the filtrate by mass spectrometry (MS).

Mass spectrometry

Products resulting from enzyme reactions as described above were analyzed by MS. Two types of mass measurements were performed on the samples: firstly, a mass profile was recorded by matrix-assisted laser desorption/ionization (MALDI)-time of flight (TOF) MS. Then, ions of interest were selectively analyzed by electrospray ionization (ESI) MS/MS upon a fragmentation by collision-induced dissociation (CID). For MALDI-TOF MS measurements, an ionic preparation of 2,5-dihydroxybenzoic acid (DHB) and *N*,*N*-dimethylaniline (DMA) was used as the MALDI matrix, according to Ropartz *et al*. [65]. Samples (1 µL) were deposited on a polished steel MALDI target plate and covered by the matrix (1 µL). After a complete drying of the deposits, MALDI measurements were performed on an Autoflex Speed MALDI-TOF/TOF spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a Smartbeam laser (355 nm, 1000 Hz) and controlled using the Flex Control 3.0 software package. The mass spectrometer was operated with positive polarity in a reflectron mode, and spectra were acquired in the range of 400–2500 *m/z*. For ESI MS/MS measurements, experiments were

performed on a Synapt G2Si high-definition mass spectrometer (Waters Corp., Manchester, UK). Fragmentation was performed by CID in the transfer cell of the instrument, using appropriate collision energies depending on the precursor. Samples were diluted ten-fold in MeOH/H_2O (1:1, v/v) and infused at a flow rate of 5 µL.min^{-1}. Acquisitions were conducted in positive polarity, as well as in 'sensitivity' mode.

Protein crystallization

Initial crystallization screening was performed with the two high-throughput sparse-matrix screening kits Index HT and Crystal (Hampton Research, USA) in sitting drops using the vapor diffusion method. In general, a 1 µL drop of precipitant was added to a 1 µL drop containing 5 or 10 mg/mL holo-*Pl*AA10, 50 mM MES (pH 6.5), and 500 µM $CuSO_4$ and equilibrated against a 80 µL reservoir of precipitant solution at 21°C. Among initial hits, needle and rod-like crystals were obtained with Crystal screen condition E11 consisting of 0.01 M cobalt chloride, 0.1 M sodium acetate trihydrate (pH 4.6), and 1.0 M 1,6-hexanediol. Crystallization was further optimized by a grid screen with the following precipitant conditions: 0 to 0.01 M cobalt chloride, 0.1 M sodium acetate (pH 4.3 to 5.2), and 0.70 to 1.15 M 1,6-hexanediol. Typically, crystals appeared after one to three days and grew to full dimensions (~ 100 x 200 x 700 µm) within a week.

Data collection and structure refinement

Prior to data collection, crystals were flash-cooled in liquid nitrogen after transfer to a cryoprotectant solution consisting of precipitant solution supplemented with 15% (v/v) glycerol. The best diffracting crystal (1.6 Å) was obtained as previously described with a 10 mg/mL protein solution (50 mM MES (pH 6.5), and 500 µM $CuSO_4$) and a precipitant solution containing 0.1 M sodium acetate (pH 5.2) and 0.7 M 1,6-hexanediol. X-ray diffraction data were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on beamline ID29 [66] at 100 K at a wavelength of 0.976 Å. A continuous oscillation range of

101° was collected in 74.74 s with an X-ray flux of 4.48 10^{10} photons.s^{-1}, resulting in an average diffraction-weighted dose absorbed by the exposed region of 178 kGy calculated using RADDOSE-3D available at http://raddo.se [67]. Data were integrated and reduced using the autoPROC pipeline [68] chosen from the series of automated reduction pipelines [69], whose results are available through the web information management system ISPyB [70]. The structure belonged to space group $P2_12_12_1$ with three protein molecules in the asymmetric unit. The structure was solved by molecular replacement using PHASER [71] implemented in PHENIX [72] with the atomic coordinates (PDB accession code 2BEM) of the wild-type AA10 LPMO from *Serratia marcescens* (*Sm*AA10, also known as CBP21) less ion, and solvent used as a search probe for rotation and translation function calculations. Each model was refined with iterative cycles of refinement in PHENIX [72] and manual model rebuilding in COOT [73]. Ions and solvent molecules were added after several rounds of refinement. In the final stages, translation libration screw (TLS) refinement was performed. TLS groups were automatically determined using PHENIX [72]. Anisotropic temperature factor refinement was performed at the last stage of refinement for all atoms except water molecules. Side chains of residues that were disordered were excluded from the final model and include E167 in monomer A; E167 in monomer B; and K96, E104, Q151, K153, and R158 in monomer C. Few electron density peaks that were not confidently interpretable were left unmodeled (such as close to D49 in monomer C). Detailed data collection and refinement statistics are recorded in Table 1.

