Functional diversity of three tandem C-terminal carbohydrate-binding modules of a β-mannanase

Marie Sofie Møller1,*°, Souad El Bouaballati1, Bernard Henriicat1,2,°, and Birte Svensson1,°

From the 1Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark; 2Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille Université, Marseille, France; and 3Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

Received for publication, September 28, 2020, and in revised form, March 29, 2021 Published, Papers in Press, April 7, 2021, https://doi.org/10.1016/j.jbc.2021.100638

Abstract

Carbohydrate active enzymes, such as those involved in plant cell wall and storage polysaccharide biosynthesis and deconstruction, often contain repeating noncatalytic carbohydrate-binding modules (CBMs) to compensate for low-affinity binding typical of protein–carbohydrate interactions. The bacterium Saccharophagus degradans produces an endo-β-mannanase of glycoside hydrolase family 5 subfamily 8 with three phylogenetically distinct family 10 CBMs located C-terminally from the catalytic domain (SdGH5_8-CBM10x3). However, the functional roles and cooperativity of these CBM domains in polysaccharide binding are not clear. To learn more, we studied the full-length enzyme, three stepwise CBM family 10 (CBM10) truncations, and GFP fusions of the individual CBM10s and all three domains together by pull-down assays, affinity gel electrophoresis, and activity assays. Only the C-terminal CBM10-3 was found to bind strongly to microcrystalline cellulose (dissociation constant, \(K_d = 1.48 \mu M\)). CBM10-3 and CBM10-2 bound galactomannan with similar affinity (\(K_d = 0.2–0.4 \text{mg/ml}\)), but CBM10-1 had 20-fold lower affinity for this substrate. CBM10 truncations barely affected specific activity on carob galactomannan and konjac glucomannan. Full-length Sd/GH5_8-CBM10x3 was twofold more active on the highly galactose-decorated viscous guar gum galactomannan and crystalline ivory nut mannan at high enzyme concentrations, but the specific activity was fourfold to ninefold reduced at low enzyme and substrate concentrations compared with the enzyme lacking CBM10-2 and CBM10-3. Comparison of activity and binding data for the different enzyme forms indicates unproductive and productive polysaccharide binding to occur. We conclude that the C-terminal most CBM10-3 secures firm binding, with contribution from CBM10-2, which with CBM10-1 also provides spatial flexibility.

Interactions between proteins and carbohydrates play a vital role in life and have through evolution been optimized to match the environments where they take place. Affinities of protein–carbohydrate complexes range from millimolar to nanomolar. Low-affinity binding (millimolar range) is considered a key factor in dynamic systems, such as plant cell wall synthesis and degradation, microbe–host interplay, and synthesis and mobilization of storage polysaccharides. The carbohydrate active enzymes catalyzing these different reactions very often contain one or more noncatalytic carbohydrate-binding modules (CBMs) that facilitate the formation of enzyme–substrate complexes and can specifically bind with the substrate polysaccharides or with different polysaccharides located nearby, such as in the plant cell wall. It is common that polysaccharide hydrolases contain several CBMs, but it is difficult to understand the function of these individual CBMs because of their interplay with each other, substrate, other polysaccharides, and the catalytic domain (CD), and perhaps even more than one CD is present (1). Certain insights are established, for example, on enzymes in cellulosesomes with domains of the polyspecific CBM32 family interacting with galactose, lactose, polygalacturonic acid, and N-acetyllactosamine (2). There are many examples where a functional role has been depicted for specific domains, while being unknown for others in multidomain architectures (3). Clearly, there are relatively few three-dimensional structures available of the more complex multidomain enzymes because of a number of challenges. One being the recombinant production of multimodular proteins containing interdomain linker regions prone to limited proteolysis, another is the dynamic flexibility inherent to the multimodular domain architectures (2, 3).

Mannan polysaccharides (Fig. 1) are found in cell walls of most plants and as storage polysaccharides in seeds and tubers (4). The four prominent types of β-mannans include galactomannans having a β-1,4 mannoe backbone with different extents of α-1,6 galactose decoration, gluco-mannans with an essentially linear backbone of glucosyl and mannosyl units, which in the galactoglucomannans are substituted by α-1,6-galactose (5, 6); essentially linear β-mannans are found in plant stems and as highly packed storage polysaccharides (7, 8) (Fig. 1). Some mannans are acetylated to varying degree depending on the type and botanical source (9).

