JBC RESEARCH ARTICLE

Structural and functional variation of chitin-binding domains of a lytic polysaccharide monooxygenase from *Cellvibrio japonicus*

Received for publication, April 27, 2021, and in revised form, August 11, 2021 Published, Papers in Press, August 17, 2021,
https://doi.org/10.1016/j.jbc.2021.101084

Eva Madland[1,‡], Zarah Forsberg[2,‡], Yong Wang[3], Kresten Lindorff-Larsen[3], Axel Niebisch[4], Jan Modregger[4], Vincent G. H. Eijsink[2], Finn L. Aachmann[1], and Gaston Courtade[1,*]

From the [1]Norwegian Biopolymer Laboratory (NOBIPOL), Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Trondheim, Norway; [2]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), Ås, Norway; [3]Structural Biology and NMR Laboratory, Department of Biology, Linderstrøm-Lang Centre for Protein Science, University of Copenhagen, Copenhagen, Denmark; [4]Eucodis Bioscience GmbH, Wien, Austria

Edited by Gerald Hart

Among the extensive repertoire of carbohydrate-active enzymes, lytic polysaccharide monooxygenases (LPMOs) have a key role in recalcitrant biomass degradation. LPMOs are copper-dependent enzymes that catalyze oxidative cleavage of glycosidic bonds in polysaccharides such as cellulose and chitin. Several LPMOs contain carbohydrate-binding modules (CBMs) that are known to promote LPMO efficiency. However, structural and functional properties of some CBMs remain unknown, and it is not clear why some LPMOs, like *Cj*LPMO10A from the soil bacterium *Cellvibrio japonicus*, have multiple CBMs (*Cj*CBM5 and *Cj*CBM73). Here, we studied substrate binding by these two CBMs to shine light on their functional variation and determined the solution structures of both by NMR, which constitutes the first structure of a member of the CBM73 family. Chitin-binding experiments and molecular dynamics simulations showed that, while both CBMs bind crystalline chitin with K_d values in the micromolar range, *Cj*CBM73 has higher affinity for chitin than *Cj*CBM5. Furthermore, NMR titration experiments showed that *Cj*CBM5 binds soluble chitohexaose, whereas no binding of *Cj*CBM73 to this chitooligosaccharide was detected. These functional differences correlate with distinctly different arrangements of three conserved aromatic amino acids involved in substrate binding. In *Cj*CBM5, these residues show a linear arrangement that seems compatible with the experimentally observed affinity for single chitin chains. On the other hand, the arrangement of these residues in *Cj*CBM73 suggests a wider binding surface that may interact with several chitin chains. Taken together, these results provide insight into natural variation among related chitin-binding CBMs and the possible functional implications of such variation.

Chitin is a linear and water insoluble polysaccharide composed of β-1,4-linked GlcNAc units found in the cell wall matrix of fungi and the exoskeletons of arthropods. Despite being the second most abundant polymer in nature, after cellulose, chitin does not accumulate in most ecosystems and tends to be absent in fossils (1). This is testimony to the capacity of nature to depolymerize and recycle chitin.

Chitinases (Enzyme Commission no.: 3.2.1.14) catalyze the hydrolytic degradation of chitin and belong to the glycoside hydrolase class of carbohydrate-active enzymes. Even though glycoside hydrolases efficiently degrade amorphous regions of chitin (2–4), they are inefficient at degrading crystalline chitin (5). The discovery of lytic polysaccharide monooxygenases (LPMOs) (6, 7) has given new insights into the degradation of chitin and other structural polysaccharides. LPMOs are copper-dependent enzymes that catalyze oxidative cleavage of glycosidic bonds in crystalline polysaccharides (6, 8). Aside from chitin, LPMOs have been reported to act on polysaccharides, such as cellulose (8–11), various hemicelluloses (12), starch (13), and significant activity on soluble substrates has been detected for some LPMOs (14, 15). In the degradation of chitin, LPMOs act in synergy with chitinases (4, 7). It is thought that LPMOs oxidize crystalline surfaces, causing "nicks" that lead to reduced crystallinity and introduction of new access points for chitinases (6, 10, 16).

Carbohydrate Active enZymes (CAZymes), such as chitinases and LPMOs, may just be composed of a single catalytic domain (CD) or may contain one or more non-CDs such as carbohydrate-binding modules (CBMs). Currently (as of September 2021), the CAZy database (17) contains 88 families of CBMs with a wide variety of binding specificities, including crystalline polysaccharides and short, soluble oligosaccharides (18, 19). The major role of CBMs is to keep an enzyme in close proximity of a substrate, thereby enhancing the effective concentration of the enzyme and overall reaction efficiency (18). In the context of LPMOs, CBMs may have a particularly important role because proximity to the substrate not only contributes to enzyme efficiency but also protects the enzyme from autocatalytic inactivation (20–22). Several studies have shown that removal of CBMs has a negative effect on LPMO performance (21–25). There are multiple families of chitin-binding and cellulose-binding CBMs, which may have different substrate specificities (*e.g.*, (23, 26)). For example,

‡ These authors contributed equally to this work.
* For correspondence: Gaston Courtade, gaston.courtade@ntnu.no.

Lehtiö *et al.* (26) showed that cellulose-binding modules belonging to two different CBM families bind to different parts of cellulose. Therefore, it is not trivial to predict or determine the role of CBMs, and a better understanding of the ways in which they bind their substrates is needed.

To address functional variation among chitin-binding CBMs, we have used chitin-active *Cj*LPMO10A from *Cellvibro japonicus* as a model system. The CD of this LPMO, which belongs to the auxiliary activity family 10 (AA10) in CAZy, is appended to two type A (18, 19) chitin-binding CBMs: an internal family 5 CBM (*Cj*CBM5) and C-terminal family 73 CBM (*Cj*CBM73) (Fig. 1). The three domains of *Cj*LPMO10A are connected by linkers that are rich in serine residues and are both approximately 30 amino acids long (Fig. 1). A previous study has shown that both CBMs bind to α- and β-chitin, thus enhancing substrate binding by the LPMO, and that the full-length (FL) protein is more efficient in comparison to the CD alone (25). In the present study, we have compared multiple truncated variants of *Cj*LPMO10A (Fig. 1*A*) to understand the roles of the appended CBMs in LPMO functionality. Furthermore, we have used NMR spectroscopy to elucidate the solution structures of the two CBMs: *Cj*CBM5 and *Cj*CBM73, where the latter is the first structure to be determined for a member of the CBM73 family. We also used NMR titration experiments to investigate binding of the CBMs to chitohexaose. These results were complemented with molecular dynamics simulations to gain more insights into CBM binding to α-chitin. Overall, the results show that while *Cj*CBM5 and *Cj*CBM73 are similar in overall structure and both bind to crystalline chitin, they differ in apparent *T*m, binding site architecture, and the ability to bind individual chitin chains.

Results

The effect of CBMs on chitin oxidation is substrate concentration dependent

To better understand the functional roles of *Cj*CBM5 and *Cj*CBM73 in relation to full-length *Cj*LPMO10A, we started by testing the performance of the three catalytically active

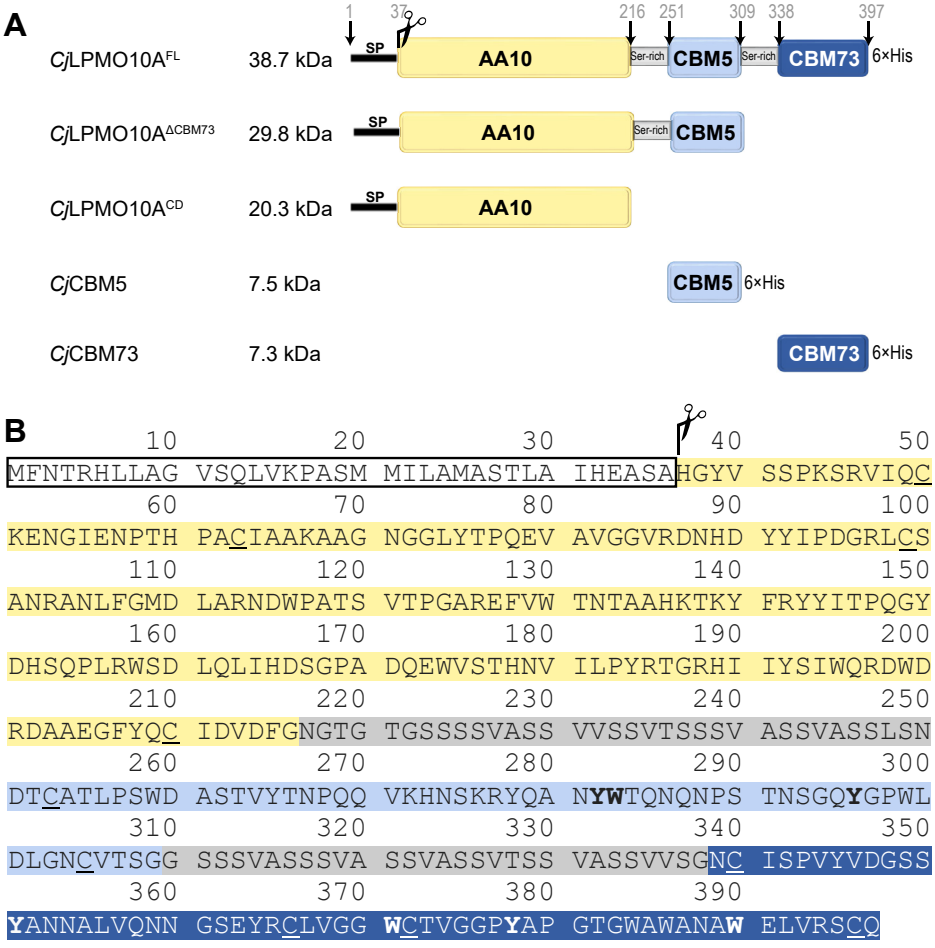

Figure 1. Domain architecture and primary structure of *Cj*LPMO10A. *A*, domain architecture and molecular weight of *Cj*LPMO10A and the truncated variants used in this study. The numbers above the full-length enzyme show the transitions between the domains and the linkers. The signal peptide (residues 1–37) is cleaved off during secretion. The indicated molecular weights are based on the mature protein, that is, enzymes without signal peptides. *B*, primary structure of *Cj*LPMO10A^{FL} with color coding according to panel *A*. Aromatic residues located on the binding surfaces of the two CBMs, as determined in this study, are printed in *bold face*; cysteine residues involved in disulfide bonds are *underlined*. CBM, carbohydrate-binding module; CD, catalytic domain, Ser-rich linker; FL, full-length; His ×6, polyhistidine tag; SP, signal peptide.