Author contribution

CD designed the research and coordinated the project. Molecular biology (cloning) was done by AM and CD. AM and BEK expressed and purified the protein. BEK performed fluorescence experiments. AM and AJS recorded EPR spectra. AJS simulated the EPR spectra. AM and BEK ran activity assays on chitin. MF performed mass spectrometry experiments. MF, HR, and DR analyzed mass spectrometry data. CD and BEK crystallized the protein. AR collected diffraction data at the synchrotron. CD and AR solved the crystal structure. Bioinformatic analysis was carried out by AM and CD. All the data were compiled and analyzed by CD, JS, and MR. All authors contributed to the writing and approved the final version of the manuscript.

Acknowledgements

CNRS and Aix Marseille Université are acknowledged for their financial support. AM is a recipient of a French Ministère de l'Enseignement Supérieur et de la Recherche PhD fellowship. The ESRF is acknowledged for access to beamline ID29 *via* its in-house research program.

Conflicts of Interest

The authors declare no conflict of interest.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Origin, name, and GenBank accession number of all AA10 LPMOs sequences included in the unrooted phylogenetic tree.

Table S2. Geometry of the copper active site of *Pl*AA10.

Table S3. Amino acid variability at the conserved alanine position in bacterial AA10 LPMOs referenced in the CAZy database and GenBank accession number of the corresponding sequences.

Table S4. Amino acid sequence of *Pl*AA10 with its native signal peptide or the pelB signal peptide, and nucleotide sequence of the corresponding genes (*plu2352* in *Photorhabdus luminescens* subsp. *laumondii* TT01 and the gene cloned in pET-22b(+))

Figure S1. Titration of apo-*Pl*AA10 with Cu(II).

Figure S2. EPR spectroscopy of *Pl*AA10 in the absence of glycerol.

Figure S3. Activity of *Pl*AA10 towards α-chitin at different pH values.

Figure S4. Activity of *Pl*AA10 towards β-chitin at different pH values.

Figure S5. MS analysis of oxidized chitooligosaccharides produced by *Pl*AA10.

Figure S6. Second copper-binding site in the crystal structure of *Pl*AA10.

Figure S7. Coordination of the active site copper (in monomer B and C) by the *C*-terminal A199 carboxyl group of a neighboring monomer in an adjacent asymmetric unit.

Figure S8. Superimposition of structures of bacterial chitin-active AA10 LPMOs and bacterial cellulose-active LPMOs highlighting the variability of the position for the side chain alanine residue.

Figure S9. Comparison of the substrate-binding surface of *Ef*AA10, a chitin-active LPMO from *Enterococcus faecalis* with a conserved alanine residue and *Pl*AA10.

References

1. Cosgrove DJ (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* 6, 850-861.

2. Cohen E (2001) Chitin synthesis and inhibition: a revisit. *Pest Manag Sci* 57, 946-950.

3. Reese ET, Siu RGH & Levinson HS (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J Bacteriol* 59, 485-497.

4. Vaaje-Kolstad G, Horn SJ, van Alten DMF, Synstad B and Eijsink VGH (2005) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem* 280, 28492-28497.

5. Harris PV, Welner D, McFarland KC, Re E, Poulsen J-CN, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S & 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.

6. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M & Eijsink VGH (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330, 219-222.

7. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen J-CN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, Tryfona T, Walter CP, Dupree P, Xu F, Davies GJ & Walton PH (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci USA* 108, 15079-15084.