Mannans are depolymerized by endo-β-mannanases found in glycoside hydrolase (GH) families GH5, GH26, GH113, and GH134 (10), some of which also contain CBMs facilitating contact with the substrates. Several applications are reported of endo-β-mannanases, for example, for production of...
oligosaccharides and in saccharification of plant biomass and lignocellulose degradation, where CBMs are of potential importance (11–14).

CBMs are classified into 88 families based on amino acid sequence similarity (CAZy, database of Carbohydrate Active enZYmes: http://cazy.org) (10). CBM family 10 (CBM10), despite the very small size of 29 to 42 amino acids, presents a diversity of carbohydrate-binding specificities and affinity ranges (Table 1) (15–21). CBM10s are thus associated with different CDs in plant cell wall polysaccharide-degrading enzymes, including cellulases (22–28), xylanases (15, 19, 26, 29), mannanases (17, 18, 26, 30–35), and lytic polysaccharide monoxygenases (LPMOs) (20). Typically, these enzymes have more than one CBM10 or contain also CBMs from other families, CBM2s being particularly common (Table 1).

CBM10 is a type A surface-binding module with a characteristic “planar” face containing aromatic residues capable of binding to regularly organized surfaces, for example, of the crystalline polysaccharides, cellulose and chitin (36). In the case of CBM10, a typical binding motif has three regularly spaced aromatic side chains, tyrosine, tryptophan, and tryptophan, initially identified in a CBM10 from xylanase A from Cellvibrio japonicus (CjGH10CBM10) (16, 37) (Table 1). These residues were pivotal for binding to microcrystalline cellulose (Avicel). Furthermore, CjGH10CBM10 bound to bacterial microcrystalline cellulose (BMCC) and phosphoric acid–swollen cellulose (PASC) with a dissociation constant (K_d) of 0.2 to 0.3 μM but not to hydroxyethyl cellulose (HEC) (15). It moreover bound differentially to plant cell walls depending upon cell type, tissue, and plant taxon (38). CjGH10CBM10 is the only structure determined CBM10 (Protein Data Bank ID: 1E8R) (37). Another cellulose-specific C. japonicus CBM10 (CjAA10CBM10) from LPMO (of auxiliary activities family 10; AA10A) showed a similar polysaccharide-binding preference and bound with K_d = 7.5 to 17.3 μM to disordered (PASC) as well as to crystalline (Avicel and BMCC) cellulose (20), which were also substrates of the LPMO (24). Finally, CjGH10CBM10 (CjGH10CBM10) of the C. japonicus GH74 endo-xylanoglucanase that also contains a CBM2, when analyzed as a GFP fusion did not bind to soluble tamarind seed xylanoglucan and barley 1,3,1,4-β-glucan, both of which are substrates of the GH74 endo-xylanoglucanase, or to HEC (Table 1), but bound with high affinity (K_d = 1.5 μM) to Avicel, which is not a substrate (19).

The single CBM10 from a mannanase of GH family 5 subfamily 8 (GH5_8) from the probiotic bacterium Biofloc-bacterium animalis subsp. lactis Bl-04 (BaGH5CBM10) was able to bind to low-viscosity carob galactomannan (CGM-lv) with a K_d of 0.31 mg/ml but not to microcrystalline cellulose (Avicel), soluble cellulose (HEC), or insoluble ivory nut mannan (INM) (18) (see Fig. 1 for polysaccharide structures) (Table 1). Notably, even though the single BaGH5CBM10 was critical for binding onto galactomannan, the enzyme kinetic parameters for this substrate did not change when the CBM10 was removed by truncation, whereas the truncated enzyme was not able to bind galactomannan (18). The lack of binding onto the crystalline cellulose and INM was speculated to be due to substitution of an otherwise conserved tyrosine residue from the typical type A planar carbohydrate-binding motif (16) and also by insertion of five residues, which expand a loop. These characteristics were seen for a phylogenetic subgroup distinct from the classical CBM10s (18).

