Insights into an unusual Auxiliary Activity 9 family member lacking the histidine brace motif of lytic polysaccharide monooxygenases

Frandsen, Kristian E H; Tovborg, Morten; Jørgensen, Christian I.; Spodsberg, Nikolai; Rosso, Marie-Noëlle; Hemsworth, Glyn R; Garman, Elspeth F; Grime, Geoffrey W; Poulson, Jens-Christian N; Batth, Tanveer S.; Miyauchi, Shingo; Lipzen, Anna; Daum, Chris; Grigoriev, Igor V; Johansen, Katja S; Henrissat, Bernard; Berrin, Jean-Guy; Lo Leggio, Leila

Published in:
The Journal of Biological Chemistry

DOI:
10.1074/jbc.RA119.009223

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):

Download date: 12. sep., 2020
Insights into an unusual Auxiliary Activity 9 family member lacking the histidine brace motif of lytic polysaccharide monoxygenases

Kristian E. H. Frandsen, a,b Morten Tovborg, c Christian I. Jørgensen, a Nikolaj Spodsberg, a Marie-Noëlle Rosso, b Glyn R. Hemsworth, d,e Elspeth F. Garman, a Geoffrey W. Grime, a Jens-Christian N. Poulsen, a Tanveer S. Batth, f Shingo Miyachi, b Anna Lipzen, c Chris Daum, c Igor V. Grigoriev, a Katja S. Johansen, k Bernard Henriassat, c,m,n Jean-Guy Berrin, b and Leila Lo Leggio a,b

From the aDepartment of Chemistry, University of Copenhagen, 2100 Copenhagen, Denmark, the bINRA, Aix-Marseille Université, UMR1163 BBF (Biodiversité et Biotechnologie Fongiques), 13009 Marseille, France, the Novozymes A/S, 2880 Bagsvaerd, Denmark, the School of Molecular and Cellular Biology and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom, the Department of Chemistry, University of York, York Y010 5DD, United Kingdom, the Department of Biochemistry, University of Oxford, Oxford OX1 3QZ, United Kingdom, The Ion Beam Centre, Advanced Technology Institute, University of Surrey, Guildford GU2 7XH, United Kingdom, The Jen W. Kimball Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California 94720, the Department of Geosciences and Natural Resource Management, University of Copenhagen, 1958 Frederiksberg C, Denmark, the aArchitecture et Fonction des Macromolécules Biologiques (AFMB), CNRS, Aix-Marseille Université, 13009 Marseille, France, INRA, USC 1408 AFMB, 13009 Marseille, France, and the Department of Biological Sciences, King Abdulaziz University, 21589 Jeddah, Saudi Arabia

Edited by Jeffrey E. Pessin

Lytic polysaccharide monoxygenases (LPMOs) are redox-enzymes involved in biomass degradation. All characterized LPMOs possess an active site of two highly conserved histidine residues coordinating a copper ion (the histidine brace), which are essential for LPMO activity. However, some protein sequences that belong to the AA9 LPMO family display a natural N-terminal His to Arg substitution (Arg-AA9). These are found almost entirely in the phylogenetic fungal class Agaricomycetes, associated with wood decay, but no function has been demonstrated for any Arg-AA9. Through bioinformatics, transcriptomic, and proteomic analyses we present data, which suggest that Arg-AA9 proteins could have a hitherto unidentified role in fungal degradation of lignocellulosic biomass in conjunction with other secreted fungal enzymes. We present the first structure of an Arg-AA9, LsAA9B, a naturally occurring protein from Lentinus similis. The LsAA9B structure reveals gross changes in the region equivalent to the canonical LPMO copper-binding site, whereas features implicated in carbohydrate binding in AA9 LPMOs have been maintained. We obtained a structure of LsAA9B with xylotetraose bound on the surface of the protein although with a considerably different binding mode compared with other AA9 complex structures. In addition, we have found indications of protein phosphorylation near the N-terminal Arg and the carbohydrate-binding site, for which the potential function is currently unknown. Our results are strong evidence that Arg-AA9s function markedly different from canonical AA9 LPMO, but nonetheless, may play a role in fungal conversion of lignocellulosic biomass.

Fungal degradation of plant biomass involves a collection of carbohydrate active enzymes (CAZymes), 2 which include

2 The abbreviations used are: CAZymes, Carbohydrate Active enZymes; LPMO, Lytic polysaccharide monoxygenase; AA, auxiliary activity; GH, glycoside hydrolases; CE, carbohydrate esterases; RNA-seq, RNA sequencing.
An unusual AA9 naturally lacking the His-brace motif

![Schematic representation of oxidative cleavage of cellooligosaccharides/cellulose by LPMOs.](image)

Figure 1. Schematic representation of oxidative cleavage of cellooligosaccharides/cellulose by LPMOs.

glycoside hydrolases, carbohydrate esterases, polysaccharide lyases, and oxidoreductases. Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent oxidoreductases shown to be pivotal in efficient plant biomass degradation (1). These enzymes activate molecular oxygen or hydrogen peroxide to introduce oxidative chain breaks into polysaccharide chains (2, 3) (Fig. 1). Some LPMOs are single domain enzymes, whereas others are linked to CBMs (carbohydrate binding modules) or domains of unknown function (4, 5). LPMOs are classified into seven Auxiliary Activity families in the CAZY database, AA9-AA11 and AA13-AA16 (6), and have been found in bacteria, viruses, fungi, and recently in arthropod species and in plants (7, 8). In the active site a copper ion is held by a motif denoted the histidine brace (His-brace) formed by the highly conserved N-terminal histidine and a second histidine later in the sequence (9). The His-brace motif is strictly conserved in all LPMOs studied to date (10–12) and is considered essential for LPMO activity. This is supported by mutational analysis of both AA9 and AA10 LPMOs (first performed on TtAA9E and SmAA10A, followed by later studies on MtAA9D and ScAA10C) in which replacement of these critical residues within the His-brace or the secondary coordination sphere of the copper ion resulted in reduced LPMO activity and hence reduced glycoside hydrolase boosting (13–16).

LPMOs display different substrate specificities, with activity demonstrated on cellulose, hemicelluloses, chitin, and starch (2, 9, 14, 17–19) and interact with their polysaccharide substrates through a surface on which the His-brace is found (20–24). This surface can have varying contour and polarity, presumably dictating substrate specificity (19, 23, 25, 26). Additionally, for some AA9 LPMOs activity has been demonstrated on soluble β-1,4-linked polysaccharides and also oligosaccharides (17, 22, 23, 27, 28), a feature that was instrumental in the determination of the first enzyme-substrate complex structures of an AA9 from *Lentinus similis* (LsAA9A) (22, 29). The structures revealed polar residues in a loop denoted L3 (after notation in Ref. 25) interacting with cellooligosaccharides near the active site and demonstrated for the first time that a conserved Tyr (Tyr-203 in LsAA9A) located on the same surface was involved in substrate binding (22, 23). In addition these structures revealed a new and highly unusual lone pair-π aromatic interaction between the pyranose ring O5 and the imidazole of the N-terminal His of the His-brace, further highlighting its importance for LPMO function (29).