8. Phillips CM, Beeson WT, Cate JH & Marletta MA (2011). Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol* 6, 1399-1406.

9. Lombard V, Golaconda Ramula H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42, D490-D495.

10. Levasseur A, Drula E, Lombard V, Coutinho PM & Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels* 6, 41.

11. Hemsworth GR, Henrissat B, Davies GJ & Walton PH, 2014. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* 10, 122-126.

12. Vu VV, Beeson WT, Span EA, Farquhar ER & Marletta MA (2014) A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci USA* 111, 13822-13827.

13. Lo Leggio L, Simmons TJ, Poulsen J-CN, Frandsen KEH, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L, Harris PV, Soong C-L, Dupree P, Tryfona T, Lenfant N, Henrissat B, Davies GJ & Walton PH (2015) Structure and boosting activity of a starch-degrading polysaccharide monooxygenase. *Nat Commun* 6, 5961.

14. Couturier M, Ladevèze S, Sulzenbacher G, Ciano L, Fanuel M, Moreau C, Villares A, Cathala B, Chaspoul F, Frandsen KEH, Labourel A, Herpoël-Gimbert I, Grisel S, Haon M, Lenfant N, Rogniaux H, Ropartz D, Davies GJ, Rosso MN, Walton PH, Henrissat B & Berrin J-G (2018) Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol* 14, 306-310.

15. Sabbadin F, Hemsworth GR, Ciano L, Henrissat B, Dupree P, Tryfona T, Marques RDS, Sweeney ST, Besser K, Elias L, Pesante G, Li Y, Dowle AA, Bates R, Gomez LD, Simister R, Davies GJ, Walton PH, Bruce NC & McQueen-Mason SJ (2018) An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun* 9, 756.

16. Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoël-Gimbert I, Chevret D, Fanuel M, Henrissat B, Heiss-Blanquet S, Margeot A & Berrin J-G (2019) AA16, a new lytic

polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol Biofuels* 12, 55.

17. Forsberg Z, Røhr AK, Mekasha S, Andersson KK, Eijsink VGH, Vaaje-Kolstad G & Sørlie M (2014) Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry* 53, 1647-1656.

18. Nakagawa YS, Kudo M, Loose JS, Ishikawa T, Totani K, Eijsink VGH & Vaaje-Kolstad G (2015) A small lytic polysaccharide monooxygenase from *Streptomyces griseus* targeting α- and β-chitin. *FEBS J* 282, 1065-1079.

19. Gregory RC, Hemsworth GR, Turkenburg JP, Hart SJ, Walton PH & Davies GJ (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.

20. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stentrøm Y, MacKenzie A, Sørlie M, Horn SJ & Eijsink VGH (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* 20, 1479-1483.

21. Beeson WT, Phillips CM, Cate JH & Marletta MA (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc* 134, 890-892.

22. Forsberg Z, Mackenzie AK, Sørlie M, Røhr AK, Helland R, Arvai AS, Vaaje-Kolstad G & Eijsink VGH (2014) Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci USA* 111, 8446-8451.

23. Gardner JG, Crouch L, Labourel A, Forsberg Z, Bukhman YV, Vaaje-Kolstad G, Gilbert HJ & Keating DH (2014) Systems biology defines the biological significance of a redox-active proteins during cellulose degradation in an aerobic bacterium. *Mol Microbiol* 94, 1121-1133.

24. Ghatge SS, Telke AA, Waghmode TR, Lee Y, Lee KW, Oh DB, Shin HD & Kim SW (2015) Multifunctional cellulolytic auxiliary activity protein *Hc*AA10-2 from *Hahella chejuensis* enhances enzymatic hydrolysis of crystalline cellulose. *Appl Microbiol Biotechnol* 99, 3041-3055.

25. Forsberg Z, Bissaro B, Gullesen J, Dahlus B, Vaaje-Kolstad G & Eijsink VGH (2018) Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J Biol Chem* 293, 1397-1412.

26. Fowler CA, Sabbadin F, Ciano L, Hemsworth GR, Elias L, Bruce N, McQueen-Mason S, Davies GJ & Walton PH (2019) Discovery, activity and characterization of an AA10 lytic polysaccharide oxygenase from the shipworm symbiont *Teredinibacter turnerae*. *Biotechnol Biofuels* 12, 232.

