Fungal lytic polysaccharide monooxygenases bind starch and -cyclodextrin similarly to amylolytic hydrolases

Nekiunaite, Laura; Isaksen, Trine; Vaaje-Kolstad, Gustav; Abou Hachem, Maher

Published in:
F E B S Letters

Link to article, DOI:
10.1002/1873-3468.12293

Publication date:
2016

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

FEBS Letters

FEBS PRESS science publishing by scientists

Fungal lytic polysaccharide monooxygenases bind starch and β-cyclodextrin similarly to amylolytic hydrolases

Laura Nekiunaite[1], Trine Isaksen[2], Gustav Vaaje-Kolstad[2] and Maher Abou Hachem[1]

1 Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark
2 Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway

Correspondence
M. Abou Hachem, Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Elektrovej, Building 375, DK-2800 Kgs. Lyngby, Denmark
Fax: +45 4588 6307
Tel: +45 4525 2732
E-mail: maha@bio.dtu.dk

(Received 27 April 2016, revised 30 June 2016, accepted 2 July 2016, available online 26 July 2016)

doi:10.1002/1873-3468.12293

Edited by Judit Ovádi

Starch-binding modules of family 20 (CBM20) are present in 60% of lytic polysaccharide monooxygenases (LPMOs) catalyzing the oxidative breakdown of starch, which highlights functional importance in LPMO activity. The substrate-binding properties of starch-active LMPOs, however, are currently unexplored. Affinities and binding-thermodynamics of two recombinant fungal LPMOs toward starch and β-cyclodextrin were shown to be similar to fungal CBM20s. Amplex Red assays showed ascorbate and Cu-dependent activity, which was inhibited in the presence of β-cylodextrin and amylose. Phylogenetically, the clustering of CBM20s from starch-targeting LPMOs and hydrolases was in accord with taxonomy and did not correlate to appended catalytic activity. Altogether, these results demonstrate that the CBM20-binding scaffold is retained in the evolution of hydrolytic and oxidative starch-degrading activities.

Keywords: AA13; carbohydrate-binding module; CBM20; lytic polysaccharide monooxygenase; starch binding; β-cyclodextrin

Starch is a major renewable energy storage polysaccharide in plants and an important resource not only as a food but also as an industrial feedstock in biofuels, pharmaceuticals, detergents, and cosmetics [1–3]. Starch consists of two types of homo-glucose polymers: the mainly linear α-1,4-linked amylose and amylopectin, constituting 65–82% (w/w) of the starch granule and differing from amylose by having a larger molecular mass and roughly 5% α-1,6-branches of 12–15 glucosyl units long on average [2,4,5]. Starch is biosynthesized as insoluble granules, varying in size, morphology, crystal packing, and crystallinity that ranges from 15 to 45% depending on botanical origin [6–8]. Radially alternating amorphous and semicrystalline layers in the starch granule arise from the packing of double helices formed by adjacent branches in amylopectin, with the semicrystalline regions contributing to resistance of starch to enzymatic degradation [9,10]. Many industrial applications require the disruption of starch granules through hydrothermal, harsh chemical, or enzymatic treatments [11–13]. Despite development of relatively efficient α-amylases and other starch-degrading enzymes, there is still a significant margin for improving starch hydrolysis yields and shortening processing time, which would significantly reduce energy and costs of the process [14,15].

Typically, glycoside hydrolases (GHs) that degrade complex polysaccharides possess carbohydrate-binding modules (CBMs) that promote enzyme-substrate proximity and thereby enhance catalytic efficiency [16,17]. Moreover, CBMs can also modulate the specificity and activity of cognate enzymes against plant cell wall

Abbreviations
AA, auxiliary activity; CAZy, carbohydrate-active enzymes; CBM, carbohydrate-binding module; DSC, differential scanning calorimetry; GH, glycoside hydrolase; ITC, isothermal titration calorimetry; LPMO, lytic polysaccharide monooxygenase; SBS, starch-binding site; SDS/PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; β-CD, β-cyclodextrin.

polysaccharides [18,19]. Similarly, starch-specific CBMs, which are assigned into CBM20 in the CAZy database [20,21], are reported to potentate the activity of fungal amylolytic enzymes, for example, α-amylases and glucoamylases on granular starch [22,23]. Catalytic efficiency gains have been also conferred to amylolytic enzymes by fusion to CBM20s, attesting the importance of these ancillary modules in the deconstruction of starch [24,25].

Lytic polysaccharide monooxygenases are copper-dependent enzymes that use molecular oxygen and an external electron donor to cleave glycosidic bonds in various polysaccharides, such as cellulose [26,27], hemicelluloses [28,29], and chitin [30]. These enzymes are assigned into auxiliary activity (AA) families 9, 10, 11, and 13 in the CAZy database. Similar to other complex polysaccharide active enzymes, CBMs occur frequently (~ 30%) together with LPMO catalytic modules [31]. Knowledge on CBMs occurring with LPMOs, however, is scarce. Recently, modular LPMOs comprising AA13 catalytic modules joined to C-terminal starch-binding CBM20 have been shown to be active on starch [32,33]. No activity, however, could be demonstrated for the truncated *Aspergillus oryzae* enzyme (*Ao*AA13) lacking the CBM20 that is present in most enzymes from this family. The purification of AA13 enzymes using amylose-affinity columns demonstrates their affinity to starchy ligands, but currently there are no data on their binding properties.

In this study, we have analyzed the binding of two AA13 enzymes from *Aspergillus terreus* and the cereal pathogen, *Magnaporthe oryzae*, to starch and the model ligand, β-cyclodextrin, which is commonly used as a starch mimic substrate. Our data establish the ability of starch-active AA13 LPMOs to bind starch granules and β-cyclodextrin with comparable affinities to typical amylolytic hydrolases, which highlights the common and important function of CBMs in granular starch degradation in both types of enzymes.