Both in *Basidiomycetes* and *Ascomycetes* fungal species multiple AA9-encoding genes are found (more than 20 and 30 in *Schizophyllum commune* and *Podospora anserina*, respectively), the vast majority of which encode proteins with N-terminal His (either demonstrated or predicted). For several of these genes the expression and subsequent secretion of the proteins is readily induced by growth on plant biomass (30). Interestingly, some AA9 members have an intriguing substitution of the N-terminal His to an Arg (from here on referred to as Arg-AA9). This can be observed in genomes of species belonging to the *Lentinus* genus, where for example, *Lentinus tigrinis* has 16 AA9s one of which is an Arg-AA9 (https://genome.jgi.doe.gov/mycocosm/proteins-browser/browse;3gQkyP?p=Lentil7_1). Similarly, for *L. similis* at least seven proteins classified as AA9s are described in patent literature (31, 32), and one of these proteins (GH61-5 in Ref. 32 and denoted LsAA9B from this point) has an N-terminal Arg.

This type of His to Arg substitution was first noted by Yakovlev and co-workers (33) in the *Russulales* fungus *Heterobasidion irregulare* in an AA9 denoted HIGH61G (HiAA9G from this point). Yakovlev et al. (33) found that genes encoding six AA9s, including HiAA9G, were up-regulated during growth on lignocellulosic biomass. We have previously reported comprehensive integrative omics studies of the saprotrophic white-rot fungi *Pycnoporus coccineus* BRFM310 (34) and *Polyporus brumalis* BFM985 (35) of the Polyporales order. There we reported co-regulation of glycoside hydrolases (GH) commonly involved plant biomass breakdown (e.g. GH5_5) and Arg-AA9 members (protein ID 1430659 and 1403153 on the JGI Mycocosm public database), which in each species was highly up-regulated during fungal growth on wheat straw (for the latter during solid-state fermentation). In addition, in a recent bioinformatics study (5) it was shown that Arg-AA9s are found in several fungal species almost entirely restricted to one phylogenetic class of largely wood-decaying members namely the *Agaricomycetes*. These Arg-AA9 sequences clustered together.
could be detected for AA9 LPMOs. Thus, perhaps predictably, no LPMO activity
activation sphere of the copper (essential for fully functioning
those forming the copper-binding site and the secondary coor-
differentiation, reveals extensive changes in residues equivalent to
AA9 LPMOs, but in addition to the N-terminal His to Arg sub-
logical role of Arg-AA9s.
starting point for the elucidation of the function and exact bio-
investigation, but the characterization and structure determi-
tional significance of both the N-terminal His to Arg substitu-
tional for these Arg-AA9 proteins, which seem to be strikingly
different compared with canonical AA9 LPMOs. The func-
tional significance of both the N-terminal His to Arg substitu-
tion, the phosphoserine, and the Xyl4 ligand require further
Our results strongly suggest a role related to biomass degra-
dation for these Arg-AA9 proteins, which seem to be strikingly
different compared with canonical AA9 LPMOs. The func-
tional significance of both the N-terminal His to Arg substitu-
tion, the phosphoserine, and the Xyl4 ligand require further
characterization and structure determination of LsAA9B presented here will serve as an important
starting point for the elucidation of the function and exact bio-
logical role of Arg-AA9s.

Results
Phylogenetic distribution and regulation of AA9s with N-terminal Arg

To assess the phylogenetic distribution of fungal AA9 family
sequences with N-terminal Arg (Arg-AA9s), we analyzed AA9
sequences belonging to cluster 63 (5) in the CAZy database (36),
which we found were restricted to species of the Polyporales,
Agaricales, and Russulales orders (all of the Agaricomycetes
class in Basidiomycetes fungi). Notably, the sequence alignment
showed that residues critical for function in canonical AA9
LPMOs (His-1, His-68, His-142, Gln-151, and Tyr-153 in
TtAA9E (14)) were all changed in the Arg-AA9 sequences (Fig.
S1). Interestingly, among the Agaricales and Russulales
sequences a small number displayed N-terminal Lys instead
(and in a few occasions Asn). The Arg and Lys sequences clus-
tered together phylogenetically; meaning that for species of
these orders Arg and Lys are perhaps interchangeable at the
N-terminal position. We found that many Arg-AA9 members
have signal peptides that indicate they would be secreted. As an
example, we found that for the Arg-AA9 of P. coccineus
BRFM310 (JGI protein ID 1430659) the SignalP server (37)
reports with high certainty a putative signal peptide with pre-
dicted cleavage before the Arg, meaning that the protein is most
likely targeted for the secretory pathway.

To augment previous reports on transcriptional regulation of
Arg-AA9—encoding genes during fungal growth on plant bio-
mass (33–35), we sought additional evidence in other fungal
species. Using RNA-seq analysis (as previously described in Ref.
38), we identified a number of genes in several Polyporales
species (with around 70% sequence identity to LsAA9B, Fig. 2),
which were up-regulated when the fungi were cultivated on
cellolose or complex biomass compared with control growth on
maltose as an easily up-takable carbon source (Table 1). We
found Arg-AA9s in the Polyporales fungi Trametes ljubarskyi
BRFM 1659 (JGI protein ID 971189), Leiotrametes sp. BRFM
1775 (JGI protein ID 236437), Pycnoporus sanguineus BRFM
An unusual AA9 naturally lacking the His-brace motif

Table 1
Transcriptional regulation of genes coding for Arg-AA9 in six Polyporales species

<table>
<thead>
<tr>
<th>Organism</th>
<th>JGI Prot ID</th>
<th>Av</th>
<th>As</th>
<th>Pi</th>
<th>WS</th>
<th>Co-regulated genes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. coccineus</em></td>
<td>1430659</td>
<td>110</td>
<td>10</td>
<td>52</td>
<td>110</td>
<td>GH5,5, GH7, AA9, CE16, GH28, AA2</td>
<td>34</td>
</tr>
<tr>
<td><em>T. ljubarskyi</em></td>
<td>971189</td>
<td>675</td>
<td>388</td>
<td>64</td>
<td>256</td>
<td>GH5,5, GH7, GH5_7, GH12, AA9, CE16</td>
<td>This manuscript</td>
</tr>
<tr>
<td><em>P. coccineus</em></td>
<td>248104</td>
<td>14</td>
<td>9</td>
<td>5</td>
<td>6</td>
<td>–</td>
<td>This manuscript</td>
</tr>
<tr>
<td><em>Leiotranetes sp.</em></td>
<td>236437</td>
<td>104</td>
<td>10</td>
<td>3</td>
<td>10</td>
<td>GH1, GH3, GH131</td>
<td>This manuscript</td>
</tr>
<tr>
<td><em>P. sanguineus</em></td>
<td>1740610</td>
<td>78</td>
<td>7</td>
<td>nd</td>
<td>3</td>
<td>GH7, AA9</td>
<td>38</td>
</tr>
<tr>
<td><em>T. elegans</em></td>
<td>360271</td>
<td>16</td>
<td>48</td>
<td>11</td>
<td>34</td>
<td>GH10</td>
<td>This manuscript</td>
</tr>
</tbody>
</table>

Figure 3. Sequence coverage of LsAA9B peptides found in the *L. similis* secretome following growth on hardwood pulp.