A CBM10 from Clostridium cellulolyticum cellulase of GH9 adsorbed to PASC, Avicel, alkaline-pretreated sugarcane bagasse, and filter paper powder (39). In addition, by comparing 40 type A CBMs (of CBM1, CBM2, CBM3, CBM10, CBM63, CBM64, CBM78, and CBM79) fused with GFP, three CBM10s in two multimodular enzymes from Teredinibacter turnerae (GH5_2-CBM5-CBM10-GH6_4) and the marine bacterium Saccharophagus degradans (CBM2-CBM10-GH5_4) as well as the CjGH10CBM10 mentioned previously showed low affinity for Sigma cellulose type 20 (20 μm particles) and chitosan compared with the other type A CBMs (21). Finally, CBM10 in a GH134 β-1,4-mannanase from Streptomyces sp. NRRL both inhibited enzyme activity and conferred stability at neutral-to-alkaline pH as well as in the presence of organic solvents and detergents (33).
<table>
<thead>
<tr>
<th>Organism</th>
<th>GenBank accession no.</th>
<th>Modular organization</th>
<th>CBM10 binding</th>
<th>Enzymatic activity</th>
<th>Binding-site residues</th>
<th>Disulphide bridges</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE85176</td>
<td>CBM10-GH26</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. degradans</em></td>
<td>ABD79911</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. bacterium</em></td>
<td>AIF91534</td>
<td>GH4_5-CBM10-GH4_10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE84673</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE83578</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE83870</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE83439</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE84076</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. bacterium</em></td>
<td>AIF91534</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. bacterium</em></td>
<td>AIF91559</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. bacterium</em></td>
<td>AIF91534</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. degradans</em></td>
<td>ABD79918</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>T. turnerae</em></td>
<td>ABD81F896</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>R. degradans</em></td>
<td>ABD79918</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. animalis</em></td>
<td>ACS46797</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>ACE84673</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>AEE43708</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. japonicus</em></td>
<td>AEE43708</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptomyces</em></td>
<td>WP_0026892</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptomyces</em></td>
<td>ADR91085</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>St. lividans</em></td>
<td>AAA28710</td>
<td>GH5_8-CBM10-GH10</td>
<td>ND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ND, not determined.

The order of the proteins follows the grouping seen in the phylogenetic tree (Fig. 3). When an enzyme has more than one CBM10, if possible, the information regarding carbohydrate binding is given for the particular CBM10 (underlined). Biochemical information is given using the following symbols for the polysaccharides: [ ] soluble galactomannan; [ ] insoluble mannan (INM or \( \beta \)-mannan); [ ] soluble glucomannan; [ ] soluble cellulose (HEC, CMC); [ ] insoluble cellulose (Avicel, BMCC, PASC, filter paper, bagasse); [ ] other soluble polysaccharides (including xyloglucans, xylan, \( \beta \)-glucans, and lichenan). [ ] indicates that the given polysaccharide has been shown not to bind or being a substrate.
Functional diversity of CBM10s

Here, we shall dissect the function of individual CBM10s in a GH5_8 endo-β-mannanase from the marine bacterium and carbohydrate super degrader S. degradans by producing the CBM10 domains individually and together as GFP fusions and the full-length enzyme (SdGH5_8-CBM10x3) as well as its three CBM10 truncations (Fig. 2). The three consecutive CBM10s situated C-terminally to the CD are here named SdGH5CBM10-1, SdGH5CBM10-2, and SdGH5CBM10-3 and shown to belong to three distinct phylogenetic subgroups (Fig. 3). Polysaccharide binding was monitored by affinity gel electrophoresis (AGE) or pull-down assays. Enzyme activity was determined using CGM-Iv and high-viscosity CGMs and glucomannan, a highly decorated very viscous guar gum (GG) galactomannan, insoluble crystalline INM, and microcrystalline cellulose (Avicel). The three SdGH5CBM10s showed distinctly different polysaccharide-binding profiles and impact on the enzyme activity and substrate specificity.