Materials and methods

Cloning, production, and purification of recombinant enzymes

The genes (cDNA) encoding lytic polysaccharide monooxygenases of auxiliary activity 13 (AA13) family from *M. oryzae, Mo*LPMO13A (UniProt: Q2KEQ8), and from *A. terreus*, *At*LPMO13A (UniProt: Q0CGA6) were synthesized by GenScript (Piscataway, NJ, USA). The synthetic *Mo*LPMO13A gene was inserted into the pPICZα A vector (Invitrogen, Carlsbad, CA, USA) using BstBI and XbaI sites, and *At*LPMO13A using XhoI and XbaI sites. These plasmids were linearized with PmeI and transformed into electrocompetent *Pichia pastoris* X-33 cells (Invitrogen) by electroporation, and selected on YPDS plates supplemented with 100 $\mu g \cdot mL^{-1}$ zeocin following manufacturer's instructions (EasySelect™ Pichia Expression Kit; Invitrogen). Transformants were screened for protein production in BMGY medium containing 1% (v/v) glycerol and best secreting transformants were used to produce *Mo*LPMO13A and *At*LPMO13A in a 5-L Biostat B bioreactor (Sartorius Stedim Biotech, Goettingen, Germany) as previously described [34]. The culture supernatants were recovered by centrifugation (14 000 ***g***, 45 min, 4 °C) and filtered using 0.45-μm membrane filters (Millipore, Bedford, MA, USA). Both proteins were purified by affinity chromatography using β-cyclodextrin (β-CD) sepharose, followed by size-exclusion chromatography. The cell-free supernatant was supplemented with $(NH_4)_2SO_4$ to 0.5 M, followed by centrifugation (15 000 ***g***, 15 min, 4 °C), and refiltration before loading onto a 20-mL β-CD sepharose [35] column and purification as previously described [36]. The elution fractions containing the purified protein were pooled, concentrated, and loaded onto a HiLoad 16/60 Superdex G-75 size-exclusion column (GE Healthcare, Uppsala, Sweden) and eluted in 10 mM Na Acetate, 150 mM NaCl, pH 5.5 at 1 $mL \cdot min^{-1}$. Chromatographic steps were performed using an ÄKTA Explorer chromatograph (GE Healthcare) at 4 °C. Protein purity was analyzed by SDS/PAGE, and the fractions containing pure protein were pooled and concentrated with Amicon Ultra centrifugal filters (MWCO 10 kDa; Millipore). Protein concentrations were determined by absorbance A_{280}, using the theoretical extinction coefficients calculated using ExPASy server (*Mo*LPMO13A: 71 360 $M^{-1} \cdot cm^{-1}$, and *At*LPMO13A: 82 360 $M^{-1} \cdot cm^{-1}$) [37].

Mass spectrometry (MS) analysis was employed to verify correct processing of signal peptide of *Mo*LPMO13A. Purified *Mo*LPMO13A was precipitated by the addition of 100% acetone (1 : 4 vol) and incubation (−20 °C, 2 h) followed by centrifugation (10 000 ***g***, 10 min, 4 °C). The *Mo*LPMO13A pellet was dried, redissolved (100 μL, 20 mM Tris-HCl, pH 8), reduced by incubation with dithiotreitol (10 mM, room temperature, 30 min). Alkylation was performed by reaction with iodoacetamide (15 mM, room temperature, 30 min). The protein was digested with 20 μL 12.5 $ng \cdot \mu L^{-1}$ sequencing-grade modified trypsin (Promega, Madison, WI, USA) and incubated at 37 °C for 16 h. Trypsination was stopped by addition of trifluoroacetic acid (TFA) to 0.5% (v/v) and the sample was dried in a SpeedVac and purified with ZipTip C_{18} column. Samples (1 μL) were spotted directly onto an MTP AnchorChip target plate (Bruker Daltonics, Bremen, Germany), allowed to dry, and overlaid with 1 μL 0.5 $\mu g \cdot \mu L^{-1}$ α-cyano-4-hydroxycinnamic acid (CHCA) matrix in 90% acetonitrile, 0.1% TFA. The analyses were performed using an Ultraflex II MALDI-TOF/TOF MS (Bruker Daltonics, Germany). The

obtained mass spectra were processed with FlexAnalysis and BIOTOOLS software (Bruker Daltonics). The MS and MS/MS data were used as query to search NCBInr database using an in-house-licensed Mascot search engine (Matrix Science, London, UK) and the following parameters: allowed global modification, carbamidomethyl cysteine; variable modification, oxidation of methionine; missed cleavages, 1; peptide tolerance, 80 ppm; MS/MS tolerance, ± 0.5 Da.

Deglycosylation

N-glycosylation sites were predicted using the servers NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) and GlycoEP (http://www.imtech.res.in/raghava/glycoep/). O-glycosylation sites were predicted using the servers NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) and GlycoEP.

*At*LPMO13A (10 μg) was incubated with 2 μL Endo H (New England Biolabs, Ipswitch, MA, USA) in a 20-μL reaction volume at 37 °C for 4 h. *Mo*LPMO13A (1 μg) was incubated with 1 μL Jack bean α-mannosidase (Sigma-Aldrich, St. Louis, MO, USA) and/or 1 μL β-mannosidase (Megazyme, Bray, Ireland) in a 10 μL reaction volume for 2 h at 25 °C and for another 2 h at 37 °C. Mobility shifts were analyzed using SDS/PAGE.

Thermal stability of the recombinant LPMOs

Differential scanning calorimetry (DSC) was used to analyze the conformational stabilities of *Mo*LPMO13A and *At*LPMO13A using a Nano DSC instrument (TA Instruments, New Castle, DE, USA). Protein samples (10 μM) were dialyzed against 3 × 1000 volumes of 10 mM Na acetate buffer, pH 5.5 for 24 h, degassed and loaded into sample cells, and scanned (20–100 °C, 1 °C·min^{-1}) with the dialysis buffer in the reference cell. Baseline scans, collected with buffer in both reference and sample cells, were subtracted from sample scans, and the UNIVERSAL ANALYSIS software (TA Instruments) with a DSC add-on was used to model the reference cell and baseline-corrected thermograms using a two-state scaled model to determine T_m (unfolding temperature, defined as the temperature of maximum apparent heat capacity) and the calorimetric heat of unfolding ΔH_{cal}.