1264 (JGI protein ID 1740610), *P. coccineus* BRFM 1662 (JGI protein ID 248104), and *Trametes elegans* BRFM 1663 (JGI protein ID 360271), which were highly up-regulated at day 3 on Avicel, pine, aspen, and wheat straw. In addition, these were co-regulated with several other secreted CAZymes targeting plant biomass polysaccharides (Table 1 and Fig. 2).

These results strongly suggest that these Arg-AA9s may have a role in the fungal adaptive response to (ligno)cellulosic substrates in coordination with the plant cell wall targeting CAZymes that are coexpressed alongside (i.e. GH5_5, GH5_7, and GH12 endoglucanases and endomannanases, GH7 cellobiohydrolases, CE16 deacetylases, GH3 glucosidases, GH28 polygalacturonases, and canonical AA9, Table 1).

**LsAA9B, an AA9 with an N-terminal Arg from *L. similis*, is secreted upon growth on hardwood pulp**

Because the bioinformatics analysis and transcriptomic data suggested that Arg-AA9s might be functionally expressed during fungal growth on biomass, we sought to structurally and functionally characterize one of these proteins. The Polyporales fungus *L. similis* contains at least seven genes encoding proteins that are classified as AA9 of which LsAA9A has been well-characterized (22, 23, 29). Also among these proteins is a single protein ID 360271), which were highly up-regulated at day 3 on Avicel (Av), aspen, (As), pine (Pi), or wheat straw (WS) were compared to normalized read counts after 3 days growth on maltose. Coregulated genes were identified that shared similar transcript levels and transcription profiles on the five carbon sources.


The **LsAA9B X-ray crystal structure reveals a fold very similar to active AA9 LPMOs**

Crystals of LsAA9B were obtained in a range of conditions with combinations of polyethylene glycols (PEGs) (see Table 2 for further details). Structures could be determined by molecular replacement (MR) in space group *P2_12_1_2* to better than 1.60-Å resolution. We obtained structures of both a naturally glycosylated and a partly deglycosylated form of LsAA9B (LsAA9B and LsAA9B_deglyc). In addition, using the glycosy-
Table 2
Crystallization conditions for crystals from which datasets were collected

<table>
<thead>
<tr>
<th>Protein batch</th>
<th>Protein concentration</th>
<th>Primary precipitant</th>
<th>Additive*</th>
<th>Buffer* (pH)</th>
<th>Notes</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>LsAA9B</td>
<td>4 mg/ml (initial)</td>
<td>PPT4</td>
<td>NPS</td>
<td>Imidazole/MES</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>LsAA9B</td>
<td>5 mg/ml (glycosylated)</td>
<td>PPT1</td>
<td>AA</td>
<td>0.1 M (pH 6.5)</td>
<td>–</td>
<td>LsAA9B</td>
</tr>
<tr>
<td>LsAA9B</td>
<td>5 mg/ml (deglycosylated)</td>
<td>PPT4</td>
<td>NPS</td>
<td>Imidazole/MES</td>
<td>–</td>
<td>LsAA9B_deglyc</td>
</tr>
<tr>
<td>LsAA9B</td>
<td>11 mg/ml (glycosylated)</td>
<td>PPT4</td>
<td>CA</td>
<td>HEPS/MOPS</td>
<td>30 min</td>
<td>LsAA9B_metalsoak</td>
</tr>
<tr>
<td>LsAA9B</td>
<td>11 mg/ml (glycosylated)</td>
<td>PPT1</td>
<td>Divalent</td>
<td>HEPS/MOPS</td>
<td>1 h 20 min</td>
<td>LsAA9B_Xyl4</td>
</tr>
</tbody>
</table>

* Primary precipitant mixtures were: PPT1: 2:1 ratio of PEGMME 500 and PEG 20,000, PPT4, 1:1 ratio of MPP (racemic), PEG 1000, PEG 3350. Additives were: amino acids (AA): 20 mM glutamate, alanine, glycine, lysine, and serine (some racemic); NPS: 30 mM of each NaNO3, Na2HPO4, (NH4)2SO4: carboxylic acids (CA): 0.02M of each Na-formate, NH4-acetate, Na-citrate, Na/K-tartrate, Na-oxamate; and divalents: 30 mM each of CaCl2 and MgCl2.

** Table 3
Crystallographic data and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>LsAA9B</th>
<th>LsAA9B_deglyc</th>
<th>LsAA9B_metalsoak</th>
<th>LsAA9B_Xyl4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamsynchrotron</td>
<td>1911–3</td>
<td>1911–3</td>
<td>1911–3</td>
<td>ID22–1</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9900</td>
<td>1.0000</td>
<td>1.3499</td>
<td>0.9760</td>
</tr>
<tr>
<td>Temperature</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Space group</td>
<td>P21_21</td>
<td>P21_21</td>
<td>P21_21</td>
<td>P21_21</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>35.19, 72.48, 78.56</td>
<td>35.08, 72.99, 79.06</td>
<td>35.16, 72.58, 78.68</td>
<td>35.41, 72.79, 78.59</td>
</tr>
<tr>
<td>a = β = γ (%)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Molecules/ASU</td>
<td>20.00–1.60</td>
<td>20.00–1.60</td>
<td>40.00–1.58</td>
<td>40.00–1.40</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.70–1.60</td>
<td>1.70–1.60</td>
<td>1.62–1.58</td>
<td>1.44–1.40</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>10.8 (67.8)</td>
<td>8.4 (82.0)</td>
<td>5.5 (39.1)</td>
<td>10.0 (156.0)</td>
</tr>
<tr>
<td>R/σ(ı)</td>
<td>14.07 (2.74)</td>
<td>14.39 (1.94)</td>
<td>28.86 (3.0)</td>
<td>13.57 (1.36)</td>
</tr>
<tr>
<td>Cc (%)</td>
<td>99.8 (76.4)</td>
<td>99.8 (76.6)</td>
<td>99.9 (82.6)</td>
<td>99.9 (45.4)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>7.83 (6.01)</td>
<td>5.81 (3.59)</td>
<td>8.60 (2.84)</td>
<td>7.17 (7.27)</td>
</tr>
<tr>
<td>No. reflections</td>
<td>62730</td>
<td>35580</td>
<td>249333</td>
<td>292531</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>27139</td>
<td>27139</td>
<td>27805</td>
<td>40817</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>13.67 (22.0)</td>
<td>15.48 (29.9)</td>
<td>11.27 (22.2)</td>
<td>13.31 (29.9)</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>18.22 (27.5)</td>
<td>21.14 (37.5)</td>
<td>17.31 (24.4)</td>
<td>17.51 (34.9)</td>
</tr>
<tr>
<td>No. atoms/B factors (Å²)</td>
<td>1703/14.84</td>
<td>1737/17.41</td>
<td>1740/12.39</td>
<td>1741/19.21</td>
</tr>
<tr>
<td>Protein*</td>
<td>241/30.10</td>
<td>291/31.47</td>
<td>276/29.64</td>
<td>231/31.12</td>
</tr>
<tr>
<td>Solvent*</td>
<td>68</td>
<td>68R6</td>
<td>68R6</td>
<td>68R6</td>
</tr>
</tbody>
</table>

* Glycosylation (N-acetylgalactosamine units, mannosyl units) is included under “Protein.”