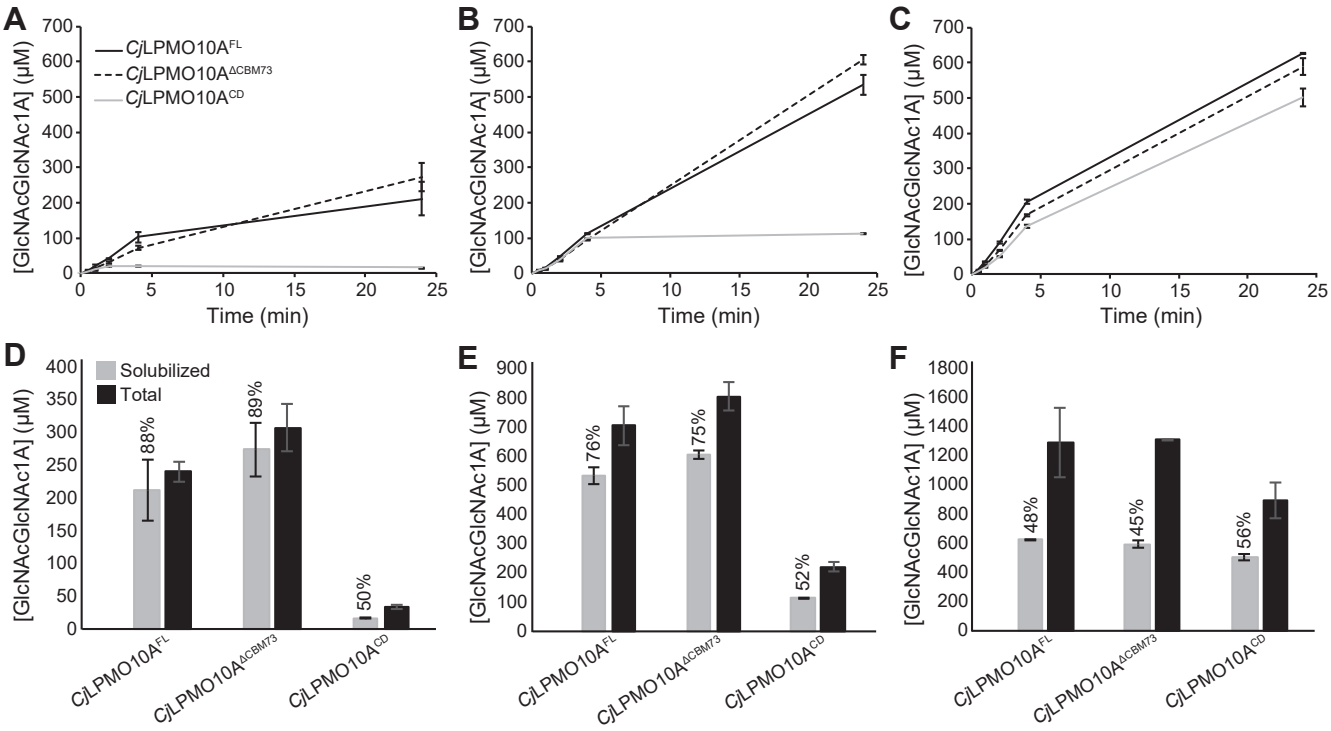

Figure 2. Chitin degradation by *Cj*LPMO10A variants. Panels *A–C* show progress curves for the formation of soluble oxidized products by *Cj*LPMO10A^{FL} (*solid black line*), *Cj*LPMO10A$^{\Delta CBM73}$ (*dashed black line*), and *Cj*LPMO10A^{CD} (*solid gray line*) at substrate concentrations of 2 g/l (*A*), 10 g/l (*B*), and 50 g/l (*C*) α-chitin. Panels *D–F* show quantification of solubilized (*gray bars*) and total oxidized sites (*black bars*) after 24 h of LPMO incubation at the various substrate concentrations, that is, 2 g/l (*D*), 10 g/l (*E*), and 50 g/l (*F*). The fraction of soluble oxidized products is given as a percentage of the total for each reaction. All reactions were carried out with 0.5 μM LPMO and 1 mM ascorbic acid in 50 mM sodium phosphate at pH 7.0 in a thermomixer set to 37 °C and 800 rpm. For quantification of soluble products, the solubilized fraction was further degraded by 0.5 μM *Sm*CHB prior to HPLC quantification. For quantification of total products (*i.e.*, soluble and insoluble fraction), samples were heat inactivated after which all α-chitin (diluted to 2 g/l) was degraded with a combination of 2.0 μM *Sm*ChiA and 0.5 μM *Sm*CHB. Note that the LPMO variants were used directly after purification and that their copper saturation levels may have varied; thus, the progress curves in panels *A–C* cannot be used for direct comparison of catalytic initial rates. The error bars show ±SD (n = 3). LPMO, lytic polysaccharide monooxygenase.

versions of *Cj*LPMO10A, namely *Cj*LPMO10A^{FL} (FL for full-length), *Cj*LPMO10A$^{\Delta CBM73}$ (for truncation of the CBM73 domain; see Fig. 1*A*) and fully truncated *Cj*LPMO10A^{CD} at different concentrations of α-chitin (2, 10, or 50 g/l; Fig. 2). At all substrate concentrations, the FL enzyme and the enzyme lacking only one CBM, *Cj*LPMO10A$^{\Delta CBM73}$, stayed active for the full duration of the experiment. At the two lowest substrate concentrations, product formation by *Cj*LPMO10A^{CD} ceased rapidly, and faster at the lowest substrate concentration, indicative of enzyme inactivation. However, at a substrate concentration of 50 g/l, all three variants showed similar progress curves and final product levels.

At the two lowest substrate concentrations, the amount of soluble oxidized products (relative to the total amount) was higher for the CBM-containing variants of *Cj*LPMO10A (>85%) compared with *Cj*LPMO10A^{CD} (about 50%) (Fig. 2, *D* and *E*). This indicates that, at these lower substrate concentrations, the presence of at least one CBM leads to more localized oxidation, generating a higher fraction of short soluble products, as discussed in Ref. (22) and later. At the highest substrate concentration (Fig. 2*F*), however, the fraction of soluble oxidized products was close to 50% for all three enzyme versions. All in all, the experiments depicted in Figure 2 did not show significant differences between the catalytic behavior of the two CBM-containing variants, but deletion of both CBMs had a major effect.

Thermal stability and oxidative performance

To assess possible functional differences between the FL enzyme and the variant lacking only the CBM73, we analyzed the effect of temperature on the oxidative performance of these variants (Fig. 3). It is believed that CAZymes with multiple CBMs have an advantage at elevated temperatures as the CBM(s) can counteract the loss of binding because of increased temperature (27–29). Interestingly, at the highest tested temperature (70 °C), *Cj*LPMO10A^{FL} showed significantly higher activity than *Cj*LPMO10A$^{\Delta CBM73}$. Thus, the presence of the CBM73 indeed has a beneficial effect on LPMO performance at higher temperatures. Determination of melting curves showed that the deletion of the *Cj*CBM73 had some effect on the shape of the curve but not on the apparent *T*m of approximately 70 °C (Fig. S1). The apparent *T*ms of the isolated CBMs were 57.2 °C for *Cj*CBM5 and 75.4 °C for *Cj*CBM73, whereas the apparent *T*m of the *Cj*LPMO10A^{CD} was 70.2 °C. In accordance with previous studies on the effect of copper binding on the stability of AA10 (30, 31) and AA9 (32) LPMOs, the apo variant of *Cj*LPMO10A^{CD} showed reduced stability ($T_{m,app}$ = 56.6 °C).

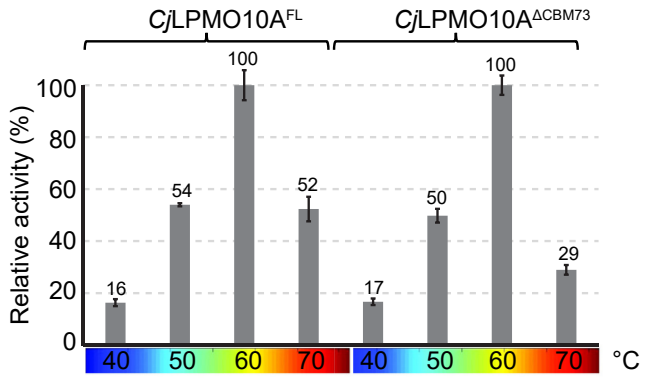

Figure 3. Catalytic performance of *Cj*LPMO10A^{FL} and *Cj*LPMO10A$^{\Delta CBM73}$ at varying temperatures. The relative activity was determined from linear progress curves for a 30-min reaction. The 100% value corresponds to 61 and 47 μM oxidized dimer (GlcNAcGlcNAc1A) for *Cj*LPMO10A^{FL} and *Cj*LPMO10A$^{\Delta CBM73}$, respectively. All reactions were carried out with 0.5 μM LPMO, 10 g/l α-chitin, and 1 mM ascorbic acid in 50 mM sodium phosphate at pH 7.0 in a thermomixer set to the indicated temperature and 800 rpm. Prior to product quantification, the solubilized fraction was further degraded with 0.5 μM *Sm*CHB. Each point represents the average of values obtained in three independent experiments. LPMO, lytic polysaccharide monooxygenase.

Solution structures of CjCBM5 and CjCBM73

The solution structures of *Cj*CBM5 (Protein Data Bank [PDB] ID: 6Z40) and *Cj*CBM73 (PDB ID: 6Z41) were determined by NMR spectroscopy (Fig. 4 and Table S1). The chemical shift assignment completion for the backbone (N, H^N, C^α, H^α, and C′) and side chains (H and C) of *Cj*CBM5 (BioMagnetic Resonance Databank [BMRB] ID: 34519) was >88% and >65%, respectively, whereas these values were >87% and >59% for *Cj*CBM73 (BMRB ID: 34520). Because of the cloning procedure, both proteins contained a Met at the N terminus and an Ala followed by a 6×His tag at the C terminus. For *Cj*CBM5, no resonances from these additional amino acids were assigned, whereas for *Cj*CBM73, the backbone resonances of the additional Ala and the first His in the 6×His tag were assigned.

The structures of *Cj*CBM5 and *Cj*CBM73 are similar (C^α RMSD = 5.6 Å) and share the same overall fold (Fig. 4). This fold has previously been described (33) as an "L" shape or "ski boot" fold because of the loop region attached perpendicularly to an antiparallel β-sheet. The structures of both CBMs are stabilized by a disulfide bridge connecting the N- and C-terminal ends of the domain. The structure of *Cj*CBM73 has a short 3_{10} helix (residues 371–374) that is linked to the central β-strand by an additional disulfide bridge. These features are unique for the CBM73 family (Fig. S2) and lack in *Cj*CBM5 and other structurally characterized members of the CBM5 family.

Most CBMs rely on exposed aromatic residues that bind carbohydrates through CH–π interactions (18, 34). Based on structural information alone, Figure 4 shows that Y282, W283, and Y296 in *Cj*CBM5, and W371, Y378, and W386 in *Cj*CBM73, could be involved in substrate binding. As shown in

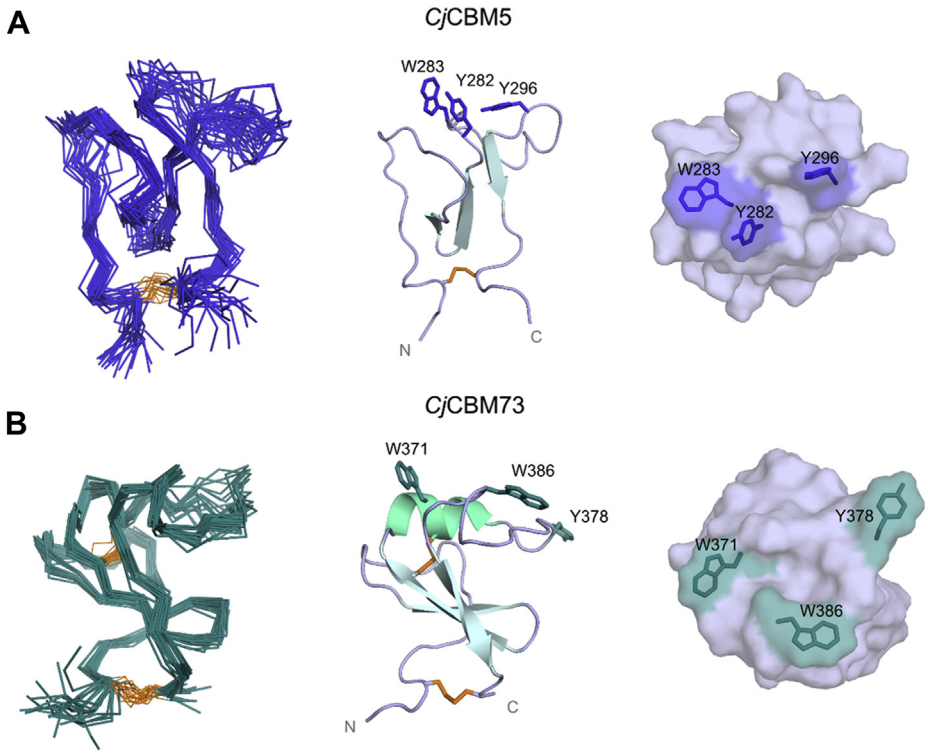

Figure 4. NMR solution structures of (*A*) *Cj*CBM5 (Protein Data Bank ID: 6Z40) and (*B*) *Cj*CBM73 (Protein Data Bank ID: 6Z41). The figures show backbone representations of the 20 conformers with the lowest CYANA target function (*left*), a *cartoon* representation of the structure with the lowest target function (*center*), and a view of the binding surfaces (*right*). The *cartoon* representations also display the secondary structure elements as well as aromatic residues of the putative binding surface. Disulfide bridges (residues 253–306 in *Cj*CBM5 and 340–396 and 366–372 for *Cj*CBM73) are highlighted in *orange*. His tags added for purification purposes (see Experimental procedures section) are not shown.