Results and discussion

Phylogenetic analysis of CBM10s

The phylogenetic analysis of CBM10 reveals a number of different subgroups (Fig. 3), which have characteristic features with regard to (i) the presence of three aromatic residues (tyrosine, tryptophan, and tryptophan) seen in a canonical type A CBM10 (16), (ii) the presence of cysteine residues potentially forming one of the two observed disulphide bonds, and (iii) the presence of a five to six residues insert between the tyrosine and the two tryptophans of the planar-binding motif (Figs. 4 and S1). For example, the BAgGH5CBM10 in the Bifidobacterium mannanase, which was needed for binding to galactomannan, but did not confer binding to cellulose or INM, had the following distinct features: (i) substitution of the first tyrosine (to a threonine) of the three aromatic residues critical for Avicel binding as demonstrated for C. japonicus CjGH10CBM10 (16), (ii) lack of two cysteines that form one of the two disulphide bridges, and (iii) a five-residue long insertion (Figs. 4 and S1).

To further investigate functional characteristics of CBM10s from different regions of the phylogenetic tree, we selected a putative GH5_8 mannanase from the marine bacterium and carbohydrate super degrader S. degradans having three C-terminal CBM10s (SdGH5_8-CBM10x3; GenBank accession no.: ABD79918) (Fig. 2). These CBM10s were hypothesized to display distinct functions as they belong to three different phylogenetic subgroups (Fig. 3).

Notably, the innermost SdGH5CBM10-1 linked to the CD of the S. degradans mannanase (Fig. 2) was found in the phylogenetic tree at a distance of characterized CBM10s. The middle domain, SdGH5CBM10-2, however belonged to the same subgroup as BAgGH5CBM10 from the Bifidobacterium mannanase, although they were not closely related. The C-terminal SdGH5CBM10-3 belongs to the same subgroup as the previously well-characterized C. japonicus CBM10s from xylanase A of GH10 (CjGH10CBM10) and of the GH74 endoglucanase active on xylglucan and mixed-linked 1,3;1,4-β-glucan (CjGH74CBM10) (19) (Fig. 3).

Effect of CBM10s on specificity and activity of GH5_8 endo-β-mannanase

The full-length enzyme, SdGH5_8-CBM10x3, and its three stepwisely truncated forms lacking one, two, or all three CBM10s were recombinantly produced in Escherichia coli (Fig. 2). For production of the individual CBM10s and all three CBM10s together, the GFP fusion strategy reported previously to enable production of CBM10 from the GH74 endoglucanase of C. japonicus (CjGH74CBM10) (19) was adopted (Fig. 2). All recombinant enzyme forms and the CBM10–GFP fusions were obtained in reasonably good yields of 0.7 to 7.5 mg/g cells after purification. Notably, some of the multimodular forms, in particular the full-length SdGH5_8-CBM10x3, undergo significant proteolytic cleavage in the 21- to 30-residue long serine–glycine rich linker regions during production and purification, which is reflected in the lower yields.

The full-length SdGH5_8-CBM10x3 and its three truncated forms were active on all tested mannan polysaccharides but not on microcrystalline cellulose (Avicel) (Table 2). The substrate specificity of SdGH5_8-CBM10x3 is similar to that of other characterized GH5_8 mannanases but with specific variances.
activities at the higher end (17, 18). Removal of the C-terminal CBM10 (SdGH5CBM10-3) did not essentially affect the specific activity toward the good substrates, that is, CGM-lv and high-viscosity CGM (locust bean gum) and konjac glucomannan (KGM) (Table 2; Fig. 1). Remarkably, however, removal of the C-terminal SdGH5CBM10-3 resulted in doubling, respectively, almost quadrupling of the specific activity (Table 2) on the poor substrates; highly viscous GG galactomannan with on average every second mannosyl backbone unit substituted by galactose (as opposed to one out of four in CGM) and the insoluble crystalline INM (Fig. 1). Removal of both the middle (SdGH5CBM10-2) and the C-terminal (SdGH5CBM10-3) domains reduced activity by 36% to 40% for CGM-lv and high-viscosity CGMs and maintained activity for KGM compared with the full-length enzyme (Table 2). Remarkably, the activity of SdGH5_8-CBM10x1 with only the innermost CBM10, increased fourfold and ninefold on the poor substrates GG and INM, respectively, compared with the full-length enzyme (Table 2). Finally, removing all three CBM10s essentially recovered activity of the full-length enzyme on CGMs and KGM. This finding resembles the lack of effect on activity for CGM-lv by removing the single C-terminal CBM10 from the Bifidobacterium GH5 mannanase (18). Again on the poor substrates GG and INM, the CD alone (SdGH5_8) showed twofold and sixfold higher activity than full-length enzyme. Overall, this suggests that the presence of only SdGH5CBM10-1 slightly hinders CGMs from interacting fully effectively with the CD. The presence of the C-terminal and middle CBM10 seems to prevent optimal productive interaction between the CD and the poor substrates INM and the highly substituted guar GG as reflected in the CD alone (SdGH5_8) and the enzyme form with only the first CBM10 (SdGH5_8-CBM10x1) having higher specific activities at low enzyme concentrations.