Amplex Red activity assay

An assay based on Amplex Red and horseradish peroxidase was used to measure the extent of H_2O_2 generation, which is a side reaction catalyzed by the reduced LPMO copper center [38]. The assay was performed in 100 mM sodium phosphate pH 6.5 by incubating 50 μM Amplex Red, 1 U·mL^{-1} horseradish peroxidase, 20 μM *At*LPMO13A, 0–200 μM ascorbate as reductant in 200 μL at 25 °C, and the absorbance at 571 nm (A_{571}) was measured in a microtiter plate reader. Removal of the catalytic copper ion was carried out by incubating 50 μM LPMO with 1 mM EDTA in 20 mM HEPES, pH 7.5 for 30 min at room temperature, followed by desalting using a Sephadex G-25 column (GE Healthcare) equilibrated and eluted with a 10 mM MES, pH 6.5. To investigate recovery of activity from copper-depleted enzyme, 100 μM $CuSO_4$ was added and excess copper was removed as described above. The inhibition of the enzyme by β-cylodextrin or amylose of degree of polymerization 17 was tested by performing the same assay in the presence of varying amounts of the carbohydrates.

Insoluble starch-binding assay

Binding of *Mo*LPMO13A to insoluble wheat starch (Sigma-Aldrich) was analyzed using wheat starch, which was washed three times with MilliQ water, followed by 10 mM Na acetate, pH 5.5. Enzyme aliquots (2.3 μM or 4.7 μM) were added to 0, 1, 5, 10, 20, 30, or 45 mg·mL^{-1} starch suspensions in the above buffer to a final volume of 300 μL. The suspensions were incubated at 4 °C for 1 h, and then centrifuged (4000 ***g***, 5 min, 4 °C) to pellet the starch. Free enzyme concentrations (A_{280}) in the supernatants from each suspension were measured and used to determine the fraction of bound protein. A one-site binding model was fit to the binding isotherms, where P is protein, S is starch, and B_{max} is the maximum binding capacity: $[P_{bound}] = B_{max}\cdot[S]/(K_d + [S])$, by nonlinear regression (Langmuir isotherm) using PRISM 6 software (GraphPad, La Jolla, CA, USA). Controls without added enzyme, and with enzyme and 2 mM β-CD were performed.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measurements were carried out using a MicroCal iTC200 calorimeter (MicroCal, Northampton, MA, USA). *At*LPMO13A (200 μM), dialyzed against 10 mM MES, pH 6.5 overnight, was titrated with 2 mM β-CD dissolved in the same buffer at 25 °C with an initial injection of 0.4 μL, followed by 21 injections of 1.8 μL. Baseline measurements were made using an identical injection regime in the absence of protein. A one-binding site model was fit to the integrated normalized data to determine the binding parameters using the MICROCAL ORIGIN software package (OriginLab, Northampton, MA, USA).

Bioinformatics analysis

The CAZy database was used to retrieve AA13 sequences. The characterized The *Aspergillus nidulans* modular *An*AA13 (UniProt: Q5B1W7), that contains a CBM20 was

used as a query in a BLAST search against the nonredundant protein database [39] to retrieve additional orthologs which were not assigned in the CAZy database. Sequences with alignment scores above 200 were retrieved. Conserved domain searches were performed using NCBI Conserved Domain Database [40], to omit sequences lacking the putative CBM20. Multiple sequence alignments of AA13 catalytic modules and CBM20s from all retrieved AA13 orthologs were performed using MUSCLE [41] and rendered using ESPRIPT [42]. Sequences of 24 CBM20s from fungal GH13 α-amylases and GH15 glucoamylases were used for comparison with 75 sequences of CBM20s from LPMOs. A phylogenetic tree was calculated from the alignment using the ClustalW2 phylogeny with default settings [41] and visualized using DENDROSCOPE 3.5 [43].

Results and Discussion

Heterologous expression of the recombinant LPMOs and enzyme stability

*Mo*LPMO13A and *At*LPMO13A were produced and purified at high yields (Fig. 1). Both enzymes migrated as smeary bands with higher apparent molecular masses than the theoretically those calculated from the sequences (42 734 Da and 37 220 Da for *At*LPMO13A and *Mo*LPMO13A, respectively). The larger observed size is likely a result of O- and/or N-glycosylations, which are predicted for both LPMOs. Particularly, O-glycosylation in the serine/threonine-rich linker, connecting the catalytic module to the C-terminal CBM20 (Fig. S1). Moreover, *At*LPMO13A has a predicted N-glycosylation site (N379) in the CBM20 (Fig. S1), at a position suggested to interact with the substrate at starch-binding site 1 [23]. Glycosylation of this site is likely to reduce affinity due to steric hindering. Treatment with EndoH, which cleaves N-glycans, however, did not result in a visible change in migration pattern on the gels (data not shown). Therefore, the larger apparent size of the enzymes was mainly attributed to O-glycosylation. A combination of α- and β-mannosidase treatment of the LPMOs resulted in an apparent decrease in size on SDS/PAGE gels (data not shown), which confirms O-mannosylation of the enzymes as observed for other LPMOs expressed in *P. pastoris* [38,44,45]. The yield of mannosidases treatment was too low to allow a preparative deglycosylation of the enzyme. Mass spectrometric analysis confirmed that the signal peptide of *Mo*LPMO13A was correctly processed to yield a native N terminus (data not shown), which is a prerequisite for LPMO activity. The stability of the LPMOs was investigated using DSC analysis, which showed that both enzymes were highly thermostable attesting their structural integrity. The denaturing temperatures (T_m) were 70.0 and 70.9 °C, for *At*LPMO13A and *Mo*LPMO13A, respectively (Fig. S2). Both thermograms featured a single peak suggesting that the unfolding processes of the catalytic and CBM20 modules were overlapping. These T_m temperatures are slightly higher but comparable to other reported T_m values for fungal LPMOs (63.0–68.9 °C) [38]. This is also in agreement with the large calorimetric enthalpy of approximately 480 kJ·mol^{-1} for both proteins.

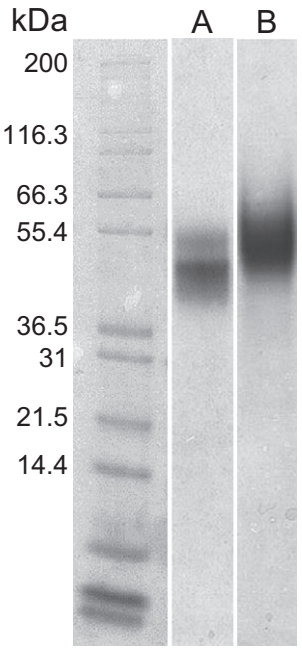

Fig. 1. SDS/PAGE showing Mark 12 protein standard (Invitrogen) in the left lane and recombinant purified *Mo*LPMO13A and *At*LPMO13A in lanes A and B, respectively.