** Table 2
Crystallization conditions for crystals from which datasets were collected

Drop sizes were 0.3 μl with protein:reservoir ratios of 3:1.

The LsAA9B batch, we obtained structures after soaking experiments with either a solution containing transition metals (LsAA9B_metalsoak) or a solution containing xylotetraose (LsAA9B-Xyl4). For all structures, all 221 residues corresponding to the mature protein could be modeled in the electron density and the final structures showed good refinement statistics (Table 3). The structure of LsAA9B revealed a typical LPMO topology (Fig. 4e and Fig. S2) consisting of a central immunoglobulin-like β-sandwich (β1:3–9; β2:52–55; β3:59–65; β4:88–94; β5:108–114; β6:134–138; β7:146–156; β8:167–178). The AA structure most closely related to LsAA9B (based on RMSD of the Cα trace calculated with Superpose (39), see Table S2) is NcAA9C (25) (RMSD of 1.04 Å) (PDB entry 4D7U) followed by NcAA9D (40) (PDB entry 4EIR), and LsAA9A (22) (PDB entry 5ACH) all harboring an extended L3 loop (25) implicated in cellulose binding. NcAA9F (PDB entry 4Q18) and in fact TrAA9E (PDB entry 3EJA, used as MR search model) are the most distantly related structures, whereas the TaAA9A (9), NcAA9M (40), and TrAA9B (41) structures (PDB entries 3ZUD, 4EIS, and 2VTC, respectively), which all possess an extended L2 loop (40) (which is absent in LsAA9B) show intermediate RMSDs. However, all are structurally quite similar to LsAA9B with RMSDs of less than 2 Å (Fig. S2 and Table S2).

The LsAA9B structures show striking differences compared with canonical LPMO active sites, but share structural features with AA9 LPMOs active on oligosaccharides

In all the LsAA9B structures the N-terminal Arg (Arg-1) and the residues, Asn-84, Leu-158, and Gln-151, Phe-169 (equivalent to His-1, His-68, His-142, Gln-151, and Tyr-153 in TrAA9E are important for AA9 LPMO activity) were clearly defined in...
the electron density map (Fig. 4b). These residues in LsAA9B (Fig. 4c and Fig. S2) are arranged in a similar configuration compared with NcAA9C (the closest structural match, PDB 4D7U), as defined by the backbone positions, but are not associated with any metal or positioned compatibly with a likely metal binding function (guanidinium group of Arg-1 points away from Asn-84) (Fig. 4). This confirms that residues important for AA9 LPMO activity are changed in LsAA9B, as was indicated by sequence comparison (note that in other LPMO families with demonstrated activity Phe is found in positions equivalent to Phe-169). Furthermore, the structure does not indicate that nearby residues could functionally compensate for the loss of the conserved residues to provide a copper-binding site/LPMO activity.

Near the N-terminal Arg, LsAA9B Tyr-206 is found (Fig. S2) in a position equivalent to Tyr-204 in NcAA9C and Tyr-203 in LsAA9A, shown to be involved in AA9 LPMO oligosaccharide interactions (21–23). In those structures the Tyr residues are well-defined and make hydrogen bonds to the backbone amide of the conserved His of the secondary copper coordination sphere (His-147 in LsAA9A and His-155 in NcAA9C). However, in LsAA9B the rotamer of the apolar Leu-158 side chain interferes with the hydrogen bonding of the backbone amide to Tyr-206, which in many of the structures was not well-defined (Fig. S3, a and b). In addition, LsAA9B has a number of residues, such as Asn-26 and His-66, which are in positions equivalent to those involved in oligosaccharide binding in LsAA9A (LsAA9A Asn-26, His-66, Asn-67, Ser-77, Glu-148, Asp-150, and Tyr-203 are equivalent to LsAA9B Asn-26, His-66, Thr-67, Thr-83, Asn-159, Glu-161, and Tyr-206) (Fig. S2). Almost all of the Arg-AA9 sequences have Asn, His/Arg, and Tyr/Phe in the positions of Asn-26, His-66 (Arg is found in this position in the oligosaccharide active CvAA9A LPMO (23)), and Tyr-206 (Fig. S1). However, for Asn-159 and Glu-161 (equivalent to residues in LsAA9A interacting with the substrate in the minus subsites) there seems to be virtually no conservation (Fig. 5 and Fig. S1).

Other structural features of LsAA9B

O-Linked and N-linked glycosylation at Thr-59 and Asn-134, respectively, were also observed in all of the LsAA9B structures. From Thr-59 one α-linked mannosyl unit could be modeled in all cases. In the LsAA9B structure, two N-acetylglucosamine (GlcNAc) units as well as one poorly defined mannosyl unit could be modeled from Asn-134. In the LsAA9B_degly structure, one well-defined and one less defined GlcNAc unit were modeled, indicating that the protein batch was not fully deglycosylated (as confirmed by MS).
An unusual AA9 naturally lacking the His-brace motif

Near Asn-134 a small pocket was found (mainly formed by Tyr-71 and a GlcNAc unit from the N-glycosylation site) (Fig. 5), which was often occupied by electron density that could be attributed to crystallization condition components (e.g. glycine, sulfate, or a MES molecule). In the LsAA9B_metalsoak structure a MES molecule could quite confidently be modeled in this small pocket (magenta). The MES molecule makes an H-bond to His-65, and makes additional interactions with Tyr-71 and a GlcNAc molecule (N-linked glycosylation at Asn-134). In the Fourier difference map. The highest peaks were found near the sulfur atoms in the Cys residues of the protein, confirming that these metals did not bind in any of these experiments. None of the structures inspected during this work showed any sign of bound metals.

Near Asn-134 a small pocket was found (mainly formed by Tyr-71 and a GlcNAc unit from the N-glycosylation site) (Fig. 5), which was often occupied by electron density that could be attributed to crystallization condition components (e.g. glycine, sulfate, or a MES molecule). In the LsAA9B_metalsoak structure a MES molecule could quite confidently be modeled in this small pocket (magenta). The MES molecule makes an H-bond to His-65, and makes additional interactions with Tyr-71 and a GlcNAc molecule (N-linked glycosylation at Asn-134). In the Fourier difference map. The highest peaks were found near the sulfur atoms in the Cys residues of the protein, confirming that these metals did not bind in any of these experiments. None of the structures inspected during this work showed any sign of bound metals.

Indications of a potential phosphorylation site in LsAA9B

Although it did not reveal any natively bound metals, the PIXE analysis detected the presence of a single atom of phosphorus in two point spectra (Fig. S4 and Table S3) for every protein molecule in the sample. No significant change of $T_m$ was found for a range of other metals by DSC.

As a final step, microPIXE (Proton-induced X-ray Emission) analysis (42) was used to detect any additional elements present in the sample with no added metals, to ascertain whether LsAA9B copurified with any metal ions that might hint at function. The analysis did not reveal significant amounts of copper, iron, or any other metal typically associated with redox enzymes present in the sample above trace levels (Fig. S4 and Table S3) (calculations suggest a lower detection limit for copper of $4 \times 10^{-3}$ atoms per protein molecule).