The kinetic parameters for full-length SdGH5_8-CBM10x3 and truncated forms toward CGM-lv are in agreement with the findings on specific activity (Table 3). Thus, $K_m$ of SdGH5_8-CBM10x1 was threefold higher than of the full-length enzyme, whereas the turnover number increased by 1.8-fold, resulting in 36% lower catalytic efficiency ($k_{cat}/K_m$). Notably, the catalytic efficiency of CD alone is doubled and 1.4-fold higher than of the SdGH5_8-CBM10x1 and full-length enzyme, respectively. The $K_m$ of CD alone is low and similar to that of full-length SdGH5_8-CBM10x3, whereas the turnover number is 68% higher. The lowest turnover number is seen for full-length enzyme, whereas SdGH5_8-CBM10x1 has the highest turnover number (Table 3).
A negative effect on activity of the presence of a CBM10 was previously observed for the GH134 \( \beta \)-1,4-mannanase that contains a single CBM10 and for which the catalytic efficiency of the CD alone on galactomannan was 2.6-fold higher than of the full-length enzyme, whereas \( K_m \) was 1.4-fold lower (33), suggesting that the CBM10 is associated with nonproductive binding. This CBM10 did not confer the binding to microcrystalline cellulose, \( \beta \)-mannan, chitin, or xylan as the CD alone can bind to these polysaccharides (33). Although it is not generally seen, there are examples from other CBM families where CBMs have an inhibitory effect on enzyme activity. Thus, removal of a cellulose-binding CBM3 from an endo-\( \beta \)-1,4-glucanase from \textit{Bacillus subtilis} resulted in both increased \( k_{cat} \) and catalytic efficiency and also improved thermal stability.