Starch-specific LPMOs activity assays

The Amplex Red assay showed that *At*LPMO13A was active and that the activity correlated with ascorbate concentration and time (Fig. 2). The depletion of copper ion almost abolished activity, which was restored when the enzyme was reincubated with $CuSO_4$ (Fig. 2). These data confirmed that the Cu-active site of the enzyme was intact. We have also observed inhibition of activity in the presence of either 1 mM β-cyclodextrin or 0.5 mM amylose with an average degree of polymerization of 17, which suggests that both these ligands block the access to the active site. We have also observed the dependence of activity on enzyme concentration. Moreover, the generation of

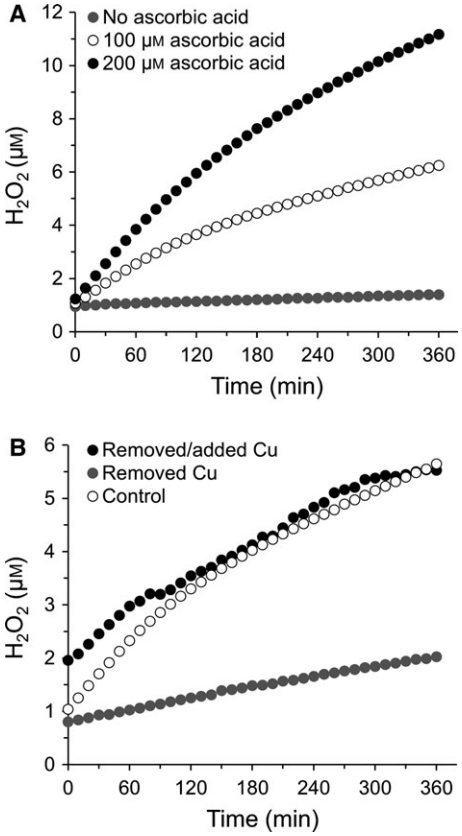

Fig. 2. Amplex Red activity assay performed on *At*LPMO13A. (A) The enzyme displays higher activity with increased ascorbate concentration; (B) The Cu ion-deprived enzyme shows markedly lower activity as compared to the control (no Cu ions are added or depleted from the enzyme). The activity of the Cu ion-deprived enzyme was restored by the addition of $CuSO_4$. These data demonstrate that the Cu-active center of the recombinant enzyme is intact.

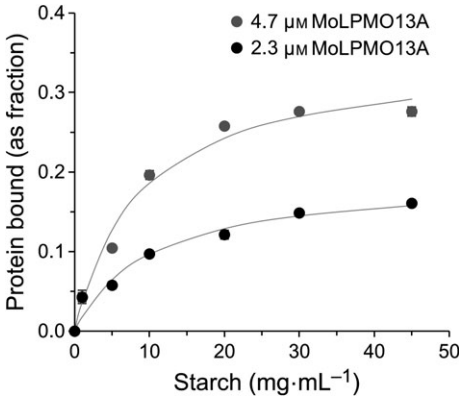

Fig. 3. *Mo*LPMO13A binding to insoluble wheat starch. The binding isotherms show the binding of 2.3 μM (black circles) and 4.7 μM (gray circles) enzyme to wheat starch (1–45 mg·mL^{-1}). The data are shown as means of duplicates ± standard deviations and the fits of a one-binding site model to the data are depicted as grey lines.

nonoxidized oligomeric products from both *At*LPMO13A and *Mo*LPMO13A was observed by thin layer chromatography and MALDI-TOF analysis, when the enzymes were incubated with amylose, but no oxidized products could be identified. No products were detected after incubation with β-cyclodextrin (data not shown).

Starch-specific LPMOs bind to starch granules and the starch mimic β-cyclodextrin

*Mo*LPMO13A and *At*LPMO13A were purified using β-cyclodextrin (β-CD) affinity chromatography suggesting that both enzymes possess affinity for starch. Indeed, the binding of *Mo*LPMO13A to wheat starch granules was demonstrated (Fig. 3). Control experiments performed in the presence of 2 mM β-CD abolished binding, which confirms the binding specificity and precludes aggregation artifacts (data not shown). An equilibrium dissociation constant $K_d = 8.70 \pm 1.90$ mg·mL^{-1} and a maximum binding capacity of 35% were determined from the binding isotherms. This affinity is roughly one order of magnitude higher than that reported for the cyclodextrin glycosyltransferase from *Bacillus circulans*, which harbors a CBM20, using potato starch as ligand ($K_d = 0.79$ mg·mL^{-1}) [46]. This enzyme, however, possesses significant affinity in its active site to β-cyclodextrin and helical structures in starch, which is likely to contribute to the higher affinity. By comparison, the affinity of the *Mo*LPMO13A to wheat starch was roughly one order of magnitude higher than the enzyme comprising the homologous CBM20 from the *Aspergillus niger* glucoamylase fused to a β-galactosidase using corn starch as ligand ($K_d = 55.6$ mg·mL^{-1}) [25]. *M. oryzae* is a rice pathogen [47,48], while *A. terreus*, for example, is frequently associated with crops, for example, wheat and potatoes [49,50], which may affect the starch type preference for enzymes from these organisms. Nonetheless, the binding of *Mo*LPMO13A to wheat starch clearly demonstrates comparable binding to amylolytic CBM20-containing hydrolases.