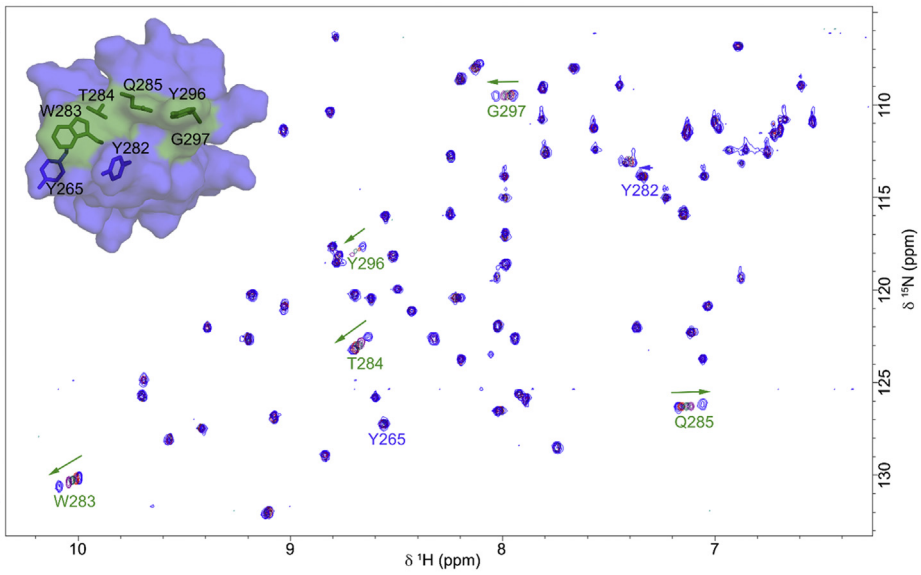

Figure 5. ^{15}N-HSQC of *Cj*CBM5 interacting with $(GlcNAc)_6$. The figure shows an overlay of ^{15}N-HSQC spectra for *Cj*CBM5 in the presence of $(GlcNAc)_6$ at various concentrations (0.2, 1.0, 2.5, and 10 mM). The *arrows* indicate the direction of change in chemical shifts as a result of the titration of *Cj*CBM5 with $(GlcNAc)_6$. Affected residues (W283, T284, Q285, Y296, and G297) are highlighted in *green* on the surface model of *Cj*CBM5. Other surface-exposed aromatic residues for which no significant chemical shift perturbation was detected (Y265 and Y282) are shown in *blue* for illustration purposes. HSQC, heteronuclear single quantum coherence.

Fig. S2, the aromatic pair Y282–W283 is almost fully conserved within the CBM5 family, whereas Y296 is less conserved. In the context of the CBM73 family (Fig. S2), W371, Y378, and W386 appear to be highly conserved. To test interactions between chitin and these aromatic patches and their neighboring polar residues, we performed NMR titrations with a soluble chitin substrate, chitohexaose $(GlcNAc)_6$.

Probing interactions between soluble chitin and CBMs by NMR

For *Cj*CBM5, titration with $(GlcNAc)_6$ led to significant $^{15}N–^{1}H^{N}$ chemical shift perturbation for W283 and Y296 as well as for residues in the neighboring loop region (T284, Q285, and G297) that are part of the putative binding surface (Fig. 5). The chemical shift perturbations were used to calculate a K_d = 2 ± 1 mM. Of note, this K_d value is some three orders of magnitude higher than the value obtained with solid α-chitin (see later).

In contrast to the experiment with *Cj*CBM5, titration of *Cj*CBM73 with up to 6.5 mM $(GlcNAc)_6$ did not result in any significant chemical shift perturbations, indicating that this CBM does not bind this soluble substrate.

*Binding of Cj**CBM5** and Cj**CBM73** to oxidized and nonoxidized α-chitin*

Previous binding studies have shown that both CBMs bind with micromolar affinity to both α- and β-chitin (25). These previous studies indicated similar K_d values (for α-chitin) for *Cj*CBM5 and *Cj*CBM73. Here, we tested binding using a similar setup, using both the same batch of α-chitin and a batch of α-chitin that had been preoxidized with *Cj*LPMO10A^{CD} as described later and in the Experimental procedures section.

Oxidized chitin was prepared to assess whether surface oxidation would affect CBM binding, one idea being that gradual oxidation of the substrate surface could facilitate release of otherwise strongly bound CBMs. The material was prepared by treating chitin with the CD of *Cj*LPMO10A^{CD}, followed by washing to remove solubilized oxidized chitooligosaccharides and residual LPMO (see Experimental procedures section for further details). The degree of oxidation of the solid fraction was determined upon complete enzymatic hydrolysis of the fraction, which entails that all oxidized sites end up as chitobionic acid. Data from six independent reactions, containing 20 mg/ml chitin, which correspond to approximately 45 mM of oxidized dimer in a theoretical 100% conversion reaction, indicated a degree of oxidation of about 0.3% (number obtained by dividing the chitobionic acid recovered from the solid fraction by the amount of chitobionic acid that would be obtained in a 100% conversion reaction). In an alternative approach, we divided the amount of chitobionic acid recovered from the solid fraction by the total amount of sugars (GlcNAc and chitobionic acid) recovered from this fraction, which indicated approximately 1% oxidation. Hence, the degree of oxidation of the insoluble fraction was estimated to be between 0.3% and 1%, and we assume that oxidation essentially happened on the substrate surface.

Figure 6 shows binding curves for the two CBMs with "nonoxidized" (panel *A* and *C*) or "preoxidized" (panel *B* and *D*) α-chitin. The data show that *Cj*CBM73 (K_d = 2.9 μM) binds with slightly higher affinity than *Cj*CBM5 (K_d = 8.5 μM). The binding studies with partly oxidized chitin showed similar results. The data showed a ∼20% increase in the K_d for *Cj*CBM5, indicating that binding by this CBM may be negatively affected by surface oxidation, but the difference was not statistically significant.

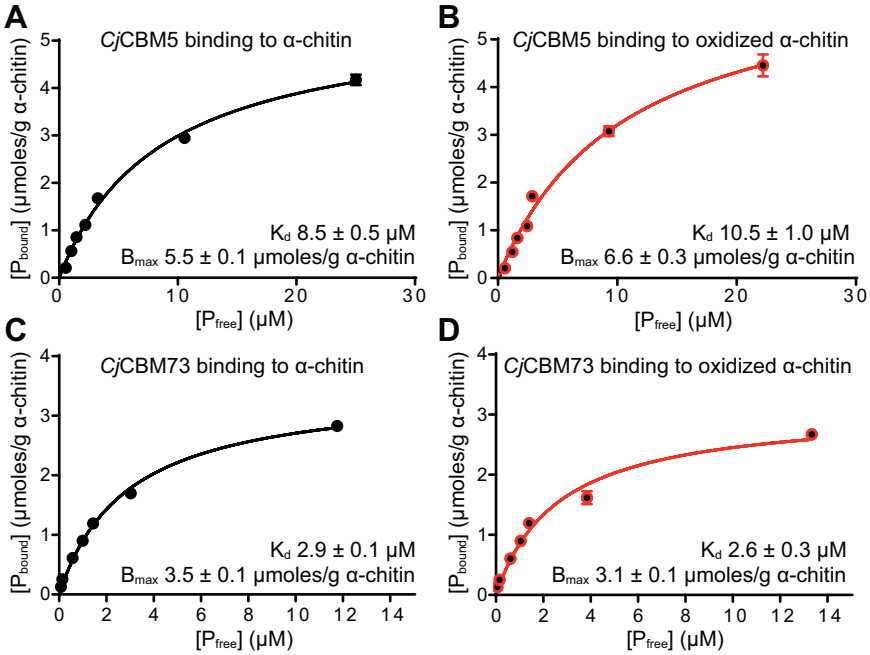

Figure 6. Binding of the CBMs of *Cj*LPMO10A to α-chitin. The plots show binding data for *Cj*CBM5 (*A* and *B*) and *Cj*CBM73 (*C* and *D*) incubated with α-chitin for 60 min. The experiments were carried out at 22 °C using 10 g/l α-chitin in 50 mM sodium phosphate buffer at pH 7.0 and show binding of *Cj*CBM5 and *Cj*CBM73 to nonoxidized (*A* and *C*) and oxidized (*B* and *D*) substrate. P_{bound} corresponds to bound protein (μmoles/g substrate), and P_{free} corresponds to nonbound protein (micromolar). The error bars show ±SD (n = 3). CBM, carbohydrate-binding module.

Simulations provide insight into binding of CBMs to α-chitin

Coarse-grained (CG) simulations were performed to further investigate interactions between CBMs and a model of the surface of α-chitin. CG models based on the Martini force field represent 3 to 4 atoms by a single "bead," thereby reducing the number of particles that are simulated (35). This allows simulations to be run longer and to sample longer timescales, compared with atomistic simulations. We combined CG models of chitin and the CBMs with well-tempered metadynamics (WT-MetaD) simulations to further enhance sampling of CBM–chitin binding/unbinding events, which occur on long timescales. In the WT-MetaD approach, protein conformations along a set of collective variables are biased by a history-dependent potential. The total bias (*i.e.*, sum of the Gaussians in the potential) forces the system to escape from local free-energy minima and explore different regions of the collective variable space. For the CBM–chitin model, we used two collective variables as proxies for binding: (i) the Euclidean distance (r_{chitin}) and (ii) number of contacts, $\langle c_w \rangle$, between aromatic residues in the putative substrate-binding surfaces (*Cj*CBM5: Y282, W283, and Y296; *Cj*CBM73: W371, Y378, and W386) and the chitin surface. Details on the calculation of these collective variables are provided in the Experimental procedures section.

To promote binding to chitin by the CBMs, it was necessary to rescale the interaction strengths between chitin beads and protein beads in the Martini model (see Experimental procedures section for details). The effect of rescaling these interactions by 0% (unchanged) or by an up to 15% increase in the strength of the chitin–protein interaction was evaluated by running umbrella-sampling simulations (Fig. S3) on the rescaled models and by comparing dissociation constants calculated from these simulations with experimentally determined values (Fig. 6). The results (Table 1) show that the best agreement with experiments was attained with a 10% increase in the chitin–protein interaction strength. The free-energy surfaces of *Cj*CBM5 and *Cj*CBM73 have similar appearances, but *Cj*CBM73 has a deeper well than *Cj*CBM5, which correlates with its experimentally observed stronger affinity for chitin (Fig. S3).

The number of contacts between all amino acids in each CBM and the α-chitin surface was calculated for every frame (n = 15,000) in the WT-MetaD simulation and reweighted using the bias from the simulation (see Experimental procedures section for details). The results (Fig. 7) show which residues have the most contacts, that is, $\langle c_w \rangle > 0.5$, with the substrate over time. For *Cj*CBM5 (Fig. 7, *A* and *C*), regions with most contacts include, and are to a large extent limited to, the three aromatic residues of the putative binding surface (Y282, W283, and Y296). In addition, the region around Y265 seems to be somewhat involved in substrate binding albeit

Table 1
Dissociation constants (K_d) for binding of the *Cj*CBMs to α-chitin, determined by experiments (Fig. 6) and simulations (Fig. 7 and Fig. S3)

Protein	Simulations (μM)				Experiments (μM)
	0%	5%	10%	15%	
*Cj*CBM5	1800	480	36	0.65	8.9 ± 0.5
*Cj*CBM73	290	76	15	0.13	2.9 ± 0.1

The modeled values were calculated from umbrella-sampling simulations in which the interaction strength between chitin beads and protein beads remained unchanged (0%) or was increased by 5, 10, and 15%.