### Table 2

<table>
<thead>
<tr>
<th>Substrates</th>
<th>Units</th>
<th>SdGH5_8-CBM10x3</th>
<th>SdGH5_8-CBM10x2</th>
<th>SdGH5_8-CBM10x1</th>
<th>SdGH5_8 (CD alone)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM-lv</td>
<td>U/mg</td>
<td>1972 ± 80</td>
<td>2285 ± 150</td>
<td>1539 ± 88</td>
<td>2906 ± 53</td>
</tr>
<tr>
<td></td>
<td>(( \mu )mol/s)/( \mu )mol protein(^a)</td>
<td>1729 ± 70</td>
<td>1763 ± 116</td>
<td>1046 ± 60</td>
<td>1690 ± 31</td>
</tr>
<tr>
<td></td>
<td>Relative (%)(^a)</td>
<td>100</td>
<td>102</td>
<td>60</td>
<td>98</td>
</tr>
<tr>
<td>CGM (high viscosity)</td>
<td>U/mg</td>
<td>2212 ± 78</td>
<td>2755 ± 109</td>
<td>1814 ± 292</td>
<td>3151 ± 304</td>
</tr>
<tr>
<td></td>
<td>(( \mu )mol/s)/( \mu )mol protein(^a)</td>
<td>1939 ± 69</td>
<td>2126 ± 84</td>
<td>1234 ± 198</td>
<td>1833 ± 177</td>
</tr>
<tr>
<td></td>
<td>Relative (%)(^a)</td>
<td>100</td>
<td>110</td>
<td>64</td>
<td>95</td>
</tr>
<tr>
<td>KGM</td>
<td>U/mg</td>
<td>3544 ± 110</td>
<td>3617 ± 112</td>
<td>5136 ± 128</td>
<td>4556 ± 108</td>
</tr>
<tr>
<td></td>
<td>(( \mu )mol/s)/( \mu )mol protein(^a)</td>
<td>3107 ± 97</td>
<td>2791 ± 87</td>
<td>3493 ± 87</td>
<td>2650 ± 63</td>
</tr>
<tr>
<td></td>
<td>Relative (%)(^a)</td>
<td>100</td>
<td>90</td>
<td>112</td>
<td>85</td>
</tr>
<tr>
<td>GG</td>
<td>U/mg</td>
<td>40 ± 8</td>
<td>24 ± 1</td>
<td>200 ± 33</td>
<td>108 ± 25</td>
</tr>
<tr>
<td></td>
<td>(( \mu )mol/s)/( \mu )mol protein(^a)</td>
<td>35 ± 7</td>
<td>19 ± 3</td>
<td>136 ± 23</td>
<td>63 ± 14</td>
</tr>
<tr>
<td></td>
<td>Relative (%)(^a)</td>
<td>100</td>
<td>54</td>
<td>389</td>
<td>180</td>
</tr>
<tr>
<td>INM</td>
<td>U/mg</td>
<td>9 ± 1</td>
<td>38 ± 3</td>
<td>106 ± 6</td>
<td>81 ± 7</td>
</tr>
<tr>
<td></td>
<td>(( \mu )mol/s)/( \mu )mol protein(^a)</td>
<td>8 ± 1</td>
<td>29 ± 2</td>
<td>72 ± 4</td>
<td>47 ± 4</td>
</tr>
<tr>
<td></td>
<td>Relative (%)(^a)</td>
<td>100</td>
<td>363</td>
<td>900</td>
<td>588</td>
</tr>
<tr>
<td>Avicel</td>
<td>U/mg</td>
<td>ND(^c)</td>
<td>ND(^c)</td>
<td>ND(^c)</td>
<td>ND(^c)</td>
</tr>
</tbody>
</table>

\(^a\) Specific activity calculated based on molecular weight.

\(^b\) Specific activity relative to full-length SdGH5_8-CBM10x3 calculated based on the specific activity given in (\( \mu \)mol/s)/\( \mu \)mol protein.

\(^c\) No activity detected.
Furthermore, when a CBM5 was removed from a *Pectobacterium chrysanthemi* endoglucanase, $K_m$ and $V_{\text{max}}$ were reduced by 4.5-fold and increased by 2.5-fold, respectively (41).

The kinetics analysis on the insoluble crystalline INM showed the C-terminal CBM10-3 of full-length SdGH5_8-CBM10x3 from *S. degradans* to be needed for maximum activity, whereas the three stepwisely CBM10 truncated enzyme forms showed similarly reduced activity relative to the full-length enzyme (Fig. 5A). In the case of GG, the full-length enzyme showed the highest activity, whereas the CD alone was the least active form (Fig. 5B). These results on the highly viscous and densely decorated guar GG indicated that the C-terminal CBM10-3 together with CBM10-1 play a role for activity at high-substrate concentrations. Thus, removal of CBM10-3 resulted in some reduction in activity, whereas removal of CBM10-2 as well did not make a difference, as the SdGH5_8-CBM10x1 and SdGH5_8-CBM10x2 showed the same activity. But the removal of CBM10-1 led to a further reduced activity. Although full substrate saturation was not possible, calculations applying the Michaelis–Menten equation on activity of the full-length SdGH5_8-CBM10x3 on INM indicated $K_m = 35 \pm 14 \text{ mg/ml}$ and $k_{\text{cat}} = 186 \pm 53 \text{ s}^{-1}$ ($k_{\text{cat}}/K_m = 5.3 \text{ ml/(mg s)}$). The $K_m$ is 1.3-fold higher as compared with a GH5_8 mannanase with one CBM10 from *Streptomyces lividans* ($K_m = 26.5 \pm 11.5 \text{ mg/ml}$), whereas the $k_{\text{cat}}$ of that enzyme was $69 \pm 1.2 \text{ s}^{-1}$ resulting in a catalytic efficiency of 2.6 ml/(mg s) (42); hence, SdGH5_8-CBM10x3 is twofold more efficient. Furthermore, the kinetics on INM of the *S. lividans* mannanase was not affected by removal of the CBM10 (42). This CBM10 is in a phylogenetic subgroup with SdGH5-CBM10-2; hence, the lack of effect on activity on ivory nut is in line with our findings (Fig. 5A).