β-CD is widely accepted as a model probe for the binding of different starch-binding proteins to starch [51–53]. The binding of recombinant *At*LPMO13A to this model ligand was measured using ITC (Fig. 4). The data established moderate affinity binding to β-

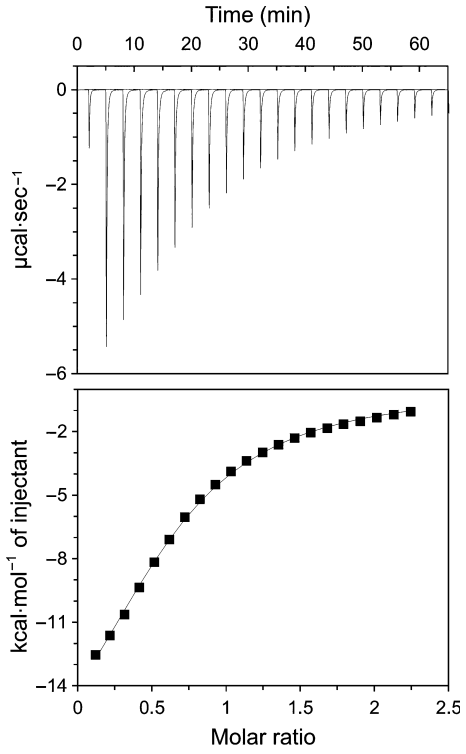

Fig. 4. Isothermal titration calorimetry (ITC) of the binding of *At*LPMO13A to β-cyclodextrin. Binding thermogram (top), and normalized integrated heat response (lower panel, black squares) are shown and the fit (black line) of a one-site binding model to the data. Titrations were performed in 10 mM MES, pH 6.5 at 25 °C.

CD with an association equilibrium constant, $K_a = (2.74 \pm 0.18) \times 10^4$ M^{-1} equivalent to a dissociation constant, $K_d = 37$ μM and a favorable free energy change, $\Delta G = -25.4$ kJ·mol^{-1}. The binding was driven by a favorable enthalpy ($\Delta H = -63.0 \pm 0.48$ kJ·mol^{-1}), which was off-set by an unfavorable entropy ($-T\Delta S = 37.6$ kJ·mol^{-1}). The binding affinity of *At*LPMO13A to β-CD is slightly lower, but in the same range as the K_d (~ 10 μM) obtained for the isolated CBM20 from the *A. niger* glucoamylase [54] and the K_d (~ 19 μM) for the full-length enzyme [55]. The thermodynamic signature of binding is also very similar, which suggests that the binding affinity of the CBM20 appended to the *A. terreus* LPMO is similar to counterparts that occur with starch-degrading hydrolases. Notably, the obtained binding stoichiometry (n) was 0.54 as compared to value of two obtained for typical CBM20 domains that possess two binding sites [55]. It cannot be excluded that the binding stoichiometry is reduced due to the presence of a fraction of the enzymes with impaired binding due to steric hindrance caused by glycosylation (Fig. S1), but the ITC affinity and thermodynamic parameters clearly demonstrate a typical binding pattern of modular GHs that target starch aided by a CBM20 module. It is not possible from these data to rule out the presence of additional binding sites on the catalytic module, which may be impaired due to glycosylation. Interestingly, activity was only possible to demonstrate on the full-length LPMO from *A. nidulans*, which suggests that the starch binding mediated by the CBM20 contributes importantly to the catalytic potency of starch-active LPMOs [32]. Solvent accessible aromatic residues are almost invariably present in the active sites of GHs, but are conspicuously lacking in LPMO counterparts, which is indicative of a more dynamic lower affinity substrate binding in the active sites of LPMOs as compared to GHs. Hence, the affinity and the specificity of the CBMs may be important for targeting of the LPMOs activity to specific sites at the surface of complex insoluble substrates. Removal of the cellulose binding CBM2 from LPMOs of AA10 enzymes (*Sc*LPMO10C and *Tf*LPMO9B) resulted in a twofold reduction in cellulolytic activity toward cellulosic substrates *in vitro* [56,57]. In contrast, the deletion of a CBM1 from an AA9 enzyme (*Nc*LPMO9C) did not affect the LPMO activity against cellulosic substrate but decreased it against xyloglucan [58]. Clearly, additional studies are needed to highlight the role and the contribution of CBMs to the mode of substrate binding and function of LPMOs.

Evolutionary conservation of CBM20 from starch-active LPMOs and GHs

Currently, the family AA13 contains only 14 LPMO sequences in the CAZy database (http://cazy.org/AA13.html), which only displays finished GenBank entries. A BLAST search using *An*AA13 as a query uncovered a total of 75 LPMO homologs possessing CBM20 modules. The *A. terreus* LPMO (*At*LPMO13A), which is currently not in the CAZy database, is a close homolog to the characterized *An*AA13, sharing 75% identity, and this study confirms the binding functionality of this starch-specific LPMO. An analysis of all AA13 LPMOs showed that the catalytic modules share 60–80% sequence identity including the conserved catalytic residues reported for other LPMOs (Fig. S3) [59]. Strikingly, approximately 60% of these sequences possess a C-terminal CBM20, frequently associated with amylolytic hydrolases [22], which corresponds to a twice as high relative occurrence of CBMs as compared to the average in LPMOs. This suggests an

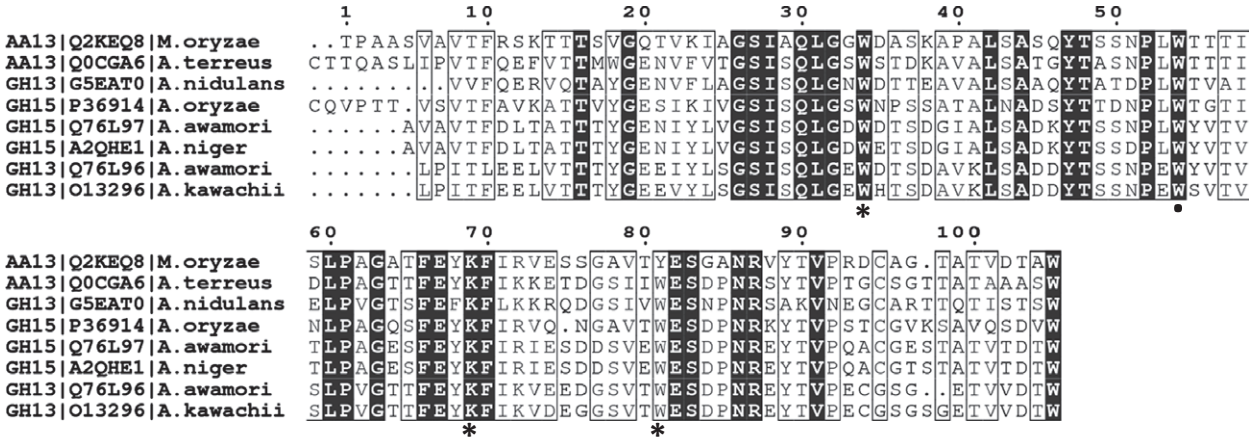

Fig. 5. Sequence alignment of CBM20 modules from *At*LPMO13A, *Mo*LPMO13A and fungal characterized GH13 α-amylases and GH15 glucoamylases. Fully conserved residues are in white letters and gray background, the SBS1 residues (two tryptophans/tyrosine and a lysine) are marked with stars, and SBS2 tryptophan is marked with black dot. The UniProt accessions and organisms are indicated.

evolutionary advantage conferred by the CBM20 in the breakdown of starch, which applies for both LPMOs and GHs.