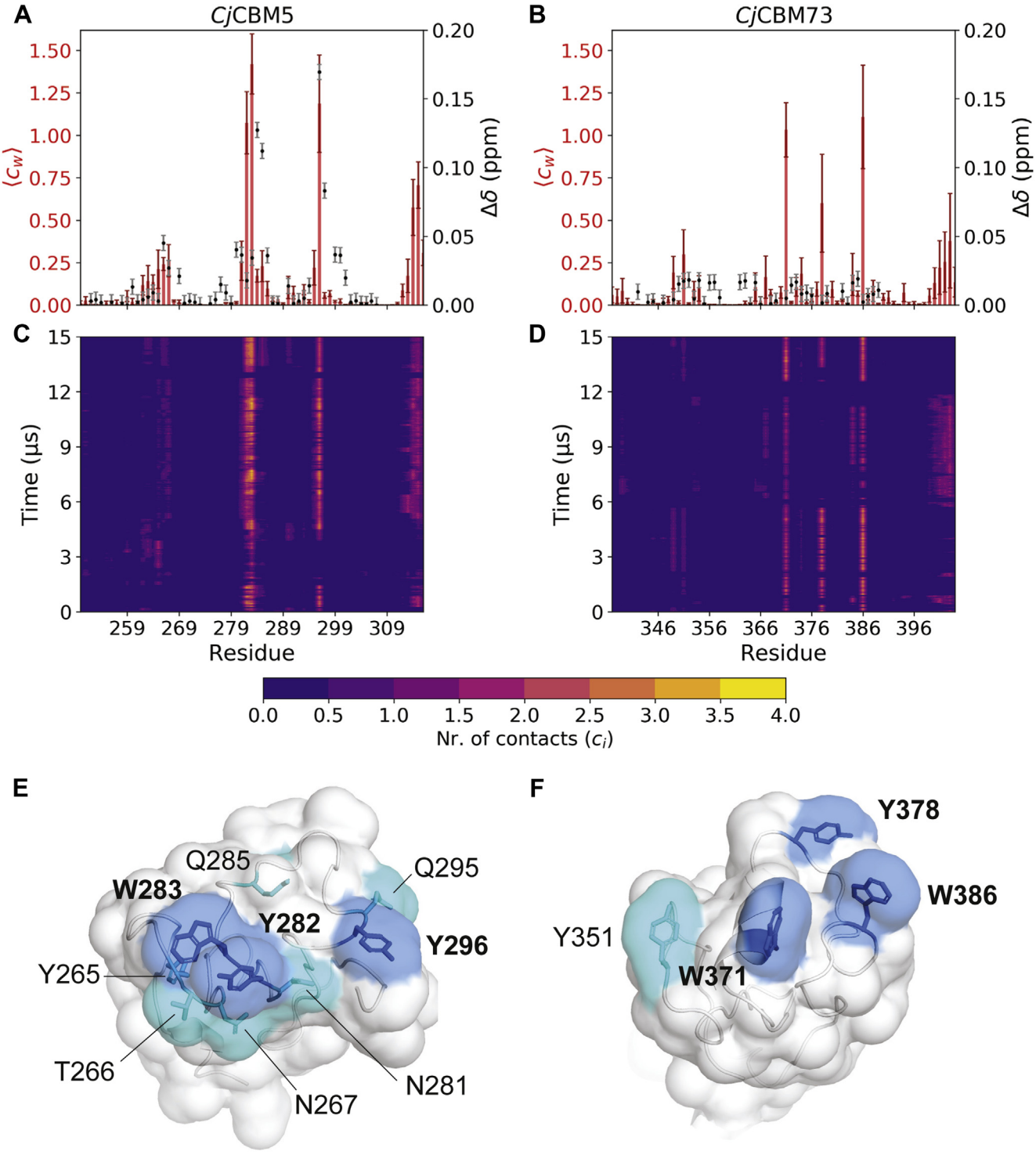

Figure 7. Chitin binding probed by NMR and simulations for *Cj*CBM5 and *Cj*CBM73. Panels *A* and *B* show the weighted average number of contacts observed during simulations ($\langle c \rangle_w$; *red bars* with *dark red error bars*) and the chemical shift perturbations (Δδ; *black dots* with *gray error bars*) observed by NMR upon addition of $(GlcNAc)_6$. The error bars for the number of contacts were calculated using block analysis (89); error bars for chemical shift perturbations correspond to 0.003 ppm. Note that no significant chemical shift perturbations were recorded for *Cj*CBM73. Panels *C* and *D* show the number of contacts between each amino acid and the α-chitin surface per frame of the 15 μs simulations, c_i, using a cutoff distance of 0.3 nm (see Experimental procedures section for details). We note that because of the use of coarse-grained models and because of the use of metadynamics, that is, enhanced sampling, the time scales do not here correspond to a physical time scale. Panels *E* and *F* show the substrate-binding surfaces of representative conformations of the bound state of *Cj*CBM5 and *Cj*CBM73, respectively. The side chains of amino acids on the binding surface that have most contacts ($\langle c_w \rangle > 0.5$) with chitin are colored *blue*, whereas the side chains of amino acids with fewer contacts ($0.2 < \langle c_w \rangle < 0.5$) are colored *cyan*.

with much fewer contacts. These observations are in good agreement with the chemical shift perturbation data for binding of $(GlcNAc)_6$. Similar observations were made for *Cj*CBM73 (Fig. 7, *B* and *D*), in the sense that also in this case the interacting regions include, and are to a large extent limited to, the three aromatic residues of the putative binding surface (W371, Y378, and W386). Furthermore, also in this case, interactions with fewer contacts ($0.2 < \langle c_w \rangle < 0.5$) with a fourth aromatic residue, Y351, were observed. In order to match the experiments as closely as possible, we included the C-terminal His in the simulations and found that these have a number of contacts with the substrate (Fig. 7, *A* and *B*). All in all, these analyses show that the amino acids on the surface of *Cj*CBM5 with most chitin contacts form a relatively linear arrangement (Fig. 7*E*), perhaps reflecting that interactions are limited to a single chitin chain, whereas the arrangement of aromatic amino acids on the surface of *Cj*CBM73 is wider and suggests a more extended substrate-binding surface (Fig. 7*F*).

Discussion

According to the Pfam database, only about 25% of AA10 LPMOs (Pfam ID: PF03067) contain one or more additional domains, and the large majority of these multimodular enzymes contain a CBM. While single-domain LPMOs can be efficient and may bind well to their substrates, as exemplified by the archetypal chitin-active LPMO CBP21 (6, 36, 37), CBMs tethered to the LPMO domain are known to have significant impact on the catalytic efficiency of multimodular LPMOs (22, 24, 25, 38). Therefore, it is important to gain a deeper understanding of the mechanisms by which CBMs recognize and bind their target substrates. Here, we have investigated two CBMs from *Cj*LPMO10A: *Cj*CBM5 and *Cj*CBM73, to illuminate structural and functional differences between these chitin-binding domains. The present results include the first structure for a member of the CBM73 family.

The NMR solution structures show that, although both CBMs have similar overall folds, *Cj*CBM73 has a 3_{10}-helix connected by an additional disulfide bridge. These features appear to be conserved in the CBM73 family (Fig. S2). To obtain further insight into the structural variation between small chitin-binding CBMs, we compared the structures of *Cj*CBM5 and *Cj*CBM73 with the structures of five CBM5s and a CBM12 (Fig. 8). The CBM12 is included because the CBMs in this family are closely related to family 5 CBMs (18). It has previously been shown (39–41) that, in addition to conserved surface-exposed aromatic residues, these CBMs share two additionally conserved aromatic amino acids (Y265 and W299 in *Cj*CBM5) that also occur in CBM73s (Y351 and W390 in *Cj*CBM73; Fig. S2). These residues are a part of the hydrophobic core of the proteins. All CBMs (39–45) in Figure 8 bind chitin.

Previous studies (40, 41) have established the importance of the two consecutive and conserved aromatic residues, Y–W or W–W, in family 5 CBMs (Y282–W283 in *Cj*CBM5). Site-directed mutagenesis studies have shown that a third aromatic residue, Y296, present on the surface of *Cj*CBM5, *Pf*ChiA_CBM5 and *Mm*Chi60_CBM5, also contributes to chitin binding (41). Whereas the NMR titration experiment with soluble chitohexaose did not show binding for *Cj*CBM73, results for *Cj*CBM5 showed that both W283 and Y296 are involved in binding $(GlcNAc)_6$. In addition, the polar residues T284 and Q285 also appear to contribute to binding $(GlcNAc)_6$. These observations suggest that chitin binding by *Cj*CBM5 likely involves a combination of CH–π interactions (34) and hydrogen bonding. Binding of chitohexaose to a CBM5 has previously been addressed by Itoh *et al.* (46) and Akagi *et al.* (40) for *Sg*ChiC_CBM5 using isothermal titration calorimetry and NMR titration, respectively. Interestingly, these studies yielded dissociation constants of 2 mM and 1.6 ± 0.3 mM, respectively, which are consistent with the dissociation constant determined here for *Cj*CBM5 (K_d = 2 ± 1 mM).

*Cj*CBM5, like other CBM5s in Figure 8, has three exposed aromatic residues with a close to linear arrangement of the side chains on the surface. This type of arrangement is often found in cellulose-binding domains (47–49), where the distance between the three aromatic residues coincides with the spacing of every second glucose ring in a single chain (48, 50). Compared with the other CBMs in Figure 8, the arrangement of the three exposed aromatic residues in *Cj*CBM73: W371, Y378, and W386, differs, which suggests that *Cj*CBM73 has a wider binding surface that may interact with several chitin chains. This could explain why *Cj*CBM73 cannot bind $(GlcNAc)_6$, a single-chain analog, whereas *Cj*CBM5 can.

The distinct arrangements of amino acids on the binding surfaces of *Cj*CBM5 and *Cj*CBM73 may also explain the experimentally and computationally observed differences in binding to α-chitin (Figs. 6 and 7). The side chains of the amino acids with most chitin contacts in the simulations (Fig. 7*E*) form a linear arrangement in *Cj*CBM5 but are distributed on a larger and wider surface in *Cj*CBM73 (Fig. 7*F*). Both experiments and simulations indicated that *Cj*CBM5 binds to chitin with lower affinity compared with *Cj*CBM73 (Table 1), whereas this is the other way around for (single chain) $(GlcNAc)_6$. The stronger affinity of *Cj*CBM73 for insoluble α-chitin can be explained by its binding surface covering a larger area than the binding surface of *Cj*CBM5.

At low substrate concentrations, the catalytic performance of *Cj*LPMO10A^{FL} and *Cj*LPMO10A$^{\Delta CBM73}$ is superior to that of *Cj*LPMO10A^{CD}, and the progress curves in Figure 2, *A* and *B* show that this is due to rapid inactivation of *Cj*LPMO10A^{CD}. Forsberg *et al.* (25) have previously shown that almost all the binding affinity for chitin in *Cj*LPMO10A^{FL} resides on the CBMs. The strong binding provided by the CBMs ensures that the LPMO stays close to its substrate, thus increasing the chances that the interaction of the reduced CD with the oxygen cosubstrate leads to a productive reaction (*i.e.*, cleavage of chitin) rather than futile turnover that may lead to autocatalytic enzyme inactivation (20), as has previously been observed for other CBM-containing LPMOs (22, 38). At the highest substrate concentration (Fig. 2*C*), inactivation of *Cj*LPMO10A^{CD} was much reduced, likely because the high substrate load favors *Cj*LPMO10A^{CD} binding to chitin, reducing the frequency of futile turnovers and the concurrent

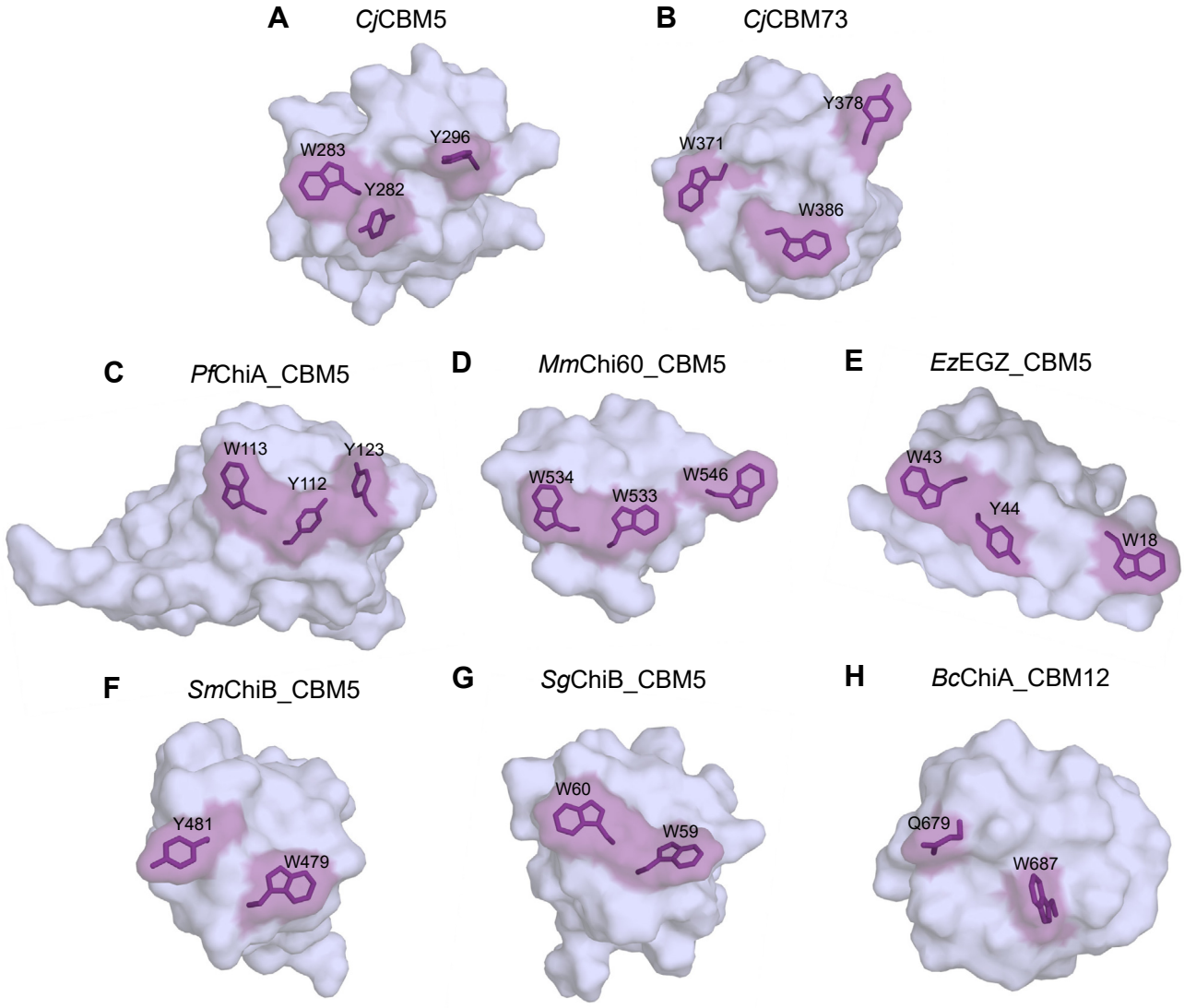

Figure 8. Comparison of the binding surfaces of the NMR structures of (*A*) *Cj*CBM5 (PDB ID: 6Z40) and (*B*) *Cj*CBM73 (PDB ID: 6Z41) with the structures of other CBM5 domains and one CBM12 domain. The other structures are derived from (*C*) PfChiA_CBM5 (41) (PDB ID: 2RTS; NMR structure), (*D*) MmChi60_CBM5 (44) (PDB ID: 4HMC; X-ray diffraction structure), (*E*) EcEGZ_CBM5 (33) (PDB ID: 1AIW; NMR structure), (*F*) SmChiB_CBM5 (43) (PDB ID: 1E15; X-ray diffraction structure), (*G*) SgChiC_CBM5 (40) (PDB ID: 2D49; NMR structure), and (*H*) BcChiA_CBM12 (39, 45) (PDB ID: 1ED7; NMR structure). Residues shown or predicted to be involved in substrate binding are highlighted in *purple*. PDB, Protein Data Bank.

risk of enzyme inactivation. This observation is in agreement with a previous study (22) showing that the negative effect of truncation of the CBM2 from a two-domain cellulose-active LPMO was smaller at higher substrate concentrations.