To understand the reason behind the differences observed between the results of the enzyme kinetics on INM and guar GG (Fig. 5) and the specific activity analyses (Table 2), one should have the differences in assay setup in mind. The specific activities were determined as is typical for mannanases at readily manageable low substrate concentrations to avoid highly viscous solutions and maintained (low) enzyme concentration with end point measurement implying substrate-dependent incubation time (18) (Table 2). The kinetics analysis contrarily was done at up to fourfold higher substrate and 10-fold to 30-fold higher enzyme concentration. The action on the insoluble INM should probably be considered as heterogeneous catalysis as found in case of cellulases acting on crystalline cellulose, characterized by the presence of different categories of binding sites and productive enzyme attack sites (43, 44). A number of crystalline forms are reported for INM (7, 8), but details on actual mannanase degradation of these forms have not been reported. It is possible that most strongly binding, albeit not productively binding, substrate sites are interacting with full-length enzyme *via* the C-terminal CBM10-3, which deprives the CD access to susceptible substrate bonds. At the higher enzyme concentration used for the kinetics analysis (Fig. 5A), this is in part overcome. For the other poor substrate guar GG, the substrate is assumed to contain enzyme target bonds placed in different structural contexts and having different preferences for the productive enzyme attack to occur (6). Although a GH26 mannanase from

### Table 3

<table>
<thead>
<tr>
<th>Enzyme form</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_m$ (mg/ml)</th>
<th>$k_{\text{cat}}/K_m$ (ml [mg s]$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SdGH5_8-CBM10x3</td>
<td>2333 ± 55</td>
<td>2.1 ± 0.1</td>
<td>1096 ± 71</td>
</tr>
<tr>
<td>SdGH5_8-CBM10x2</td>
<td>3589 ± 278</td>
<td>4.2 ± 0.7</td>
<td>855 ± 149</td>
</tr>
<tr>
<td>SdGH5_8-CBM10x1</td>
<td>4253 ± 337</td>
<td>6.1 ± 0.9</td>
<td>701 ± 114</td>
</tr>
<tr>
<td>SdGH5_8</td>
<td>3440 ± 75</td>
<td>2.4 ± 0.1</td>
<td>1413 ± 82</td>
</tr>
</tbody>
</table>

---

**Figure 5.** Kinetics of the full-length enzyme SdGH5_8-CBM10x3. The three stepwise truncated forms on ivory nut mannan (INM) (A) and guar gum galactomannan (GG) (B). SdGH5_8-CBM10x3 (■), SdGH5_8-CBM10x2 (●), SdGH5_8-CBM10x1 (▲), and SdGH5_8 (◆). The Michaelis–Menten fit to the SdGH5_8-CBM10x3 data is shown as dashed line in A.
**Functional diversity of CBM10s**

*B. subtilis* is reported with preference for the guar GG over less substituted galactomannans (45), this is not common, and we assume that initial hydrolysis on GG is slow and may open up for more accessible substrate regions possibly influenced by the fine structural preferences of the CBM10-3 and the CD. This type of action may be observed as a degradation involving apparent cooperative events also indicated by the sigmoidal curvature of the kinetic data (Fig. 5B).