The alignment of sequences showed that CBM20 appended to AA13 are highly similar to counterparts appended to GH13 and GH15 (Figs 5 and S4) [60]. Overall, the three out of four consensus starch-binding sites (SBS) 1 and 2 carbohydrate-binding residues (corresponding to W543, K578, W590, and W563 of *A. niger* glucoamylase [53]) are conserved in CBM20s from fungal AA13, GH13, and GH15. It has been shown that substitutions of these residues cause substantial affinity losses for starchy substrates [23,54]. At SBS1 one of tryptophans (W590 in *A. niger* glucoamylase) is substituted by tyrosine in several AA13 proteins, including *Mo*LPMO13A, and glycine in one AA13 protein. The subtle change from a tryptophan to a tyrosine is unlikely to cause major affinity changes. Furthermore, the two polar residues at SBS1 (corresponding to K578 and N595 of *A. niger* glucoamylase) that are known to form hydrogen bonds with starchy ligands, are conserved in all except one AA13 form *Verticillium dahliae*. The conservation of functional residues in LPMO-associated CBM20s is in agreement with the ITC affinity measurements that show comparable binding affinity and thermodynamics to fungal CBM20s. To investigate if CBM20 from LPMOs could be distinguished from counterparts in hydrolases, for example, α-amylases and glucoamylases, we performed a phylogenetic analysis of all CBM20 sequences from LPMOs and 24 amylolytic hydrolases (Fig. 6). Notably, the CBM20 sequences clustered largely based on taxonomy, with no apparent regard to the nature of the cognate catalytic module (i.e., LPMO or hydrolase). These findings suggest a strong pressure to retain the CBM20 as a common binding scaffold during the evolution of catalytically diverse oxidative LPMOs and hydrolases targeting starch. This is also consistent with the early observations of Janecek *et al.* [61] with regard to clustering of CBM20 in amylolytic hydrolases.

In conclusion, this study presents the first quantitative data on the binding properties of starch-active LPMOs possessing CBM20 modules. The measured binding affinities and thermodynamic fingerprint revealed a binding pattern typical of CBM20-containing amylolytic hydrolases. These results suggest that starch-binding of AA13 enzymes is mediated mainly by the CBM20 part, but the presence of lower affinity binding sites on catalytic modules cannot be excluded. Sequence analysis revealed the conservation of ligand-binding residues in CBM20s from fungal starch-specific LPMOs and hydrolases suggestive of a shared function of these starch-binding modules in both catalytic contexts. This is also supported by the phylogenetic clustering of CBM20 sequences according to taxonomy with no obvious correlation with the activity of appended catalytic modules. The conservation of the CBM20-binding scaffold in the evolution of different oxidative and hydrolytic starch-degrading enzymes asserts the pivotal role of these modules in targeting granular starch.