LsAA9B is not a copper-binding protein

When inspecting the LsAA9B structures no bound metals were found. We speculated whether Arg-1 would be able to adopt a different conformation in response to addition of transition metals allowing for LsAA9B to coordinate a metal cofactor at this surface. Co-crystallization and soaking experiments with transition metals were attempted with solution mixtures containing Fe$^{2+}$, Cu$^{2+}$, Mn$^{2+}$, and Co$^{2+}$. X-ray diffraction data were collected at an appropriate wavelength of 1.35 Å ~ 9.184 keV to ensure an anomalous signal would be obtained to allow location of any bound metal ions in the crystal structure. Following map generation, no peaks that could be interpreted as bound metals could be found when inspecting an anomalous Fourier difference map. The highest peaks were found near the sulfur atoms in the cysteine residues of the protein, confirming that these metals did not bind in any of these experiments. None of the structures inspected during this work showed any sign of bound metals.

Metal binding was also investigated by thermal shift assays with differential scanning fluorimetry (DSF) and differential scanning calorimetry (DSC). For three completely independent measurements with DSF, the $T_c$ for LsAA9B was 65.5 ± 0.75 °C. The presence of copper ions at slightly over stoichiometric amounts (0.048 mM copper acetate) resulted in a decrease of $T_c$ to 62.4 °C. A similar approximate 4 °C decrease of $T_m$ was found with the addition of copper measured by DSC. No significant change of $T_m$ was found for a range of other metals by DSC.

For many of the crystal structures determined during this work, Fourier difference map density was visible proximal to Ser-25. Because additional evidence suggested a putative phosphorylation site in LsAA9B (see “Indications of a potential phosphorylation site in LsAA9B”), the final structure includes a phosphoserine (pSer-25), which fits well the electron density in two of the four structures (LsAA9B_deglyc and LsAA9B_metalsoak). In these structures, the pSer-25 and Ser-25 are modeled in alternative conformations with 70 and 30% occupancy, respectively (Fig. 4b).

Near Asn-134 a small pocket was found (mainly formed by Tyr-71 and a GlcNAc unit from the N-glycosylation site) (Fig. 5), which was often occupied by electron density that could be attributed to crystallization condition components (e.g. glycine, sulfate, or a MES molecule). In the LsAA9B_metalsoak structure a MES molecule could quite confidently be modeled in this small pocket (magenta). The MES molecule makes an H-bond to His-65, and makes additional interactions with Tyr-71 and a GlcNAc molecule (N-linked glycosylation at Asn-134). In the Fourier difference map. The highest peaks were found near the sulfur atoms in the Cys residues of the protein, confirming that these metals did not bind in any of these experiments. None of the structures inspected during this work showed any sign of bound metals.

Metal binding was also investigated by thermal shift assays with differential scanning fluorimetry (DSF) and differential scanning calorimetry (DSC). For three completely independent measurements with DSF, the $T_c$ for LsAA9B was 65.5 ± 0.75 °C. The presence of copper ions at slightly over stoichiometric amounts (0.048 mM copper acetate) resulted in a decrease of $T_c$ to 62.4 °C. A similar approximate 4 °C decrease of $T_m$ was found with the addition of copper measured by DSC. No significant change of $T_m$ was found for a range of other metals by DSC.

As a final step, microPIXE (Proton-induced X-ray Emission) analysis (42) was used to detect any additional elements present in the sample with no added metals, to ascertain whether LsAA9B copurified with any metal ions that might hint at function. The analysis did not reveal significant amounts of copper, iron, or any other metal typically associated with redox enzymes present in the sample above trace levels (Fig. S4 and Table S3) (calculations suggest a lower detection limit for copper of $4 \times 10^{-3}$ atoms per protein molecule).

Indications of a potential phosphorylation site in LsAA9B

Although it did not reveal any natively bound metals, the PIXE analysis detected the presence of a single atom of phosphorus in two point spectra (Fig. S4 and Table S3) for every protein molecule in the sample. No significant change of $T_m$ was found for a range of other metals by DSC.

As a final step, microPIXE (Proton-induced X-ray Emission) analysis (42) was used to detect any additional elements present in the sample with no added metals, to ascertain whether LsAA9B copurified with any metal ions that might hint at function. The analysis did not reveal significant amounts of copper, iron, or any other metal typically associated with redox enzymes present in the sample above trace levels (Fig. S4 and Table S3) (calculations suggest a lower detection limit for copper of $4 \times 10^{-3}$ atoms per protein molecule).
An unusual AA9 naturally lacking the His-brace motif

Following this observation, we tried to identify phosphorylation using MS. Making several measurements of the Endo H-deglycosylated LsAA9B batches with matrix-associated laser desorption/ionization time of flight (MALDI-TOF) we identified the largest peaks with m/z values that could correspond to LsAA9B with glycosylation (N-GlcNAc and O-Man) and a phosphoform (23,255 Da + 203 Da + 162 Da + 80 Da = 23,670 Da). Using MALDI-TOF to measure the glycosylated LsAA9B we found that the m/z values could correspond with a single phosphorylation, two GlcNAc and varying degrees of mannosylation (m/z increasing by 162 Da for every mannosyl unit). MS techniques with a trypsin-digested preparation of LsAA9B only identified the nonmodified peptide (total sequence coverage of 55%), thus conclusive evidence for phosphorylation is lacking.

However, it is interesting to observe that in positions corresponding to Ser-24 and Ser-25 in LsAA9B most of the analyzed Arg-AA9 sequences have Ser or Thr (Figs. 4 and 5, and Fig. S1), which are potential phosphorylation sites, as predicted by the NetPhos 3.1 Server (43). Thus, at least theoretically, it seems that a number of the Arg-AA9 sequences could be phosphorylated.

Investigation of oligosaccharide binding

As already stated LsAA9B contains residues (e.g. Asn-26, His-66, and Tyr-206) equivalent to some of those involved in LsAA9A oligosaccharide binding (Fig. S2). When superimposing either the LsAA9A-Cell5 (PDB 5NLS) or LsAA9A-Xyl5 (PDB 5NLO) complex onto the LsAA9B structure, we found that some of these residues in LsAA9B came within hydrogen bonding distance of the Cell5 and Xyl5 ligands of the LsAA9A complex structures. In addition, we observed that the sugar moiety of the +2 subsite in LsAA9A (the C6-hydroxyl and the C2-hydroxyl for Cell5 and Xyl5, respectively) could interact with the putative phosphorylation in LsAA9B. Thus, LsAA9B carbohydrate binding was investigated by cocrystallization with monosaccharides and crystal soaking experiments and thermal shift assays with oligosaccharides.

Addition of oligosaccharides at 50 mM concentration in solution caused very little change in T_m (66.1 °C with maltopentaose; 65.9 °C with Cell5; 64.7 °C for Xyl5; 65.0 °C for chitopentaose). Addition of stoichiometric amounts of copper (with or without 200 mM NaCl as described in Ref. 22) decreases the T_m to 62.4 °C (T_m of 61.9 °C in the presence of NaCl), still with no upward thermal shift on addition of oligosaccharides (T_m values of 59.9–61.4 °C with copper alone and T_m values of 61.3–62.2 °C with copper and NaCl). No crystals appropriate for data collection were obtained from cocrystallization experiments. Soaking experiments with cellotetraose (Cell4), cellopentaose (Cell5), or xylotetraose (Xyl4) oligosaccharides were done in conditions with MgCl_2 and CaCl_2 (30–100 mM) because the presence of chloride ions has been shown to enhance oligosaccharide binding to canonical AA9 (22). No electron density that could be interpreted as cellooligosaccharides was identified at the protein surface.