The protective effect of the substrate, mediated by the CBMs, became more evident at higher temperatures (Fig. 3), where *Cj*LPMO10A^{FL} showed higher catalytic performance than *Cj*LPMO10A$^{\Delta CBM73}$, indicating that *Cj*CBM73 appears to provide additional protection to the enzyme from thermal inactivation. It is conceivable that the increased performance at higher temperatures translates into increased performance at lower, more physiologically relevant temperatures where the enzyme may experience other types of stress, such as very low substrate concentrations or high levels of oxidant.

The aforementioned previous study with a two-domain cellulose LPMO (22) shows that the anchoring effect of the CBMs leads to a higher fraction of soluble oxidized products relative to oxidized sites on the insoluble substrate. A similar effect was also observed for *Cj*LPMO10A^{FL} and *Cj*LPMO10A$^{\Delta CBM73}$, which produced a higher fraction of soluble oxidized products compared with *Cj*LPMO10A^{CD} (Fig. 2, *D–F*). Interestingly, this difference became less at higher substrate concentrations, which may perhaps be due to the fact that higher substrate concentrations increase the chance that a substrate-anchored, but otherwise freely moving CD, acts on a neighboring fibril rather than the fibril to which it is bound, as discussed previously (22).

Considering the LPMO reaction cycle and considering that anchoring by the CBMs could lead to multiple oxidized sites localized on the chitin surface around the CBM-binding site, it is conceivable that accumulation of oxidized sites could trigger unbinding of the CBMs. The results of our attempts to test this

hypothesis by studying CBM binding to partially oxidized chitin (Fig. 6) were not conclusive but did indicate that substrate oxidation slightly weakened chitin binding by *Cj*CBM5.

In conclusion, we have revealed structural and functional variation between the two chitin-binding domains in *Cj*LPMO10A. While it is clear that the presence of these CBMs has a significant effect on the catalytic performance of the LPMO, the question why nature has evolved enzymes with two different chitin-binding domains remains to be answered. Chitin occurs in different crystalline forms and may be intertwined with other polymers, such as β-glucans in fungal cell walls or proteins in crustacean shells. Perhaps, some CBMs are adapted to interacting with chitin in specific copolymeric contexts that are absent in the heavily processed α-chitin used in this study. Indeed, it is possible that functional differences between the CBM5 and the CBM73 remain undetected in the present experiments because of the choice of substrate. The possibility that different CBMs are adapted to different crystalline forms or different faces of a chitin fiber is supported by multiple studies in the cellulose field, which have shown that cellulose-binding domains belonging to different CBM families bind at different locations on cellulose crystals (26, 51). Interestingly, chitin-binding studies with a chitinase from *Thermococcus kodakaraensis*, which contains two CDs and three CBMs, led Kikkawa *et al.* (52) to propose that two highly similar chitin-binding CBM2 domains ensure binding to the chitin surface, whereas a CBM5 would ensure binding to single chitin chains, perhaps near chain ends. It is conceivable that a similar scenario applies to *Cj*LPMO10A; in this case, the CBM73 would promote binding to the chitin surface, whereas the CBM5, which was shown to bind soluble chitooligosaccharides, would promote binding to single chains. Eventually, insights into these two CBMs will increase our understanding of how LPMOs depolymerize insoluble polysaccharides.

Experimental procedures

Cloning, expression, and purification of CjLPMO10A variants

The gene encoding *Cj*LPMO10A^{FL} (residues 1–397) was codon optimized for *Escherichia coli* expression. *Cj*LPMO10A^{CD} (residues 1–216) was cloned into the pRSET B expression vector (Invitrogen) as previously described (25), as well as the construct lacking the CBM73 and the preceding serine-rich linker, named *Cj*LPMO10A$^{\Delta CBM73}$ (residues 1–307).

To obtain better expression of *Cj*LPMO10A^{FL}, the codon-optimized gene encoding mature *Cj*LPMO10A^{FL} (residues 37–397) was cloned behind an IPTG-inducible T7 promoter in the pD441-CH expression vector by ATUM, resulting in a fusion construct with an N-terminal *E. coli* OmpA signal peptide and a C-terminal His_6 motif (Gly-$(His)_6$). The expression vector was transformed into chemically competent *E. coli* BL21 (New England Biolabs). Production of *Cj*LPMO10A^{FL} was achieved by fed-batch fermentation of the expression strain in a 1-l fermenter (DASGIP benchtop bioreactors for cell culture; Eppendorf), essentially as described previously (53), with the following modifications: at the start of the feed phase, the temperature was switched to 25 °C, and 0.6 mM IPTG was added to the glucose feed solution for continuous induction of gene expression. After 18 h of glucose feed, the cells were removed by centrifugation. The culture supernatant containing the target protein was concentrated threefold and buffer exchanged against six volumes of working buffer (50 mM Tris–HCl, 300 mM NaCl, and pH 8.0) by crossflow filtration (Millipore Pellicon 2 mini filter, regenerated cellulose, 3 kDa molecular weight cutoff [MWCO]). After centrifugation for 30 min at 35,000*g* to remove precipitated proteins and filtration through a 0.2 μm Nalgene Rapid-Flow sterile bottle-top filter unit (Thermo Scientific), the culture filtrate was applied to a 20-ml nickel–nitrilotriacetic acid sepharose column connected to an ÄKTA express FPLC system (GE Healthcare Life Sciences). After washing with ten column volumes (CVs) of working buffer containing 20 mM imidazole, bound protein was eluted with a buffer containing 200 mM imidazole. Fractions containing the target protein were pooled and buffer exchanged into 20 mM Tris–HCl, 200 mM NaCl, pH 7.5 by gel filtration over Sephadex G25 (GE Healthcare, 4× HiPrep Desalting 26/10 columns).

*Cj*LPMO10A$^{\Delta CBM73}$ and *Cj*LPMO10A^{CD} were expressed in lysogeny broth (LB) media containing 50 μg/ml ampicillin. Cells harboring the plasmid were grown at 30 °C for 24 h, without any induction, prior to harvest. The protein was extracted from the periplasmic space using an osmotic shock method that was first described by Manoil and Beckwith (54), followed by purification using a two-step chromatography protocol. The periplasmic extract was adjusted to 50 mM Tris–HCl at pH 9.0 (loading buffer) and loaded onto a 5 ml Q Sepharose anion exchange column (GE Healthcare). Proteins were eluted using a linear salt gradient (0–500 mM NaCl) over 60 CVs using a flow rate of 2.5 ml/min. LPMO-containing fractions were pooled and concentrated to 1 ml before being loaded onto a HiLoad 16/60 Superdex 75 size exclusion column (GE Healthcare) operated with a running buffer consisting of 50 mM Tris at pH 7.5 and 200 mM NaCl, at a flow rate of 1 ml/min. Fractions containing pure LPMO were identified by SDS-PAGE and subsequently pooled and concentrated using Amicon Ultra centrifugal filters (Millipore) with an MWCO of 10 kDa. Protein concentrations were measured using the Bradford assay (Bio-Rad). The protein solutions were stored at 4 °C until further use.

Typical yields of purified protein were 95, 10, and 10 mg per liter of culture for *Cj*LPMO10A^{FL}, *Cj*LPMO10A$^{\Delta CBM73}$, and *Cj*LPMO10A^{CD}, respectively. The absence of free copper in the preparations of purified LPMOs was confirmed by measuring hydrogen peroxide production upon addition of ascorbic acid using the Amplex Red assay as described by Kittl *et al.* (55). The presence of free copper would lead to drastically increased levels of hydrogen peroxide production (56), and this was not observed.

Expression plasmids for *Cj*CBM5 (residues 251–309) and *Cj*CBM73 (residues 338–397) based on the pNIC-CH vector (Addgene) were used for cytoplasmic expression as previously described (25). This cloning procedure adds a Met residue to

the N terminus as well as one Ala residue and a polyhistidine tag (6×His tag) to the C terminus of both proteins. Precultures in 5 ml LB medium (10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl) were used to inoculate 500 ml of terrific broth medium supplemented with 50 μg/ml kanamycin. The cultures were grown at 37 °C for approximately 3 h in a LEX-24 Bioreactor (Harbinger Biotechnology) using compressed air for aeration and mixing. Expression was induced by adding IPTG to a final concentration of 0.1 mM at an absorbance of 0.6 at 600 nm (OD_{600}), followed by incubation for 24 h at 23 °C. Cells were harvested by centrifugation (5500*g*, 10 min) followed by cell lysis using pulsed sonication in a buffer containing 50 mM Tris–HCl at pH 8.0, 500 mM NaCl, and 5 mM imidazole. Cell debris was removed by centrifugation (75,000*g*, 30 min), and the supernatant was loaded onto a 5 ml HisTrap HP Ni Sepharose column (GE Healthcare) equilibrated with lysis buffer. The protein was eluted by applying a 25 CV linear gradient to reach 100% of a buffer containing 50 mM Tris–HCl at pH 8.0, 500 mM NaCl, and 500 mM imidazole, at a flow rate of 2.5 ml/min. Protein-containing fractions were analyzed by SDS-PAGE and subsequently concentrated, with concomitant buffer exchange to 20 mM Tris–HCl at pH 8.0, using an Amicon Ultra centrifugal filter (Millipore) with a 3 kDa cutoff. The concentrations of *Cj*CBM5 and *Cj*CBM73 were determined by measuring absorbance at 280 nm (A_{280}) and calculated using theoretical molar extinction coefficients ($\varepsilon_{280,\ CjCBM5}$ = 22,585 M^{-1} cm^{-1}; $\varepsilon_{280,\ CjCBM73}$ = 28,210 M^{-1} cm^{-1}). Typical yields of these procedures were 5 and 2 mg of pure protein per liter of culture for *Cj*CBM5 and *Cj*CBM73, respectively.

Production of CjCBM5 and CjCBM73 for NMR studies

*Cj*CBM5 and *Cj*CBM73 samples for NMR studies were produced both with ^{13}C and ^{15}N isotopic labeling and ^{15}N labeling only. A preculture was grown in 6 ml LB medium supplemented with 50 μg/ml kanamycin in a shaking incubator at 30 °C, 225 rpm, for 6 h. A main culture of 500 ml M9 medium (6 g/l Na_2HPO_4, 3 g/l KH_2PO_4, and 0.5 g/l NaCl) supplemented with 500 μg/ml kanamycin, 0.5 g $(^{15}NH_4)_2SO_4$, 6 ml glycerol, 5 ml ^{15}N Bioexpress Cell Growth Medium (Cambridge Isotope Laboratories), 5 ml Gibco MEM Vitamin Solution (100×), 1 ml $MgSO_4$ (1 M), and 5 ml of a trace-metal solution (0.1 g/l $ZnSO_4$, 0.8 g/l $MnSO_4$, 0.5 g/l $FeSO_4$, 0.1 g/l $CuSO_4$, and 1 g/l $CaCl_2$) was inoculated with 1% of the preculture and incubated at 22 °C in a LEX-24 Bioreactor as described previously. After 18 h, the cultures induced with 0.5 mM IPTG to a final concentration of 0.5 mM were followed by incubation at 22 °C for 24 h. Cells were harvested by centrifugation at 4 °C, 6000*g*, for 5 min. The pellet was resuspended in 20 ml lysis buffer (50 mM Tris–HCl, 50 mM NaCl, 0.05% Triton X-100, and pH 8.0) supplemented with a tablet EDTA-free cOmplete ULTRA protease inhibitor (Roche) followed by pulsed sonication. Cell debris was removed by centrifugation at 4 °C, 16,600*g*, for 45 min. The supernatant was sterilized by filtration through a 0.2 μm Sterile-flip filter unit (Nalgene). Buffer B (50 mM Tris–HCl, 400 mM imidazole, and pH 8.0) was added to the filtered lysate to obtain a final concentration of 20 mM imidazole. The proteins were purified by loading the supernatant onto a 1 ml HisTrap HP Ni-sepharose column (GE Healthcare Life Sciences) equilibrated with 5 CV of 95% buffer A (50 mM Tris–HCl, pH 8.0) and 5% buffer B with a flow rate of 1 ml/min. Impurities were removed by washing with 95% buffer A and 5% buffer B for 10 CV. The protein was eluted using a 30 CV gradient of 5 to 100% buffer B. The purity of the protein fractions was assessed with SDS-PAGE. The yields of the labeled proteins were 1 and 0.2 mg per liter of culture for *Cj*CBM5 and *Cj*CBM73, respectively.