27. Bissaro B, Várnai A, Røhr ÅK & Eijsink VGH (2018) Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. *Microbiol Mol Biol Rev* 82, e00029-18.

28. Bissaro B, Røhr ÅK, Müller G, Chylenski P, Skaugen M, Forsberg Z, Horn SJ, Vaaje-Kolstad G & Eijsink VGH (2017) Oxidative cleavage of polysaccharides by monocopper enzymes depends on H_2O_2. *Nat Chem Biol* 13, 1123-1128.

29. Wang B, Johnston EM, Li P, Shaik S, Davies GJ, Walton PH & Rovira C (2018) QM/MM studies into the H_2O_2-dependent activity of lytic polysaccharide monooxygenases: evidence for the formation of a caged hydroxyl radical intermediate. *ACS Catal* 8, 1346-1351.

30. Hedegård ED & Ryde U (2018) Molecular mechanism of lytic polysaccharide monooxygenases. *Chem Sci* 9, 3866-3880.

31. Vaaje-Kolstad G, Forsberg Z, Loose JS, Bissaro B & Eijsink VGH (2017) Structural diversity of lytic polysaccharide monooxygenases. *Curr Opin Struct Biol* 44, 67-76.

32. Aachmann FL, Sørlie M, Skjåk-Bræk G, Eijsink VGH & Vaaje-Kolstad G (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc Natl Acad Sci USA* 109, 18779-18784.

33. Hemsworth GR, Taylor EJ, Kim RQ, Gregory RC, Lewis SJ, Turkenburg JP, Parkin A, Davies GJ & Walton PH (2013) The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc* 135, 6069-6077.

34. Gudmundsson M, Kim S, Wu M, Ishida T, Momeni MH, Vaaje-Kolstad G, Lundberg D, Royant A, Ståhlberg J, Eijsink VGH, Beckham GT & Sandgren M (2014) Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of lytic polysaccharide monooxygenase by X-ray photoreduction. *J Biol Chem* 289, 18782-18792.

35. Lieberman RL and Rosenzweig AC (2005) Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* 434, 177-182.

36. Cao L, Caldararu O, Rosenzweig AC & Ryde U (2018) Quantum refinement does not support dinuclear copper sites in crystal structures of particulate methane monooxygenases. *Angew Chem Int Ed Engl* 57, 162-166.

37. Zhang L, Koay M, Maher MJ, Xiao Z & Wedd AG (2006) Intermolecular transfer of copper ions from the CopC protein of *Pseudomonas syringae*. Crystal structures of fully loaded $Cu^{I}Cu^{II}$ forms. *J Am Chem Soc* 128, 5834-5850.

38. Lawton TJ, Kenney GE, Hurley JD & Rosenzweig AC (2016) The CopC family: structural and bioinformatic insights into a diverse group of periplasmic copper binding proteins. *Biochemistry* 55, 2278-2290.

39. Fisher OS, Sendzik MR, Ross MR, Lawton TJ, Hoffman BM & Rosenzweig AC (2019) PCu_AC domains from methane-oxidizing bacteria use a histidine brace to bind copper. *J Biol Chem* in press, doi:10.1074/jbc.RA119.010093.

40. Agostoni M, Hangasky JA & Marletta MA (2017). Physiological and molecular understanding of bacterial polysaccharide monooxygenases. *Microbiol Mol Biol Rev* 81, e00015-17.

41. Duchaud E, Rusniok C, Frangeul L, Buchrieser C, Givaudan A, Taourit S, Bocs S, Boursaux-Eude C, Chandler M, Charles JF, Dassa E, Derose R, Derzelle S, Freyssinet G, Gaudriault S, Médigue C, Lanois A, Powell K, Siguier P, Vincent R, Wingate V, Zouine M, Glaser P, Boemare N, Danchin A & Kunst F (2003) The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nat Biotechnol* 21, 1307-1313.