However, we did identify a Xyl4 ligand bound at the surface of LsAA9B, but the binding mode is distinct compared with LsAA9A-Xyl5 (Fig. 6a). One explanation for this could be that glycosylation of a symmetry related molecule occludes this binding site (Fig. 6b). Instead the Xyl4 ligand bound in a small cleft formed by residues 41–49 (TGFIQPVSK) and 99–10 (TQS). Three xylosyl units of the Xyl4 ligand (numbered 1–3 from the reducing end) are well-defined and could easily be modeled in the electron density, whereas the fourth xylosyl unit at the nonreducing end is only partially defined (Fig. 6c). Many of the glycosidic torsion angles of the Xyl4 ligand deviate from ideal values, e.g. between xylosyl units 1 and 2 (in particular the Ψ-torsion angle) (Table S4).

Discussion

LsAA9B can be heterologously expressed and is a well-folded and stable protein, with melting temperatures comparable with enzymatically active members of the family (44–47), but showed no detectable LPMO activity under the condition tested. The X-ray crystal structure of LsAA9B in fact confirms a complete disruption of the copper-binding site that is the hallmark of LPMOs (9) and absence of other metal-binding sites, even though the overall-fold remains very similar to that expected. We clearly demonstrated that residues Arg-1 and Asn-84 in LsAA9B are positioned equivalently to the His-brace of AA9 LPMOs, but that neither of these (nor any other) residues coordinate copper. In all structures (with or without the putative phosphorylation site modeled and/or metals added), Arg-1 in fact assumes a rotamer conformation that is not optimal for forming a His-brace like metal-binding site. PIXE spectroscopy (42), crystal soaking experiments with metal ions, and thermal shift analyses in solution failed to indicate binding of any metal. This is consistent with a survey of metal-binding sites in protein structures reporting that Asn and Arg interact with elements such as potassium, sodium, magnesium, and calcium more frequently than with transition metals (of which only manganese is reported), whereas in contrast His is the most common protein residue to bind copper (48). Thus, our work confirms the expectation that LsAA9B and other Arg-AA9s are very unlikely to be copper-dependent LPMOs.

At the same time, evidence presented here and elsewhere shows that a number of genes (in wood-degrading fungal Polyporales species) encoding Arg-AA9 proteins are up-regulated during fungal growth on plant biomass alongside lignocellulose-degrading CAZymes (most commonly members of GH and AA families). The majority of the Arg-AA9 sequences possess a signal peptide (with predicted processing before Arg) explicitly targeting the mature N-terminal Arg proteins for the secretory pathway. Indeed both LsAA9B and the Arg-AA9 of P. brunalis BRFM 985 (GI protein ID 1403153) have been detected in fungal secretome data. In addition, the mature LsAA9B protein with N-linked glycosylation is secreted when recombinantly produced in Aspergillus from a construct with the native L. simillis signal peptide. Thus, there is circumstantial evidence that Arg-AA9 proteins, like conventional LPMOs, could be involved in plant biomass degradation. The open question remains how, and although this work does not provide a conclusive answer, the structural analysis suggest a number of new avenues for further investigation.

For example, we found that the relatively conserved L3 loop is likely structurally rigid in the majority of Arg-AA9 proteins.
and together with the Asn-134 glycosylation forms a small pocket. In the \textit{LsAA9B} _metalsoak_ structure, a MES molecule is interacting with the completely conserved His-65, Tyr-71 in this pocket. It is possible that the MES molecule mimics biologically relevant interactions and notably, MES bear some resemblance to phenolic compounds (e.g. \textit{p}-coumaryl-coniferyl, sinapyl alcohols) that make up lignin, and thus the possible connection to lignin degradation/detoxification pathways should be investigated in the future.

Structural comparison suggested that \textit{LsAA9B} could bind oligosaccharides in a manner similar to \textit{LsAA9A}, but this could not be confirmed by our structural studies, perhaps because of glycosylation from a symmetry-related molecule occluding this binding site. In contrast, we obtained an \textit{LsAA9B} structure with a Xyl4 ligand bound at a distinct binding site, in a small cleft made up mostly by relatively conserved residues (no similar cleft is present in \textit{LsAA9A} where this space is occupied by the peptide stretch VDNRVV formed by residues 43 to 48). It is difficult to establish whether the binding site is biologically relevant: the expected structural conservation of this region in \textit{Arg-AA9} and the fact that there is only one hydrogen bond of Xyl4 with a symmetry-related molecule suggest this is not a crystal artifact. On the other hand the bound conformation of Xyl4 deviates considerably from what is usually observed in xylooligosaccharides and xylan (49) (Table S4), and Xyl5 in solution caused no thermal shift. The path connecting this binding site and the binding site in \textit{LsAA9A} is definitely occluded by crystal contacts, and thus binding of longer oligosaccharides (connecting the two binding sites) is not possible in this crystal form. However, the discovery of a Xyl4-binding site warrants additional investigations of polysaccharide binding by this and other Arg-AA9.

Finally, we found indications that Ser-25 could be phosphorylated in \textit{LsAA9B} as supported by some high-resolution structures and by PIKE data, and consistent with MALDI-TOF analysis, although a phosphorylated peptide could not be conclusively identified. Intriguingly, the phosphoserine modeled in the structure interacts with the N-terminal Arg (2.9 Å hydrogen bond) and could additionally potentially interact with bound polysaccharides near the binding sites observed for \textit{LsAA9A} and \textit{LsAA9B}.

Potential similar phosphorylation sites (equivalent to Ser-24 and Ser-25 of \textit{LsAA9B}, Figs. 4b and 5d) are also predicted for Arg-AA9 protein sequences (presented here in Fig. 2 and Table 1 and previously (34, 35, 38)), for which the transcription of the corresponding genes were up-regulated during fungal growth on plant biomass. In addition, in the Fungi Phosphorylation Database (FPD) (50) documented Ser phosphorylation sites in fungal species (including \textit{Aspergillus}) can be found that are consistent with the \textit{LsAA9B} sequence surrounding the Ser-25 phos...
An unusual AA9 naturally lacking the His-brace motif

Phosphorylation, supporting the notion that phosphorylation could play a role in fungal plant biomass degradation. It is known that proteins can be phosphorylated in the secretory pathway (51, 52), in mammalian cells by the Fam20C kinase (53, 54), and phosphorylation does appear to play a role extracellularly during some biological events (e.g. during microbial host infection (55–57)). Thus, it is possible that also phosphorylated Arg- AA9s could exist extracellularly, because the evidence points to an extracellular location for these proteins.

In summary, our results indicate the importance of Arg-AAs in fungal adaptation to biomass degradation. We have shown that the transcription of Arg-AA9 genes is up-regulated in fungal species belonging to the phylogenetic class of Agaricomycetes (of which many are wood or litter decayers) when cultivated on plant biomass. We have determined this first structure of an Arg-AA9 protein, LsAA9B (belonging to the wood-decaying Lentinus genus (58)), and have identified interesting structural features, including a potential carbohydrate-binding cleft, a small conserved pocket, and a potential phosphorylation site. Sequences conservation suggests that these features are found in the majority of the Arg-AA9s as well. These findings provide a solid framework for future investigations of this subgroup of AA9 and for the pursuit of their potential role in fungal biology. The original structures of AA9 (14) and AA10 (24) family members were of key importance toward the subsequent identification of the nature of LPMO action on biomass. We hope that the structure of LsAA9B will likewise act as the starting point for further characterization of these unusual AA9 family members.