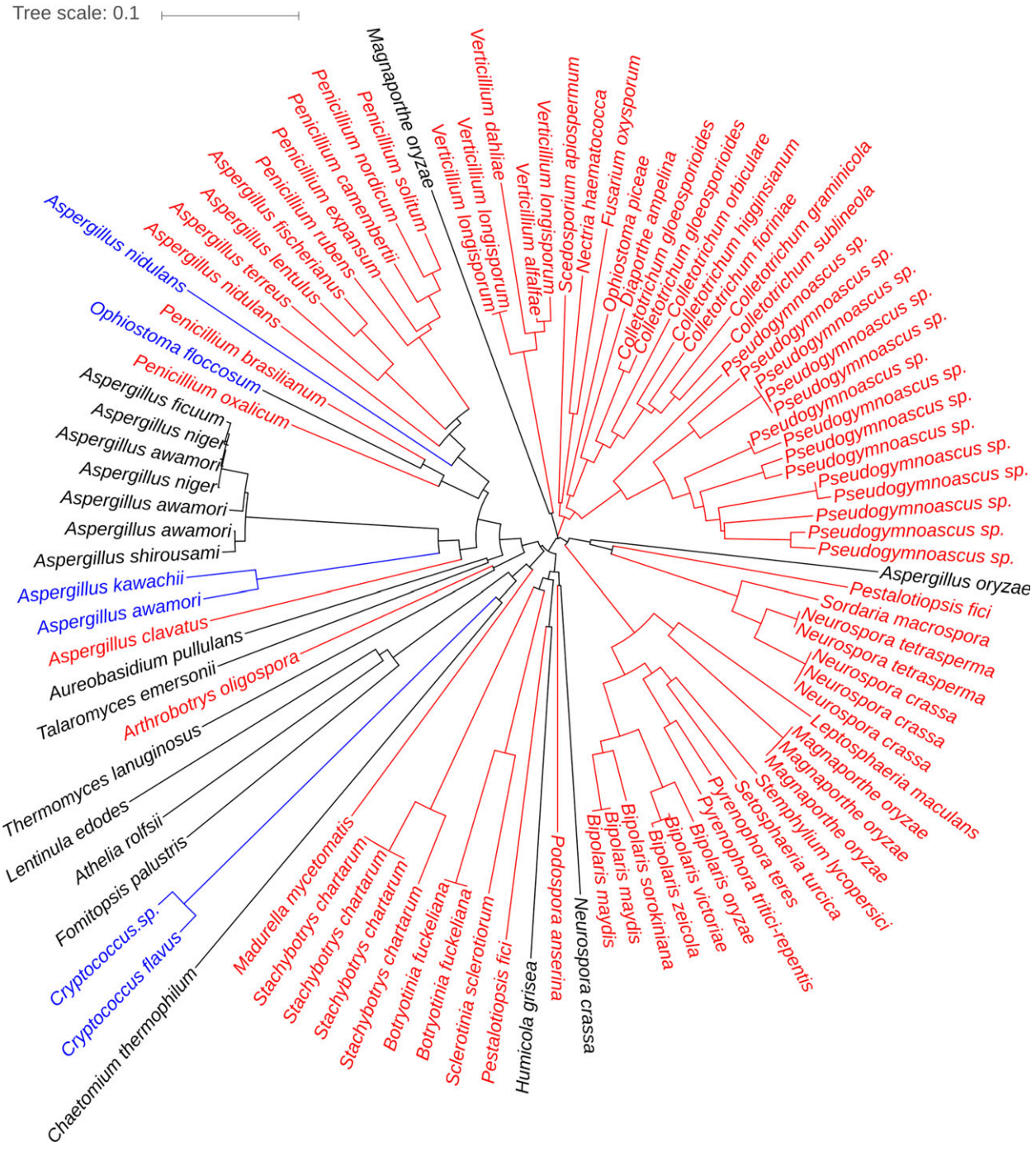

Fig. 6. Phylogram of CBM20 modules from AA13 LPMOs, fungal GH13 α-amylases, and GH15 glucoamylases. AA13 are in red, GH13 are in blue and GH15 in black. The enzyme sequences are labeled with organism names are to reveal taxonomic distribution. The tree shows that the CBM20 sequences mostly cluster based on taxonomic distribution with no apparent regard to the type of appended catalytic module.

Acknowledgements

This project was supported by a "Biotechnology-based Synthesis and Production" grant (NNF12OC0000769) from the Novo Nordisk Foundation. Carlsberg Foundation is acknowledged for instrument grants for purchase of ITC and DSC

microcalorimeters. A. H. Viborg is thanked for his CAZy research tool.

Author contributions

MAH and LN designed the experiments. LN performed all experimental work. TI performed the cloning of the AtAA13 enzyme. MAH and LN wrote the first manuscript and all authors participated in writing the manuscript and approved the final version.

References

1 Alvani K, Qi X and Tester RF (2012) Gelatinisation properties of native and annealed potato starches. *Starch/Stärke* **64**, 297–303.
2 Zeeman SC, Kossmann J and Smith AM (2010) Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol* **61**, 209–234.
3 Sajilata MG, Singhal RS and Kulkarni PR (2006) Resistant starch: a review. *Compr Rev Food Sci Food Saf* **5**, 1–17.
4 Buléon A and Colonna P (2007) Physicochemical behaviour of starch in food applications. In: *The Chemical Physics of Food* (Belton P, ed.), pp. 20–67. Blackwell Science Ltd, Oxford.
5 Singh N, Singh J, Kaur L, Sodhi NS and Gill BS (2003) Morphological, thermal and rheological properties of starches from different botanical sources. *Food Chem* **81**, 219–231.
6 Pérez S and Bertoft E (2010) The molecular structures of starch components and their contribution to the architecture of starch granules: a comprehensive review. *Starch/Stärke* **62**, 389–420.
7 Gallant DJ, Bouchet B and Baldwin PM (1997) Microscopy of starch: evidence of a new level of granule organization. *Carbohydr Polym* **32**, 177–191.
8 Buléon A, Colonna P, Planchot V and Ball S (1998) Starch granules: structure and biosynthesis. *Int J Biol Macromol* **23**, 85–112.
9 Tester RF, Karkalas J and Qi X (2004) Starch – composition, fine structure and architecture. *J Cereal Sci* **39**, 151–165.
10 Tester RF, Qi X and Karkalas J (2006) Hydrolysis of native starches with amylases. *Anim Feed Sci Technol* **130**, 39–54.
11 Jacobs H and Delcour JA (1998) Hydrothermal modifications of granular starch, with retention of the granular structure: a review. *J Agric Food Chem* **46**, 2895–2905.
12 Spier F, Zavareze EDR, Marques e Silva R, Elias MC and Dias ARG (2013) Effect of alkali and oxidative treatments on the physicochemical, pasting, thermal and morphological properties of corn starch. *J Sci Food Agric* **93**, 2331–2337.
13 van der Maarel MJEC, van der Veen B, Uitdehaag JCM, Leemhuis H and Dijkhuizen L (2002) Properties and applications of starch-converting enzymes of the α-amylase family. *J Biotechnol* **94**, 137–155.
14 Sun H, Zhao P, Ge X, Xia Y, Hao Z, Liu J and Peng M (2010) Recent advances in microbial raw starch degrading enzymes. *Appl Biochem Biotechnol* **160**, 988–1003.
15 Robertson GH, Wong DWS, Lee CC, Wagschal K, Smith MR and Orts WJ (2006) Native or raw starch digestion: a key step in energy efficient biorefining of grain. *J Agric Food Chem* **54**, 353–365.
16 Gilbert HJ, Knox JP and Boraston AB (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol* **23**, 669–677.
17 Guillén D, Sánchez S and Rodríguez-Sanoja R (2010) Carbohydrate-binding domains: multiplicity of biological roles. *Appl Microbiol Biotechnol* **85**, 1241–1249.
18 Hervé C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ and Knox JP (2010) Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci USA* **107**, 15293–15298.
19 Cuskin F, Flint JE, Gloster TM, Morland C, Baslé A, Henrissat B, Coutinho PM, Strazzulli A, Solovyova AS, Davies GJ *et al.* (2012) How nature can exploit nonspecific catalytic and carbohydrate binding modules to create enzymatic specificity. *Proc Natl Acad Sci USA* **109**, 20889–20894.
20 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM and Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, 490–495.
21 Boraston AB, Bolam DN, Gilbert HJ and Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769–781.
22 Machovic M and Janecek S (2006) Starch-binding domains in the post-genome era. *Cell Mol Life Sci* **63**, 2710–2724.
23 Janecek S, Svensson B and MacGregor EA (2011) Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb Technol* **49**, 429–440.
24 Juge N, Nøhr J, Le Gal-Coëffet MF, Kramhøft B, Furniss CSM, Planchot V, Archer DB, Williamson G and Svensson B (2006) The activity of barley α-amylase on starch granules is enhanced by fusion of a starch binding domain from *Aspergillus niger* glucoamylase. *Biochim Biophys Acta* **1764**, 275–284.
25 Chen LJ, Ford C, Kusnadi A and Nikolov ZL (1991) Improved adsorption to starch of a β-galactosidase