The protocol for production and purification of nonlabeled samples of *Cj*CBM5 and *Cj*CBM73 for NMR studies was as described above, except that 2× LB medium (20 g/l tryptone, 10 g/l yeast extract, and 5 g/l NaCl) was used instead of M9. The yields of the nonlabeled proteins were 4 and 3 mg per liter of culture for *Cj*CBM5 and *Cj*CBM73, respectively.

Fractions shown to contain *Cj*CBM73 were pooled and concentrated using Amicon Ultra protein concentrators (MWCO = 3 kDa) at 10 °C and 7000*g* to obtain a volume of ∼5 ml. This protein solution was loaded onto a size-exclusion chromatography column (HiLoad 16/600 Superdex 75 pg; 120 ml CV) that had been equilibrated with 1 CV of size-exclusion chromatography-buffer pH 7.5 (50 mM Tris–HCl and 20 mM NaCl). Protein fractions were eluted using a 1 ml/min flow rate, and the concentration was measured as mentioned previously.

The buffer in the protein-containing fractions was exchanged to NMR buffers (for structure elucidation: 25 mM sodium phosphate and 10 mM NaCl, pH 5.5; for interaction studies: 50 mM sodium phosphate [*Cj*CBM5] or 25 mM sodium phosphate [*Cj*CBM73], pH 7.0) prior to concentrating to ∼70 μM and a final volume of ∼400 μl. All steps were performed by centrifugation using Amicon Ultra protein concentration tubes (MWCO = 3 kDa) at 10 °C and 7000*g*. NMR samples were prepared by adding D_2O to a final ratio of 90% H_2O/10% D_2O.

Chitin degradation experiments

Unless stated otherwise, reactions were performed with 0.5 μM LPMO in 50 mM sodium phosphate buffer at pH 7.0 in the presence of 1 mM ascorbic acid at 37 °C and 800 rpm in an Eppendorf thermomixer. All reactions were performed in triplicates.

Preparation of oxidized chitin for binding studies

*Cj*LPMO10A^{CD}, which is known to bind weakly to α-chitin (25) and which was expected to oxidize the chitin surface more randomly compared with the full-length enzyme (Fig. 1; (22)), was used to prepare oxidized chitin. Six 1-ml reactions, each containing 20 g/l α-chitin suspended in 50 mM sodium phosphate at pH 7.0, were supplemented with 1 μM *Cj*LPMO10A^{CD} and 1 mM ascorbic acid three times with 24-h intervals (*i.e.*, to a final concentration of 3 μM enzyme and 3 mM ascorbic acid). The reactions were incubated in a

thermomixer set to 37 °C and 800 rpm. After 72 h of incubation, samples were taken from all six reactions and diluted in buffer supplemented with 2 μM *Sm*ChiA (57) and 0.5 μM *Sm*CHB (58) to a substrate concentration of 2 g/l. These reaction mixtures were incubated for 24 h at 37 °C at 800 rpm after which oxidized products were analyzed quantitatively to determine the total degree of oxidation in the LPMO-treated chitin. The rests of the 20 g/l reactions were centrifuged in an Eppendorf centrifuge (12,000*g* for 3 min), the supernatant was removed, and the soluble products in the supernatant were subjected to degradation with 2 μM *Sm*ChiA and 0.5 μM *Sm*CHB as described previously, to determine the amount of solubilized oxidized products. The pelleted oxidized chitin was washed with buffer (3 × 1 ml of 50 mM sodium phosphate at pH 7.0) by repetitively suspending the chitin in buffer and removing the supernatant after centrifugation. Finally, the oxidized chitin was suspended in buffer to 20 g/l. Again, samples were taken from all six reactions and diluted in buffer supplemented with 2 μM *Sm*ChiA and 0.5 μM *Sm*CHB to a substrate concentration of 2 g/l. The reactions were incubated for 24 h at 37 °C at 800 rpm, and the resulting samples were used to determine the amount of insoluble oxidized products. Compounds in *Sm*ChiA/*Sm*CBH degraded samples were quantified as described later.

Quantitative analysis of chitobionic acid (GlcNAcGlcNAc1A)

Prior to product quantification, LPMO-generated products were degraded with only *Sm*CHB (soluble fractions) or a combination with *Sm*ChiA and *Sm*CHB (for total or insoluble fractions) to yield a mixture of nonoxidized monomeric GlcNAc and the oxidized dimer, chitobionic acid, which consists of a GlcNAc and an oxidized GlcNAc in the aldonic acid form (GlcNAc1A). Analysis and quantification of GlcNAcGlcNAc1A were carried out using an RSLC system (Dionex) equipped with a 100 × 7.8 mm Rezex RFQ-Fast Acid H+ (8%) (Phenomenex) column operated at 85 °C. Samples of 8 μl were injected to the column, and sugars were eluted isocratically using 5 mM sulphuric acid as mobile phase with a flow rate of 1 ml/min. Standards of GlcNAcGlcNAc1A (10–500 μM) were used for quantification. GlcNAcGlcNAc1A was generated in house by complete oxidation of *N*-acetylchitobiose (Megazyme; 95% purity) by the *Fusarium graminearum* chitooligosaccharide oxidase as previously described (58, 59).

Determination of apparent Tm

The apparent T_m of the proteins was determined according to a protein thermal shift assay (Thermo Fisher Scientific) based on using SYPRO orange, a fluorescent dye, to monitor protein unfolding (60). The quantum yield of the dye is significantly increased upon binding to hydrophobic regions of the protein that become accessible as the protein unfolds. The fluorescence emission (relative fluorescence unit) was monitored using a StepOnePlus real-time PCR machine (Thermo Fisher Scientific). T_m was calculated as the temperature corresponding to the minimum value of the derivative plot (−d[relative fluorescence unit]/dT *versus* T; Fig. S1). 0.1 g/l LPMO in 50 mM sodium phosphate buffer (pH 7.0) was heated in the presence of the dye in a 96-well plate from 25 to 95 °C, over 50 min. For each protein, the experiment was carried out in quadruplicates (*i.e.*, n = 4).

Binding studies with Cj*CBM5 and* Cj*CBM73*

Binding studies were performed as previously described (25). The equilibrium binding constants (K_d) and binding capacity (B_{max}) were determined for *Cj*CBM5 and *Cj*CBM73 by mixing protein solutions of varying concentrations (0, 20, 50, 75, 100, 150, 300, and 500 μg/ml for *Cj*CBM5 and 0, 10, 20, 50, 75, 100, 150, and 300 μg/ml for *Cj*CBM73; protein concentration was determined by A_{280}) with 10 mg/ml preoxidized (see aforementioned one) or untreated α-chitin. Before adding the chitin, A_{280} was measured for each of the prepared protein solutions (in 50 mM sodium phosphate buffer, pH 7.0), to create individual standard curves for each protein. After addition of chitin, the solutions were placed at 22 °C in an Eppendorf Comfort Thermomixer set to 800 rpm for 60 min. Subsequently, samples were filtered using a 96-well filter plate (Millipore), and the concentration of free protein in the supernatant was determined by measuring A_{280}. All assays were performed in triplicate and with blanks (buffer and 10 mg/ml α-chitin). The equilibrium dissociation constants, K_d (μM), and substrate-binding capacities, B_{max} (μmol/g α-chitin), were determined by fitting the binding isotherms to the one-site binding equation, where *P* represents protein: $[P_{bound}] = B_{max}[P_{free}]/(K_d + [P_{free}])$, by nonlinear regression using the Prism 6 software (GraphPad Software, Inc).

NMR spectroscopy

NMR spectra of 70 μM *Cj*CBM5 and *Cj*CBM73 in NMR buffer (25 mM sodium phosphate and 10 mM NaCl, pH 5.5) containing 10% D_2O were recorded at 25 °C on a Bruker Ascend 800 MHz spectrometer with an Avance III HD (Bruker Biospin) console equipped with a 5 mm Z-gradient CP-TCI (H/C/N) cryogenic probe at the NV-NMR-Centre/Norwegian NMR Platform at NTNU, the Norwegian University of Science and Technology. ^{1}H chemical shifts were referenced internally to the water signal, whereas ^{13}C and ^{15}N chemical shifts were referenced indirectly to water based on the absolute frequency ratios (61). Backbone and side-chain assignments of *Cj*CBM5 and *Cj*CBM73 were obtained using ^{15}N-heteronuclear single quantum coherence (HSQC), ^{13}C-HSQC, HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCANHHNCACB, CBCA(CO)NH, and H(C)CH-TOCSY. For *Cj*CBM5, the band-selective excitation short-transient (62) versions of HNCA, HN(CO)CA, HNCO, HN(CA)CO, and HN(CO)CACB were recorded. The assignments have been deposited in the BMRB under the IDs 34519 (*Cj*CBM5) and 34520 (*Cj*CBM73).

Structure elucidation

The NMR data were recorded and processed with TopSpin version 3.6 (Bruker), and analyzed with CARA version 1.5.5 (63). For structure determination, 3D ^{13}C-edited and ^{15}N-edited NOESY-HSQC spectra as well as 2D ^{1}H-^{1}H NOESY

spectra were recorded. NOE crosspeaks were manually identified, assigned, and integrated using the NEASY program within CARA version 1.5.5. Dihedral torsion angles (ϕ and ψ) were calculated from chemical shift data (C^{α}, C^{β}, H^{N}, H^{α}, H^{β}, N, and C′) by TALOS-N (64). Structures were calculated using the torsion angle dynamics program CYANA version 3.97 (65). The structure calculation started by generating 200 conformers with random torsion angles and the dihedral angles in each conformer were optimized using simulated annealing in 10,000 steps to fit the restraints. The 20 conformers with the lowest CYANA target function values were energy minimized using YASARA (66), first *in vacuo*, followed by using water as the explicit solvent and calculating electrostatics by applying the particle mesh Ewald method (67). In both these steps, the YASARA force field (68) was applied. The coordinates of the minimized CBM conformers have been deposited in the PDB under the IDs 6Z40 (*Cj*CBM5) and 6Z41 (*Cj*CBM73). The two structures were aligned using the combinatorial extension algorithm, which determines the longest continuous alignment between fragment pairs (69).

Titration of CBMs with chitohexaose

The interaction between *Cj*CBM5 and chitohexaose, $(GlcNAc)_6$ was investigated using NMR spectroscopy. A ^{15}N-HSQC spectrum was recorded of a sample of ^{15}N-labeled *Cj*CBM5 (70 μM) in 50 mM sodium phosphate containing 10% D_2O and used as reference. Another sample of ^{15}N-labeled *Cj*CBM5 (70 μM) with 10 mM $(GlcNAc)_6$ in 50 mM sodium phosphate containing 10% D_2O was prepared. After recording a ^{15}N-HSQC spectrum of this latter sample, it was mixed with the reference sample to obtain the following concentrations of $(GlcNAc)_6$ 0.2, 0.5, 1.0, and 5.0 mM, while maintaining a constant protein concentration. A new ^{15}N-HSQC spectrum was recorded at each $(GlcNAc)_6$ concentration. Chemical shift perturbations (Δδ) were calculated using the equation $\Delta\delta = \sqrt{(\Delta\delta_H)^2+(\Delta\delta_N/6.5)^2}$, where $\Delta\delta_H$ is the change in chemical shift for the amide proton and $\Delta\delta_N$ for the amide nitrogen, in ppm (70). The K_d was estimated using Gnuplot 5.2 (www.gnuplot.info) based on an average of the amide chemical shift perturbation (Δδ) from the five most affected amino acids (W283, T284, Q285, Y296, and G297). The function used for fitting was $\Delta\delta = \Delta\delta_{max}[S]/(K_d+[S])$, where $\Delta\delta_{max}$ describes the binding capacity as the maximum value of Δδ. Error bars in the chemical shift measurements correspond to 0.003 ppm.