Experimental procedures

Fungal transcriptomic and secretomic data

L. similis was cultivated at 30 °C on a basic fungal media including 0.5% (w/v) hardwood BCTMP (bleached chemithermo mechanical pulp) for induction experiments. Samples were taken after 0, 3, 5, 7, and 10 days to analyze secreted proteins and identify LsAA9B with ESI-MS/MS. Tryptic digests were prepared by a filter-aided sample preparation method. Following digestion the extracted peptides were analyzed on a nano LC-MS/MS system: UltiMate 3000 RSLC nano/LTQ Orbitrap Velos Pro (Thermo/Dionex). For protein identification the data were searched against the complete L. similis proteome (internal database) using the Mascot search engine (Matrix science) in the Genedata Expressionist software (1% false discovery rate cutoff). Relative protein concentrations were calculated by label-free quantification from peptide volumes using a Hi3 standard method in Genedata Expressionist.

P. coccineus BRFM 1662, T. ljubariskyi BRFM 1659, Leio-trametes sp. BRFM 1775, and T. elegans BRFM 1663 strains were obtained from the CIRM collection at the National Institute of Agricultural Research. Transcriptome and secretome data were collected from triplicated independent 3-day cultures in the presence of 20 g x liter⁻¹ of maltose, 15 g x liter⁻¹ of Avicel, 15 g x liter⁻¹ of ground wheat straw, 15 g x liter⁻¹ of ground pine wood, or 15 g x liter⁻¹ of ground aspen wood as the sole carbon source. RNA libraries were prepared and sequenced on Illumina HighSeq-2500 as described in Ref. 59. Secreted proteins were collected from the same cultures, dialyzed and identified by ESI-MS/MS (as in Ref. 59). Transcript reads were analyzed as described in Ref. 38. For each strain, the genes with similar transcription profiles on the five carbon sources were grouped into nodes using the Self-organizing maps Harboring Informative Nodes with Gene Ontology Pipeline (SHIN+GO (34)). All sequence data are available on the Mycosom public database at Joint Genome Institute (https://genome.jgi.doe.gov/programs/fungi) (60).

LsAA9B protein production, purification, and activity measurements

An AA9 protein from L. similis (LsAA9B) was recombinantly produced in A. oryzae MT3568 and purified as described in the patent literature (32) (similar to the previously characterized LsAA9A (22, 31)). The mycelium was removed by filtration and the broth collected for protein purification by chromatography with a 50-mm diameter 167-ml Butyl-ToyoPearl 650 column (ToSoh BioSciences, Stuttgart, Germany) with a gradient of 0–100% of buffer A (25 mM Tris-HCl, 1.0 M ammonium sulfate, pH 7.5) and buffer B (25 mM Tris-HCl, pH 7.5). To ensure purity suitable for structural studies, two further purification steps were applied. After buffer change by ultrafiltration (Vivaspin, Sartorius) the sample was applied to a 100-ml Q-Sepharose column (Sigma) in 20 mM Tris-HCl, pH 8.0. The protein was eluted with a gradient of 0 to 500 mM NaCl. Fractions containing protein were further purified on a 26-mm Superdex 75 size exclusion (Sigma) column with isoosmotic 20 mM MES, 125 mM NaCl, pH 6.0. All protein batches were buffered in 20 mM MES, pH 6.0. Activity was measured as previously described by depolymerization of AZCL-cellulose (61) and PASC (9) in parallel on LsAA9A and LsAA9B in the presence of a variety of electron donors (pyrogallol, 4-OH-5-CH3–3-furanone, ascorbate, and cysteine) and under similar conditions. In addition, LsAA9B activity was measured (according to manufacturer instructions) on AZO-xylan (purchased from Megazyme; product code S-AXBL) without and with equimolar concentrations (1 μM) of copper or manganese in the presence of ascorbate (1 mM).

Mass spectrometry analysis of purified LsAA9B

MS was performed using MALDI-TOF. The matrix was prepared by saturating a TA solution (0.1% (v/v) trifluoroacetic acid (TFA) and acetonitrile in a 2:1 (v/v) ratio) with sinapic acid. Matrix and protein sample were sequentially applied to the target plate (in a final 2:1 ratio) and left for solvent to evaporate. The TOF experiments were performed in linear mode calibrated for 10–50 kDa molecules.

Products following a trypsin digest of the heterologously expressed LsAA9B batches were analyzed using a Thermo Fisher Scientific Q-Exactive HF-X Orbitrap mass spectrometer that was operated in positive mode with a MS1 resolution of 120,000. A top 10 method was utilized and a resolution of 45,000 for MS2 fragment scans was utilized. Peptides were separated on a 15-cm column (75 μm inner diameter) packed in-house with 1.9 μm C18 particles. EASY-nLC 1200 nano-LC system coupled to the mass spectrometer was utilized to separate injected peptides over an increasing gradient of buffer B (80%...
acetonitrile, 0.1% formic acid). The collected raw data were analyzed using MaxQuant (version 1.6.1.11).

**Crystallization and data collection**

Deglycosylation of protein for crystallization was performed in 20 mM MES, pH 6.0, 125 mM NaCl, by incubating about 10 μl of endoglycosidase H (from Roche Diagnostics, 11643053001) per mg of LsAA9B protein overnight at room temperature. Crystallization was carried out by sitting drop vapor diffusion in 96-well MRC-2 plates using an Oryx8 crystallization robot (Douglas Instruments) with drop sizes of 0.3–0.4 μl (with protein:reservoir:Milli-Q water ratios of 3:1:1, 3:1:0, 1:1:1, or 1:1:0) and with 100-μl reservoirs in 24-well VDX plates with drop sizes of 2–4 μl and reservoirs of 1 ml. Morpheus screens (62) were set up using both a naturally glycosylated batch (11 mg/ml) and a deglycosylated (with endoglycosidase H from Roche Diagnostics, 11643053001) batch (4–5 mg/ml) of LsAA9B. For all batches, crystals could be obtained in quite similar conditions (Table 2).

With an initial protein batch a crystal grew in Morpheus condition 28 in well C4 (Table 2), from which a complete 1.6 Å resolution X-ray diffraction dataset of 155 frames (155°) was collected. With a glycosylated LsAA9B, intergrown crystal plates were obtained in Morpheus condition 85 (Table 2). A single plate was mounted separately from a cluster and a dataset of 200 images (200°) was collected. Crystals were also obtained with a deglycosylated LsAA9B (Morpheus condition 6 in well A6) and were reproducible in MRC-2 plates with protein concentrations from 2 to 5 mg/ml and 24–48% (w/v) PPT4 (a 1:1:1 mixture of MPD/racemic, PEG100, and PEG3350), and a dataset of 324 images (162°) was collected (Table 2). For all datasets the data collection were carried out with the crystals mounted in nylon loops at 100 K without additional cryoprotectant, because the Morpheus screen composition is already cryo-protecting.