fusion protein containing the starch-binding domain from *Aspergillus* glucoamylase. *Biotechnol Prog* **7**, 225–229.
26 Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A *et al.* (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci USA* **108**, 15079–15084.
27 Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ and Eijsink VGH (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci* **20**, 479–1483.
28 Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH and Westereng B (2014) Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci USA* **111**, 6287–6292.
29 Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SWA, Koetsier MJ, van Berkel WJH, Gruppen H and Kabel MA (2015) Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels* **8**, 101.
30 Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M and Eijsink VGH (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219–222.
31 Crouch LI, Labourel A, Walton PH, Davies GJ and Gilbert HJ (2016) The contribution of non-catalytic carbohydrate binding modules to the activity lytic polysaccharide monooxygenases. *J Biol Chem* **291**, 7439–7449.
32 Lo Leggio L, Simmons TJ, Poulsen JCN, Frandsen KEH, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L *et al.* (2015) Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun* **6**, 5961.
33 Vu VV, Beeson WT, Span EA, Farquhar ER and Marletta MA (2014) A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci USA* **111**, 13822–13827.
34 Vester-Christensen MB, Abou Hachem M, Naested H and Svensson B (2010) Secretory expression of functional barley limit dextrinase by *Pichia pastoris* using high cell-density fermentation. *Protein Expr Purif* **69**, 112–119.
35 Silvanovich MP and Hill RD (1976) Affinity chromatography of cereal α-amylase. *Anal Biochem* **73**, 430–433.
36 Juge N, Andersen JS, Tull D, Roepstorff P and Svensson B (1996) Overexpression, purification, and characterization of recombinant barley α-amylases 1 and 2 secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Expr Purif* **8**, 204–214.
37 Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In *Proteomics Protocols Handbook* (Walker JM, ed.), pp. 571–607. Humana Press Inc., Totowa, NJ.
38 Kittl R, Kracher D, Burgstaller D, Haltrich D and Ludwig R (2012) Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels* **5**, 79.
39 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PS I-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
40 Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, 222–226.
41 Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N and Lopez R (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* **43**, 580–584.
42 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, 320–324.
43 Huson DH and Scornavacca C (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol* **61**, 1061–1067.
44 Dimarogona M, Topakas E, Olsson L and Christakopoulos P (2012) Lignin boosts the cellulase performance of a GH61 enzyme from *Sporotrichum thermophile*. *Bioresour Technol* **110**, 480–487.
45 Koseki T, Mese Y, Fushinobu S, Masaki K, Fujii T, Ito K, Shiono Y, Murayama T and Iefuji H (2008) Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl Microbiol Biotechnol* **77**, 1279–1285.
46 Penninga D, van der Veen BA, Knegtel RMA, van Hijum SAFT, Rozeboom HJ, Kalk KH, Dijkstra BW and Dijkhuizen L (1996) The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J Biol Chem* **271**, 32777–32784.
47 Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J *et al.* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* **13**, 414–430.
48 Wilson RA and Talbot NJ (2009) Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. *Nat Rev Microbiol* **7**, 185–195.

49 Louis B, Waikhom SD, Roy P, Bhardwaj PK, Singh MW, Chandradev SK and Talukdar NC (2014) Invasion of *Solanum tuberosum* L. by *Aspergillus terreus*: a microscopic and proteomics insight on pathogenicity. *BMC Res Notes* **7**, 350.
50 Kück U, Bloemendal S and Teichert I (2014) Putting fungi to work: harvesting a cornucopia of drugs, toxins, and antibiotics. *PLoS Pathog* **10**, e1003950.
51 Glaring MA, Baumann MJ, Abou Hachem M, Nakai H, Nakai N, Santelia D, Sigurskjold BW, Zeeman SC, Blennow A and Svensson B (2011) Starch-binding domains in the CBM45 family - low-affinity domains from glucan, water dikinase and α-amylase involved in plastidial starch metabolism. *FEBS J* **278**, 1175–1185.
52 Sigurskjold BW, Christensen T, Payre N, Cottaz S, Driguez H and Svensson B (1998) Thermodynamics of binding of heterobidentate ligands consisting of spacer-connected acarbose and β-cyclodextrin to the catalytic and starch-binding domains of glucoamylase from *Aspergillus niger* shows that the catalytic and starch-binding sites are in close proximity in space. *Biochemistry* **37**, 10446–10452.
53 Sorimachi K, Le Gal-Coëffet MF, Williamson G, Archer DB and Williamson MP (1997) Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to β-cyclodextrin. *Structure* **5**, 647–661.
54 Williamson MP, Le Gal-Coëffet MF, Sorimachi K, Furniss CSM, Archer DB and Williamson G (1997) Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. *Biochemistry* **36**, 7535–7539.
55 Sigurskjold BW, Svensson B, Williamson G and Driguez H (1994) Thermodynamics of ligand-binding to the starch-binding domain of glucoamylase from *Aspergillus niger*. *Eur J Biochem* **225**, 133–141.
56 Forsberg Z, Mackenzie AK, Sorlie M, Rohr ÅK, Helland R, Arvai AS, Vaaje-Kolstad G and 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.
57 Arfi Y, Shamshoum M, Rogachev I, Peleg Y and Bayer EA (2014) Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc Natl Acad Sci USA* **111**, 9109–9114.
58 Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Várnai A, Røhr ÅK, Payne CM, Sørlie M, Sandgren M *et al.* (2015) Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem* **290**, 22955–22969.
59 Hemsworth GR, Johnston EM, Davies GJ and Walton PH (2015) Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol* **33**, 747–761.
60 Christiansen C, Abou Hachem M, Janecek S, Viksø-Nielsen A, Blennow A and Svensson B (2009) The carbohydrate-binding module family 20 - diversity, structure, and function. *FEBS J* **276**, 5006–5029.
61 Janecek S and Sevcík J (1999) The evolution of starch-binding domain. *FEBS Lett* **456**, 119–125.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:
Fig. S1. Predicted N- and O-glycosylation sites in *At*LPMO13A and *Mo*LPMO13A (see materials and methods).
Fig. S2. Differential scanning calorimetry (DSC) Analysis representing a two-state protein unfolding curves for *At*LPMO13A and *Mo*LPMO13A.
Fig. S3. Sequence alignment of the catalytic modules of LPMOs assigned into AA13 in the CAZy database.
Fig. S4. Alignment of CBM20 modules from AA13 LPMOs and fungal GH13 α-amylases and GH15 glucoamylases.