The same procedure was applied for the NMR titration of *Cj*CBM73 with $(GlcNAc)_6$, using *Cj*CBM73 (500 μM) and the following $(GlcNAc)_6$ concentrations: 0.2, 0.6, 1.2, 2.8, 5.6, and 6.5 mM.

Modeling CG α-chitin

The crystal structure of α-chitin at 300 K (71) was used to generate an all-atom α-chitin surface composed of 12 chains with 20 residues each, in UCSF Chimera, version 1.13.1 (72). The all-atom model was coarse-grained using the bead mapping and topology parameters proposed by Yu and Lau (73) and bead types from the CG Martini version 3.0.beta.4.17 force field (35). A rectangular simulation box was defined with the same size as the surface in the *y* (110 Å) and *z* (100 Å) dimensions and 150 Å in the *x* dimension.

CG simulations

The CG Martini version 3.0.beta.4.17 force field was used in combination with GROMACS version 5.1.4 (74) to simulate interactions between the CBMs and α-chitin. The constructs were coarse-grained using the Martinize2 program (75). An elastic network model (76) was used to constrain the overall structure of the CBMs. The beads in the chitin surface were kept in place by applying a harmonic potential with a force constant of 1000 kJ mol^{-1} nm^{-1} on the *x*, *y*, and *z* positions. Models of *Cj*CBM5 or *Cj*CBM73 were manually placed in the simulation box above the chitin surface by using PyMOL (Schrödinger, Inc) (77). The simulation box was filled with water beads, and the system was neutralized with beads corresponding to Na^+ and Cl^- ions to an ionic strength of 0.15 M. The complex was energy minimized using a steepest-descent algorithm (100 steps, 0.03 nm maximum step size) prior to being relaxed for 1 ns, with a time step of 5 fs, using the velocity-rescale thermostat (78), Parrinello–Rahman barostat (79), and Verlet cutoff scheme (80). Simulations were run on the relaxed models with a time step of 20 fs, using the velocity-rescale thermostat, Parrinello–Rahman barostat, and Verlet cutoff scheme in the isothermal–isobaric (NPT) ensemble. Frames were written every 1 ns for each trajectory. Representative conformations of CG models of *Cj*CBM5 and *Cj*CBM73 were backmapped to atomistic models by using the CG2AT2 program (81).

Adjusting the protein–chitin interaction strength in the Martini force field

Initially, binding between CBMs and chitin was not observed; therefore, drawing inspiration from Larsen *et al.* (82), we considered the following approach to modify the Martini version 3.0.beta.4.17 force field to promote CBM binding to chitin. First, the interaction strength (*i.e.*, ε parameter in the Lennard–Jones potential) between chitin beads and protein beads was increased by 10%, and WT-MetaD simulations (see later) were run until binding between the CBMs and chitin was observed. Then, a binding path for each CBM, which included bound and unbound conformations, was selected from the WT-MetaD simulations, and umbrella-sampling simulations were performed on these conformations as described later. Finally, we tested different interaction strengths by generating topologies where the interaction strength was modified from 0% (unchanged) to 15% increase of the chitin–protein interaction strength and ran umbrella-sampling simulations for each interaction strength. Fig. S3 shows the free-energy surfaces calculated for each umbrella-sampling simulation using the weighted histogram analysis method (WHAM) (http://membrane.urmc.rochester.edu/?page_id=126) (83). Dissociation constants calculated from each free-energy surface are shown in Table 1.

WT-MetaD simulations

WT-MetaD simulations (84) were performed using the PLUMED 2.5 plugin (85–87) and the Martini model where chitin–protein interactions had been increased by 10%. We used the distance (r_{chitin}) and number of contacts (cutoff distance, r_0 = 0.7 nm) between aromatic residues in the putative substrate-binding surfaces (*Cj*CBM5: Y282, W283, and Y296; *Cj*CBM73: W371, Y378, and W386) and the chitin surface in 15-μs long (15,000 frames) simulations as collective variables to sample the binding of CBMs to chitin. Gaussian hills were added every 10 ps, with a starting height of 2.0 kJ mol^{-1}, width of 0.5, and bias factor of 50. Fig. S4 shows the evolution of the collective variables and deposition of Gaussian hills over the course of the simulation.

Since the modeled chitin surface corresponds to the *x*,*y* plane in the simulation box, r_{chitin}, was calculated using Equation 1 (shown below with *Cj*CBM5 as an example), which uses the geometric center of the *z* coordinates, z_{gc}, calculated using Equation 2. Here, z_i is the z coordinate of each amino acid or chitin bead, and N is the number of beads in the amino acid or chitin.

$$r_{chitin} = \sqrt{\left(z_{gc,Y282} - z_{gc,chitin}\right)^2 + \left(z_{gc,W283} - z_{gc,chitin}\right)^2 + \left(z_{gc,Y296} - z_{gc,chitin}\right)^2} \quad (1)$$

$$z_{gc} = \frac{\sum_i^N z_i}{N} \quad (2)$$

The number of contacts between all beads in each amino acid and all beads in the chitin surface, c_i, were calculated using the COORDINATION routine in PLUMED 2.5 (*i.e.*, Equation 3), with a cutoff distance r_0 = 0.3 nm, r_i is the distance between all beads in each amino acid in the CBMs and all beads in the chitin surface.

$$c_i = \frac{1-(r_i/r_0)^6}{1-(r_i/r_0)^{12}} \quad (3)$$

Weights for each frame, w_i, were calculated from the bias in the WT-MetaD simulation using the REWEIGHT_BIAS routine in PLUMED 2.5 (see Ref. (88) and https://www.plumed.org/doc-v2.5/user-doc/html/_r_e_w_e_i_g_h_t__b_i_a_s.html for details). The weighted average number of contacts, $\langle c \rangle_w$, was calculated using Equation 4, and errors in $\langle c \rangle_w$ were estimated using block analysis (89).

$$\langle c \rangle_w = \sum w_i \cdot c_i \quad (4)$$

Dissociation constants from umbrella-sampling simulations

A binding path for each CBM that included bound and unbound conformations was selected from the WT-MetaD simulations. Umbrella-sampling simulations were performed by running twenty-three 100 ns-long replicas using the PLUMED 2.5 plugin, where the distance between the putative binding surface and the chitin surface (r) was restrained from 0.0 to 4.4 nm in steps of 0.2 nm using a harmonic restraint with a force constant of 100 kJ mol^{-1} nm^{-1}. Free-energy surfaces were calculated using the WHAM (http://membrane.urmc.rochester.edu/?page_id=126) (83), where the errors were estimated by Monte–Carlo resampling. Dissociation constants (K_d) were calculated from the free-energy surfaces (Fig. S3) by using Equations 5–7, where r_0 = 0.0 nm (*i.e.*, the minimum distance that for which WHAM calculated a nonzero probability), r_c = 1.85 nm, k_B is Boltzmann's constant, T is the temperature in Kelvin, P_0 is the protein concentration in the simulations, N_A is Avogadro's number, V_{box} is the volume of the simulation box, and $\dot{\alpha} = 4.38$ nm is the upper limit of r_{chitin}.

$$K_d = e^{\frac{-\Delta G_{binding}}{k_B T}} \quad (5)$$

$$\Delta G_{binding} = k_B T ln \left[\frac{\int_{r_c}^{\dot{\alpha}} exp\left(\frac{-PMF(r_{chitin})}{k_B T}\right) dr_{chitin}}{\int_{r_0}^{r_c} exp\left(\frac{-PMF(r_{chitin})}{k_B T}\right) dr_{chitin}} \right] + k_B T \; ln \; P_0 \quad (6)$$

$$P_0 = \frac{(1/N_A)}{V_{box}} = \frac{(1/6.02 \cdot 10^{23} mol)}{16 nm^3} \approx 0.001 mol/l \quad (7)$$

Data availability

The data for NMR structures and their restraints have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (https://www.rcsb.org/) under the following PDB ID codes: 6Z40 (*Cj*CBM5) and 6Z41 (*Cj*CBM73). NMR chemical shift assignments have been deposited in the BioMagnetic Resonance Databank (BMRB) under the IDs 34519 (*Cj*CBM5) and 34520 (*Cj*CBM73). All the data and PLUMED input files required to reproduce the simulation results reported in this article are available online at https://github.com/gcourtade/papers/tree/master/2021/CBM5-CBM73-MetaD-US and on PLUMED-NEST (www.plumed-nest.org), the public repository of the PLUMED consortium (85) as plumID: 21.015.

All other data are available in the main text and SI Appendix.

Supporting information—This article contains supporting information (60, 83, 88, 90, 91).

Acknowledgments—This work was funded by the Research Council of Norway through projects 226244, 247001, 262853, and 269408.

Author contributions—Z. F., Y. W., K. L.-L., V. G. H. E., F. L. A., and G. C. conceptualization; E. M., Z. F., Y. W., K. L.-L., A. N., J. M., V. G. H. E., F. L. A., and G. C. methodology; Y. W. and K. L.-L. software; E. M., Z. F., Y. W., K. L.-L., V. G. H. E., F. L. A., and G. C. validation; E. M., Z. F., Y. W., K. L.-L., V. G. H. E., F. L. A., and G. C. formal analysis; E. M., Z. F., Y. W., A. N., J. M., V. G. H. E., and G. C. investigation; Y. W., K. L.-L., V. G. H. E., ad F. L. A. resources; E. M., Z. F., Y. W., and G. C. data curation; E. M., Z. F., Y. W., A. N., J. M., V. G. H. E., F. L. A., and G. C. writing–original draft; E. M., Z. F., Y. W., K. L.-L., V. G. H. E., F. L. A., and G. C. writing–review and editing; E. M., Z. F., and G. C. visualization; Y. W., K. L.-L., V. G. H. E., F. L. A., and G. C. supervision; V. G. H. E., F. L. A., and G. C. project administration; Z. F., V. G. H. E., F. L. A., and G. C. funding acquisition.

Funding and additional information—This work was funded by The Novo Nordisk Foundation, project numbers NNF18OC0055736 (to Z. F.) and NNF18OC0032242 (to G. C.). Y. W. and K. L.-L. were supported by the BRAINSTRUC structural biology initiative from the Lundbeck Foundation.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AA10, auxiliary activity family 10; BMRB, BioMagnetic Resonance Databank; CAZyme, Carbohydrate Active enZyme; CBM, carbohydrate-binding module; CD, catalytic domain; CG, coarse-grained; *Cj*CBM5, internal family 5 CBM; *Cj*CBM73, C-terminal family 73 CBM; CV, column volume; D_2O, heavy water; FL, full length; HSQC, heteronuclear single quantum coherence; LB, lysogeny broth; LPMO, lytic polysaccharide monooxygenase; MWCO, molecular weight cutoff; PDB, Protein Data Bank; WHAM, weighted histogram analysis method; WT-MetaD, well-tempered metadynamics.

References

1. Gooday, G. W. (1990) The ecology of chitin degradation. In *Advances in Microbial Ecology*, Springer, Boston, MA: 387–430
2. Igarashi, K., Uchihashi, T., Uchiyama, T., Sugimoto, H., Wada, M., Suzuki, K., Sakuda, S., Ando, T., Watanabe, T., and Samejima, M. (2014) Two-way traffic of glycoside hydrolase family 18 processive chitinases on crystalline chitin. *Nat. Commun.* **5**, 3975
3. Hult, E., Katouno, F., Uchiyama, T., Watanabe, T., and Sugiyama, J. (2005) Molecular directionality in crystalline β-chitin: Hydrolysis by chitinases A and B from *Serratia marcescens* 2170. *Biochem. J.* **388**, 851–856
4. Vaaje-Kolstad, G., Horn, S. J., Sørlie, M., and Eijsink, V. G. H. (2013) The chitinolytic machinery of *Serratia marcescens*-a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS J.* **280**, 3028–3049
5. Nakagawa, Y. S., Eijsink, V. G. H., Totani, K., and Vaaje-Kolstad, G. (2013) Conversion of α-chitin substrates with varying particle size and crystallinity reveals substrate preferences of the chitinases and lytic polysaccharide monooxygenase of *Serratia marcescens*. *J. Agric. Food Chem.* **61**, 11061–11066
6. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V. G. H. H. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219–222
7. Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M. F., Synstad, B., and Eijsink, V. G. H. (2005) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J. Biol. Chem.* **280**, 28492–28497
8. Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., Krogh, K. B. R. M., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., *et al.* (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15079–15084
9. Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S. J., and Eijsink, V. G. H. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20**, 1479–1483
10. Eibinger, M., Ganner, T., Bubner, P., Rosker, S., Kracher, D., Haltrich, D., Ludwig, R., Plank, H., and Nidetzky, B. (2014) Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J. Biol. Chem.* **289**, 35929–35938
11. Phillips, C. M., Beeson, W. T., Cate, J. H., and Marletta, M. A. (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406
12. Agger, J. W., Isaksen, T., Várnai, A., Vidal-Melgosa, S., Willats, W. G. T., Ludwig, R., Horn, S. J., Eijsink, V. G. H., 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. U. S. A.* **111**, 6287–6292
13. Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R., and Marletta, M. A. (2014) A family of starch-active polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13822–13827
14. Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., Ludwig, R., Haltrich, D., Eijsink, V. G. H., and Horn, S. J. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J. Biol. Chem.* **289**, 2632–2642
15. Frandsen, K. E. H., Simmons, T. J., Dupree, P., Poulsen, J.-C. C. N., Hemsworth, G. R., Ciano, L., Johnston, E. M., Tovborg, M., Johansen, K. S., von Freiesleben, P., Marmuse, L., Fort, S., Cottaz, S., Driguez, H., Henrissat, B., *et al.* (2016) The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298–303
16. Vermaas, J. V., Crowley, M. F., Beckham, G. T., and Payne, C. M. (2015) Effects of lytic polysaccharide monooxygenase oxidation on cellulose structure and binding of oxidized cellulose oligomers to cellulases. *J. Phys. Chem. B* **119**, 6129–6143
17. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, 490–495
18. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781
19. Gilbert, H. J., Knox, J. P., and Boraston, A. B. (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* **23**, 669–677
20. Bissaro, B., Røhr, Å. K., Müller, G., Chylenski, P., Skaugen, M., Forsberg, Z., Horn, S. J., Vaaje-Kolstad, G., and Eijsink, V. G. H. (2017) Oxidative cleavage of polysaccharides by monocopper enzymes depends on H_2O_2. *Nat. Chem. Biol.* **13**, 1123–1128
21. Forsberg, Z., Bissaro, B., Gullesen, J., Dalhus, B., Vaaje-Kolstad, G., and Eijsink, V. G. H. (2018) Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J. Biol. Chem.* **293**, 1397–1412

22. Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H., and Aachmann, F. L. (2018) The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J. Biol. Chem.* **293**, 13006–13015
23. Crouch, L. I., Labourel, A., Walton, P. H., Davies, G. J., and Gilbert, H. J. (2016) The contribution of non-catalytic carbohydrate binding modules to the activity lytic polysaccharide monooxygenases. *J. Biol. Chem.* **291**, 7439–7449
24. Chalak, A., Villares, A., Moreau, C., Haon, M., Grisel, S., Orlando, A., Gimbert, I. H., Labourel, A., Cathala, B., and Berrin, J. G. (2019) Influence of the carbohydrate-binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol. Biofuels* **12**, 206
25. Forsberg, Z., Nelson, C. E., Dalhus, B., Mekasha, S., Loose, J. S. M., Crouch, L. I., Røhr, Å. K., Gardner, J. G., Eijsink, V. G. H., and Vaaje-Kolstad, G. (2016) Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J. Biol. Chem.* **291**, 7300–7312
26. Lehtiö, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 484–489
27. Bayer, E. A., Chanzy, H., Lamed, R., and Shoham, Y. (1998) Cellulose, cellulases and cellulosomes. *Curr. Opin. Struct. Biol.* **8**, 548–557
28. Brunecky, R., Alahuhta, M., Xu, Q., Donohoe, B. S., Crowley, M. F., Kataeva, I. A., Yang, S.-J., Resch, M. G., Adams, M. W. W., Lunin, V. V., Himmel, M. E., and Bomble, Y. J. (2013) Revealing nature's cellulase diversity: The digestion mechanism of *Caldicellulosiruptor bescii* CelA. *Science* **342**, 1513–1516
29. Bayer, E. A., Shoham, Y., and Lamed, R. (2006) Cellulose-decomposing bacteria and their enzyme systems. *Prokaryotes* **2**, 578–617
30. Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J., and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135**, 6069–6077
31. Tuveng, T. R., Jensen, M. S., Fredriksen, L., Vaaje-Kolstad, G., Eijsink, V. G. H., and Forsberg, Z. (2020) A thermostable bacterial lytic polysaccharide monooxygenase with high operational stability in a wide temperature range. *Biotechnol. Biofuels* **13**, 194
32. Kracher, D., Andlar, M., Furtmüller, P. G., and Ludwig, R. (2018) Active-site copper reduction promotes substrate binding of fungal lytic polysaccharide monooxygenase and reduces stability. *J. Biol. Chem.* **293**, 1676–1687
33. Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) Solution structure of the cellulose-binding domain of the endoglucanase Z secreted by *Erwinia chrysanthemi*. *Biochemistry* **36**, 16074–16086
34. Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L., and Woolfson, D. N. (2015) Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160
35. Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S. J. (2008) The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **4**, 819–834
36. Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H. K., Eijsink, V. G. H., and van Aalten, D. M. F. (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J. Biol. Chem.* **280**, 11313–11319
37. Vaaje-Kolstad, G., Forsberg, Z., Loose, J. S. M., Bissaro, B., and Eijsink, V. G. H. (2017) Structural diversity of lytic polysaccharide monooxygenases. *Curr. Opin. Struct. Biol.* **44**, 67–76
38. Mutahir, Z., Mekasha, S., Loose, J. S. M., Abbas, F., Vaaje-Kolstad, G., Eijsink, V. G. H., and Forsberg, Z. (2018) Characterization and synergistic action of a tetra-modular lytic polysaccharide monooxygenase from *Bacillus cereus*. *FEBS Lett.* **592**, 2562–2571
39. Ikegami, T., Okada, T., Hashimoto, M., Seino, S., Watanabe, T., and Shirakawa, M. (2000) Solution structure of the chitin-binding domain of *Bacillus circulans* WL-12 chitinase A1. *J. Biol. Chem.* **275**, 13654–13661
40. Akagi, K. I., Watanabe, J., Hara, M., Kezuka, Y., Chikaishi, E., Yamaguchi, T., Akutsu, H., Nonaka, T., Watanabe, T., and Ikegami, T. (2006) Identification of the substrate interaction region of the chitin-binding domain of *Streptomyces griseus* chitinase C. *J. Biochem.* **139**, 483–493
41. Mine, S., Nakamura, T., Sato, T., Ikegami, T., and Uegaki, K. (2014) Solution structure of the chitin-binding domain 1 (ChBD1) of a hyperthermophilic chitinase from *Pyrococcus furiosus*. *J. Biochem.* **155**, 115–122
42. Simpson, H. D., and Barras, F. (1999) Functional analysis of the carbohydrate-binding domains of *Erwinia chrysanthemi* Cel5 (endoglucanase Z) and an *Escherichia coli* putative chitinase. *J. Bacteriol.* **181**, 4611–4616
43. van Aalten, D. M., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H., and Wierenga, R. K. (2000) Structure of a two-domain chitotriosidase from *Serratia marcescens* at 1.9-A resolution. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5842–5847
44. Malecki, P. H., Raczynska, J. E., Vorgias, C. E., and Rypniewski, W. (2013) Structure of a complete four-domain chitinase from *Moritella marina*, a marine psychrophilic bacterium. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 821–829
45. Hara, M., Sugimoto, H., Uemura, M., Akagi, K., Suzuki, K., Ikegami, T., and Watanabe, T. (2013) Involvement of Gln679, in addition to Trp687, in chitin-binding activity of the chitin-binding domain of chitinase A1 from *Bacillus circulans* WL-12. *J. Biochem.* **154**, 185–193
46. Itoh, Y., Kawase, T., Nikaidou, N., Fukada, H., Mitsutomi, M., and Itoh, Y. (2002) Functional analysis of the chitin-binding domain of a family 19 chitinase from *Streptomyces griseus* HUT6037: Substrate-binding affinity and *cis*-dominant increase of antifungal function. *Biosci. Biotechnol. Biochem.* **66**, 1084–1092
47. Mattinen, M. L., Kontteli, M., Kerovuo, J., Linder, M., Annila, A., Lindeberg, G., Reinikainen, T., and Drakenberg, T. (1997) Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein Sci.* **6**, 294–303
48. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) Crystal structure of a bacterial family-III cellulose-binding domain: A general mechanism for attachment to cellulose. *EMBO J.* **15**, 5739–5751
49. Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. *Biochemistry* **34**, 6993–7009
50. Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J., and Orosz, L. (1998) All three surface tryptophans in type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. *FEBS Lett.* **429**, 312–316
51. Carrard, G., Koivula, A., Söderlund, H., and Béguin, P. (2000) Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10342–10347
52. Kikkawa, Y., Tokuhisa, H., Shingai, H., Hiraishi, T., Houjou, H., Kanesato, M., Imanaka, T., and Tanaka, T. (2008) Interaction force of chitin-binding domains onto chitin surface. *Biomacromolecules* **9**, 2126–2131
53. Mekasha, S., Tuveng, T. R., Askarian, F., Choudhary, S., Schmidt-Dannert, C., Niebisch, A., Modregger, J., Vaaje-Kolstad, G., and Eijsink, V. G. H. (2020) A tri-modular bacterial enzyme combining hydrolytic activity with oxidative glycosidic bond cleavage efficiently degrades chitin. *J. Biol. Chem.* **295**, 9134–9146
54. Manoil, C., and Beckwith, J. (1986) A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403–1408
55. 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
56. Stepnov, A. A., Forsberg, Z., Sørlie, M., Nguyen, G. S., Wentzel, A., Røhr, Å. K., and Eijsink, V. G. H. (2021) Unraveling the roles of the reductant and free copper ions in LPMO kinetics. *Biotechnol. Biofuels* **14**, 28
57. Brurberg, M. B., Eijsink, V. G. H., and Nes, I. F. (1994) Characterization of a chitinase gene (chiA) from *Serratia marcescens* BJL200 and one-step purification of the gene product. *FEMS Microbiol. Lett.* **124**, 399–404

58. Loose, J. S. M., Forsberg, Z., Fraaije, M. W., Eijsink, V. G. H., and Vaaje-Kolstad, G. (2014) A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett.* **588**, 3435–3440
59. Heuts, D. P. H. M., Winter, R. T., Damsma, G. E., Janssen, D. B., and Fraaije, M. W. (2008) The role of double covalent flavin binding in chito-oligosaccharide oxidase from *Fusarium graminearum*. *Biochem. J.* **413**, 175–183
60. Huynh, K., and Partch, C. L. (2015) Analysis of protein stability and ligand interactions by thermal shift assay. *Curr. Protoc. Protein Sci.* **79**, 28.9.1–28.9.14
61. Zhang, H., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* **25**, 173–195
62. Lescop, E., Schanda, P., and Brutscher, B. (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–169
63. Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial*, CANTINA Verlag, Goldau, Switzerland
64. Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241
65. Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378
66. Krieger, E., Koraimann, G., and Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* **47**, 393–402
67. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593
68. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77**, 114–122
69. Shindyalov, I. N., and Bourne, P. E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11**, 739–747
70. Mulder, F. A. A., Schipper, D., Bott, R., and Boelens, R. (1999) Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilisins*. *J. Mol. Biol.* **292**, 111–123
71. Sikorski, P., Hori, R., and Wada, M. (2009) Revisit of α-chitin crystal structure using high resolution X-ray diffraction data. *Biomacromolecules* **10**, 1100–1105
72. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
73. Yu, Z., and Lau, D. (2015) Development of a coarse-grained α-chitin model on the basis of MARTINI forcefield. *J. Mol. Model.* **21**, 128
74. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25
75. de Jong, D. H., Singh, G., Bennett, W. F. D., Arnarez, C., Wassenaar, T. A., Schäfer, L. V., Periole, X., Tieleman, D. P., and Marrink, S. J. (2013) Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–697
76. Periole, X., Cavalli, M., Marrink, S.-J., and Ceruso, M. A. (2009) Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **5**, 2531–2543
77. DeLano, W. L., and Lam, J. (2005) PyMOL: A communications tool for computational models. *Abstr. Pap. Am. Chem. Soc.* **230**, 1371–1372
78. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101
79. Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190
80. Páll, S., and Hess, B. (2013) A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650
81. [preprint] Vickery, O. N., and Stansfeld, P. J. (2021) CG2AT2: An enhanced fragment-based approach for serial multi-scale molecular dynamics simulations. *bioRxiv*. https://doi.org/10.1101/2021.03.25.437005
82. Larsen, A. H., Wang, Y., Bottaro, S., Grudinin, S., Arleth, L., and Lindorff-Larsen, K. (2020) Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution. *PLoS Comput. Biol.* **16**, e1007870
83. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021
84. Barducci, A., Bussi, G., and Parrinello, M. (2008) Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603
85. The PLUMED Consortium (2019) Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670–673
86. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613
87. Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A., and Parrinello, M. (2009) PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **180**, 1961–1972
88. Tiwary, P., and Parrinello, M. (2015) A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **119**, 736–742
89. Flyvbjerg, H., and Petersen, H. G. (1989) Error estimates on averages of correlated data. *J. Chem. Phys.* **91**, 461–466
90. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 41
91. Sievers, F., Wilm, F., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539