Cocrystallization experiments with transition metals were performed in VDX plates with 33–42% (w/v) of PPT1 (a 2:1 mixture of PEGMMME500 and PEG20,000) or PPT4 at pH 6.5, 7.5, or 8.5 and a metal mixture additive of 5–50 mM CuSO4, MnCl2, and CoCl2, but failed to provide crystals of sufficient quality. For soaking experiments, crystals were grown in PPT4 at pH 7.5 (with either carboxylic acids or amino acids or monosaccharides additives). Transition metal mixture solutions were added to an approximate final concentration of 6–7 mM of each of Fe2+, Cu2+, Mn2+, and Co2+. Crystals were mounted at different time points from 1 min to 5 h 20 min and datasets (generally to better than 1.7 Å resolution) were collected as before but at a wavelength of 1.35 Å (9.184 keV) at which the anomalous signal should be significant to detect all of the metals. No peaks that could account for bound metals were found when inspecting an anomalous Fourier difference map. In fact, the highest peaks were found near the sulfur atoms in the cysteine residues.

Optimized cocrystallization with monosaccharides additives (20–100 mM) were carried out in VDX plates with 33–42% (w/v) of either PPT1 or PPT4 at pH 6.5, 7.5, or 8.5 (because intergrown protein crystals in conditions with monosaccharide additives were observed in the screens). However, no crystals appropriate for data collection were obtained. Soaking experiments with oligosaccharides were carried out in MRC-2 plates using crystals grown with 6–11 mg/ml of protein in 27–42% (w/v) PPT1 at either pH 6.5 or 8.5 with additives of either 30 or 100 mM MgCl2 and CaCl2. Crystals were transferred to reservoir solutions containing 400–800 mM of cellotetraose (Cell4), cellopentaose (Cell5), or xylotetraose (Xyl4) and soaked for 40–80 min. All data were collected at the MX beamline I911-3 at the MAX-IV laboratory, Lund (Sweden), or by remote access at beamline ID23-1 at the ESRF, Grenoble (France). All crystals were isomorphous and processed and scaled with XDS/XSCALE (63) in P2_1_2_1 with similar cell dimensions (Table 3).

**Structure determination and refinement**

From sequence alignment the closest AA9s with structures available at the time were TlAA9E (PDB entry 3EJA (14)) and NcAA9D (PDB entry 4EIR (40)) with 42.9 and 45.1% sequence identity, respectively. MR with MOLREP (64) using only the protein coordinates (or Sculptor (65) modified models) of either of PDB entries 3EJA or 4EIR as search models gave solutions, which followed one round (10 cycles) of restrained refinement (with Refmac5 (66), CCP4 suite (67)) resulted in R-factors of ~40% after which parts of the map could be relatively easily interpreted. Modelling in COOT (68, 69) followed by several rounds of restrained refinement resulted in an initial preliminary structure, which was used to refine additional structures obtained later. Structures of LsAA9B were refined in Refmac5 (CCP4 suite) (66, 67) using Rfree flags imported from the previous structure factor file. All LsAA9B structures obtained after this point were refined with anisotropic B factors for all protein atoms (including glycosylation) and isotropic B factors for all other atoms. Refinement and validation statistics are shown in Table 3.

**Construction of multiple sequence alignments and structural comparison**

Superpose (CCP4 suite) was used for calculation of the RMSD Cα trace using secondary-structure matching (39, 67). Multiple sequence alignment was constructed using a standalone version of STRAP (STRuctural Alignments of Proteins) and programs within (70, 71). Multiple sequence alignment, MAFFT was used (72). For the structure-based sequence alignment TMalign was used to align the Cα trace of the AA9 3D structures (73). Mapping of sequence on the LsAA9B structure was done using CAMPO (74) available in the PyMod 2.0 plugin for PyMOL (75).

**Thermal shift assays**

Thermal shift assays with DSF were carried out using intrinsic fluorescence using the NanoTemper Tycho NT.6 (Nano Temper Technologies) according to the manufacturer’s instructions. Tycho NT.6 follows the unfolding process by recording sample fluorescence at 330 and 350 nm during thermal unfolding. A constant heating rate of 30 °C/min is applied to the sample, heating from 35 to 95 °C. A shift in Tm to higher temperatures is usually indicative of ligand binding. 10 μl of each sample were prepared, incubated for 5 min, and then loaded in capillaries for measurements. LsAA9B was at a concentration of 1 mg/ml in 20 mM MES, pH 6.0. All measurements were carried out at least in triplicates. Thermal shift assays with capillary DSC (Northampton, MA) were carried out with a
heating rate of 1.5 °C/min with LsAA9B at pH 5.0 in the presence of 1 mM divalent metal ions (Cu²⁺, Ni²⁺, Fe²⁺, Zn²⁺, Mn²⁺, and Ca²⁺) or 1 mM DTPA (as chelating agent).

**Particle-induced X-ray emission (microPIXE) analysis**

LsAA9B, purified as described above, was prepared for microPIXE analysis by passing down a 16/600 Superdex 75 (GE Healthcare) column to buffer exchange into 20 mM ammonium acetate, pH 5.5, 200 mM KBr. The final sample was then concentrated to 7.7 mg/ml using a 10-kDa cut off VivaSpin Concentrator (Sartorius).

0.1 µl of sample was placed onto a 4-µm polypropylene film (Prolene supplied by Fluxana GmbH & Co., Germany) which was then allowed to dry naturally in a closed environment to prevent dust contamination. The sample was mounted in the path of a 2.5-MeV proton beam with a beam diameter of ~2.5 µm at the Ion Beam Centre, University of Surrey. Prior to the experiment a glass standard was analyzed and validated to ensure accurate elemental quantitation. Data were collected by first scanning the proton beam across the sample to generate a set of elemental maps. Point spectra were then measured from two distinct regions of the sample based on the distribution of sulfur in the map, signifying where protein was located. The collected spectra were processed using the Q-factor method (76) as implemented in the OMDAQ-3 software.

A standard was analyzed and validated to ensure accurate elemental quantitation. Data were collected by first scanning the proton beam across the sample to generate a set of elemental maps. Point spectra were then measured from two distinct regions of the sample based on the distribution of sulfur in the map, signifying where protein was located. The collected spectra were processed using the Q-factor method (76) as implemented in the OMDAQ-3 software.

**References**


function of a large, enigmatic family. *Biochemistry* **49**, 3305–3316


21. Courtade, G., Wimmer, R., Rohr, Å. K., Preims, M., Felice, A. K., Dimaro-Courtade, G., Wimmer, R., Rohr, Å. K., Preims, M., Felice, A. K., Dimaro-
An unusual AA9 naturally lacking the His-brace motif


An unusual AA9 naturally lacking the His-brace motif


Insights into an unusual Auxiliary Activity 9 family member lacking the histidine brace motif of lytic polysaccharide monooxygenases

doi: 10.1074/jbc.RA119.009223 originally published online August 30, 2019

Access the most updated version of this article at doi: 10.1074/jbc.RA119.009223

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 74 references, 15 of which can be accessed free at http://www.jbc.org/content/294/45/17117.full.html#ref-list-1