42. Nielsen-LeRoux C, Gaudriault S, Ramarao N, Lereclus D & Givaudan A (2012) How the insect pathogen bacteria *Bacillus thuringiensis* and *Xenorhabdus*/*Photorhabdus* occupy their hosts. *Curr Opin Microbiol* 15, 220-231.

43. Garcia-Gonzalez E, Poppinga L, Fünfhaus A, Hertlein G, Hedtke K, Jakubowska A & Genersch E (2014). *Paenibacillus larvae* chitin-degrading protein *Pl*CBP49 is a key virulence factor in American foulbrood of honey bees. *PLoS Pathog* 10, e1004284.

44. Chen G, Zhang Y, Li J, Dunphy GB, Punja ZK & Webster JM (1996) Chitinase activity of *Xenorhabdus* and *Photorhabdus* species, bacterial associates of entomopathogenic nematodes. *J Invertebr Pathol* 68, 101-108.

45. Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VGH & van Alten DMF (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* 280, 11313-11319.

46. Bissaro B, Isaksen I, Vaaje-Kolstad G, Eijsink VGH & Røhr ÅK (2018) How a lytic polysaccharide monooxygenase binds crystalline chitin. *Biochemistry* 57, 1893-1906.

47. Book AJ, Yennamalli RM, Takasuka TE, Currie CR, Phillips GN Jr & Fox BG (2014) Evolution of substrate specificity in bacterial AA10 lytic polysaccharide monooxygenases. *Biotechnol Biofuels* 7, 109.

48. Walter S & Schrempf H (2008) Characteristics of the surface-located carbohydrate-binding protein CbpC from *Streptomyces coelicolor* A3(2). *Arch Microbiol* 190, 119-127.

49. Vaaje-Kolstad G, Bøhle LA, Gåseidnes S, Dahlus B, Bjørås M, Mathiesen G & Eijsink VGH (2012) Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J Mol Biol* 416, 239-254.

50. Yu MJ, Yoon SH & Kim YW (2016) Overproduction and characterization of a lytic polysaccharide monooxygenase in *Bacillus subtilis* using an assay based on ascorbate consumption. *Enzyme Microb Technol* 93-94, 150-156.

51. Russo DA, Zedler JAZ, Wittman DN, Möllers B, Singh RK, Batth TS, van Oort B, Olsen JV, Bjerrum MJ & Jensen PE (2019) Expression and secretion of a lytic polysaccharide monooxygenase by a fast-growing cyanobacterium. *Biotechnol Biofuels* 12, 74.

52. Rodrigues KB, Macêdo JKA, Teixaira T, Barros JS, Araújo ACB, Santos FP, Quirino BF, Brasil BSAF, Salum TFC, Abdelnur PV & Fávaro LCL (2017) Recombinant expression of *Thermobifida fusca* E7 LPMO in *Pichia pastoris* and *Escherichia coli* and their functional characterization. *Carbohydr Res* 448, 175-181.

53. Chaplin AK, Wilson MT, Hough MA, Svistunenko DA, Hemsworth GR, Walton PH, Vijgenboom E & Worrall JAR (2016) Heterogeneity in the histidine-brace copper coordination sphere in Auxiliary Activity Family 10 (AA10) lytic polysaccharide monooxygenases. *J Biol Chem* 291, 12838-12850.

54. Peisach J & Blumberg WE (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins. *Arch Biochem Biophys* 165, 691-708.

55. Hemsworth GR, Ciano L, Davies GJ & Walton PH (2018) Production and spectroscopic characterization of lytic polysaccharide monooxygenases. *Methods Enzymol* 613, 63-90.

56. Ciano L, Davies GJ, Tolman WB & Walton PH (2018) Bracing copper for the catalytic oxidation of C-H bonds. *Nat Catal* 1, 571-577.

57. Chiu E, Hijnen M, Bunker RD, Boudes M, Rajendran C, Aizel K, Oliéric V, Schulze-Briese C, Mitsuhashi W, Young V, Ward VK, Bergoin M, Metcalf P & Coulibaly F (2015) Structural basis for the enhancement of virulence by viral spindles and their *in vivo* crystallization. *Proc Natl Acad Sci USA* 112, 3973-3978.

58. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G & Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37, 420-423.

59. Käll L, Krogh A & Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. *Nucleic Acids Res* 35, W429-W432.

60. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD & Higgins DG (2011) Fast scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539.

61. Letunic I & Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127-128.

62. Jeong JY, Yim HS, Ryu JY, Lee HS, Lee JH, Seen DS & Kang SG (2012) One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl Environ Microbiol* 78, 5440-5443.

63. Gill SC & von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182, 319-326.

64. Stoll S & Schweiger A (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J Magn Res* 178, 42-55.

65. Ropartz D, Bodet PE, Przybylski C, Gonnet F, Daniel R, Fer M, Helbert M, Bertrand D & Rogniaux H (2011) Performance and evaluation on a wide set of matrix-assisted laser desorption ionization matrices for the detection of oligosaccharides in a high-throughput mass spectrometry screening of carbohydrate depolymerizing enzymes. *Rapid Commun Mass Spectrom* 25, 2059-2070.

66. de Sanctis D, Beteva A, Caserotto H, Dobias F, Gabadinho J, Giraud T, Gobbo A, Guijarro M, Lentini M, Lavault B, Mairs T, McSweeney S, Petitdemange S, Rey-Bakaikoa V, Surr J, Theveneau P, Leonard GA & Mueller-Dieckmann C (2012) ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J Synchrotron Radiat* 19, 455-461.

67. Zeldin OB, Gerstel M & Garman EF (2013) *RADDOSE-3D*: time- and space-resolved modelling of dose in macromolecular crystallography. *J Appl Crystallogr* 46, 1225-1230.

68. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T & Bricogne G (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr* D67, 293-302.

69. Monaco S, Gordon E, Bowler MW, Delagenière S, Guijarro M, Spruce D, Svensson O, McSweeney SM, McCarthy AA, Leonard G & Nanao MH (2013) Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *J Appl Crystallogr* 46, 804-810.

70. Delagenière S, Brenchereau P, Launer L, Ashton AW, Leal R, Veyrier S, Gabadinho J, Gordon EJ, Jones SD, Levik KE, McSweeney SM, Monaco S, Nanao M, Spruce D, Svensson O, Walsh MA & Leonard GA (2011) ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics* 27, 3186-92.

71. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* 40, 658-674.

72. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr* D66, 213-221.

73. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr* D66, 486-501.

Tables

Table 1. Data collection and refinement statistics

	Wild-type *Pl*AA10
Unit cell	
space group symmetry	$P2_12_12_1$
a, b, c (Å)	76.5, 80.4, 90.9
α, β, γ (deg)	90, 90, 90
Data collection	
wavelength (Å)	0.976
resolution limits (Å)	58.55-1.60
total/unique reflections	259629/72679
R_{merge}[a,b]	0.078 (0.547)
$CC_{1/2}$[a,c]	1.00 (0.80)
$I/\sigma(I)$[a]	9.5 (2.1)
redundancy[a]	3.6 (3.6)
completeness (%)[a]	95.2 (99.7)
Refinement	
reflections used in refinement/test set	70875/3556
R_{cryst}[a,d]	0.188 (0.249)
R_{free}[a,e]	0.218 (0.274)
protein atoms[f]	4247
water molecules[f]	751
Cu(I/II) ions[f]	6
Rms deviations from ideal geometry	
bonds (Å)	0.006
angles (°)	0.9
Ramachandran plot (%)[g]	
favored	98.9
allowed	1.1
outliers	0.0

PDB accession code	6T5Z

[a] Values in parentheses refer to the highest shell of data. [b] $R_{merge} = \sum | I_h - \langle I \rangle_h | / \sum I_h$, where $\langle I \rangle_h$ is the average intensity for reflection *h* calculated from replicate reflections. [c] Pearson correlation coefficient between random half-datasets. [d] $R_{cryst} = \sum | |F_o| - |F_c| | / \sum |F_o|$ for reflections contained in the working set. $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. [e] $R_{free} = \sum | |F_o| - |F_c| | / \sum |F_o|$ for reflections contained in the test set held aside during refinement. [f] Per asymmetric unit. [g] Calculated with MolProbity.