**Polysaccharide-binding analysis**

The affinities for soluble CGM-lv, HEC, insoluble micro-crystalline cellulose (Avicel), and INM were assessed by retardation in AGE and pull-down assays, respectively. Neither the different enzyme forms nor the CBM10s individually or all three CBM10s together were able to bind HEC in AGE. By contrast, both full-length enzyme *Sd*GH5_8-CBM10x3 and the three CBM10s together (*Sd*GH5CBM10x3) showed high affinity for CGM-lv (Table 4). Although each of the three *Sd*GH5CBM10s when tested individually were able to bind CGM-lv with reasonable affinity, the strong binding depended on the C-terminal domain (*Sd*GH5CBM10-3) but was also clearly a combined effect of all the three CBM10s. The full-length enzyme and the three CBM10s together thus bound very strongly with *Kd* of 0.13 mg/ml, whereas neither of the truncated enzyme forms lacking one, two, or all three CBM10 domains showed significant retardation in AGE. Both the C-terminal CBM10-3 and the middle CBM10-2 domains, however, bound reasonably well to galactomannan, the data suggesting that they together secures the full-length enzyme binding. This is opposed to the analysis of the innermost *Sd*GH5CBM10-1, which showed ninefold to 18-fold reduced affinity compared with the other two CBM10s having *Kd* of 0.4 and 0.2 mg/ml, respectively (Table 4). The second domain (*Sd*GH5CBM10-2_GFP), which has an insert of five amino acid residues and lacks a potentially stabilizing disulphide bridge similar to *Ba*GH5CBM10 from the *Bifidobacterium* mannanase (Figs. 4 and S1), had a *Kd* similar to that of *Ba*GH5CBM10 (*Kd* = 0.3 g 1⁻¹) (18) that it also resembles phylogenetically (Fig. 3). But surprisingly, *Sd*GH5CBM10-2_GFP had ninefold higher affinity than the *Sd*GH5CBM10-1_GFP (Table 4). A possible explanation is that *Sd*GH5CBM10-1, although it probably contains the two stabilizing disulphide bridges, has a serine replacing the first of the three aromatic binding residues and a tyrosine replacing the tryptophan at the position of the third binding site residue (Table 1; Fig. 4). It belongs to a phylogenetic region shared by few CBM10s. Notably, the LPMO *CjAAb*CBM10 also has the third tryptophan replaced (by asparagine) but still has good affinity for insoluble cellulose; however, it does have the conserved tyrosine as the first of the three binding residues (Table 1; Fig. 4).

A qualitative pull down by insoluble INM of the three individual *Sd*GH5CBM10 and the combined three domain GFP fusions showed that *Sd*GH5CBM10-3_GFP and *Sd*GH5CBM10x3_GFP interacted strongly with INM, whereas the amount of protein bound in the case of both *Sd*GH5CBM10-1_GFP and *Sd*GH5CBM10-2_GFP was in the same range as the amount of GFP alone (Fig. 6). This is in line with the kinetic assay showing that when the C-terminal CBM10-3 was removed, the activity was reduced (Fig. 5A). Moreover, the qualitative pull down by Avicel showed only weak interaction with *Sd*GH5CBM10-1_GFP and *Sd*GH5CBM10-2_GFP, whereas the three CBM10s together (*Sd*GH5CBM10x3_GFP) and the C-terminal domain *Sd*GH5CBM10-3_GFP bound to Avicel (Fig. 6). Quantitative binding isotherms gave *Kd* for *Sd*GH5CBM10x3_GFP and *Sd*GH5CBM10-3_GFP of 1.04 ± 0.16 and 1.48 ± 0.12 μM, respectively (Fig. 7). The maximum amount of bound *Sd*GH5CBM10x3_GFP and *Sd*GH5CBM10-3_GFP was 0.13 ± 0.00 and 0.27 ± 0.04 μmol/g Avicel, respectively. The binding of *Sd*GH5CBM10-1_GFP and *Sd*GH5CBM10-2_GFP (and GFP alone) was too weak to obtain a *Kd*. Hence, *Sd*GH5CBM10-3 is responsible for Avicel binding, in accordance with findings for other members of the same phylogenetic subgroup (Fig. 3), which had essentially the same affinity as reported also for *CjAAb*CBM10-GFP of GH74 endo-xylloglucanase (19). Notably, the first and second CBM10s in the *S. degradans* mannanase belong to different phylogenetic subgroups and lacked two and one, respectively, of the three residues in the planar aromatic binding motif (16) (Figs. 3 and 4). Previously, it was suggested that the absence of a potentially stabilizing disulphide bond explained a similar lack of Avicel binding for the *Bifidobacterium* full-length mannanase with one CBM10 (18), but this cannot explain why *Sd*GH5CBM10-1 is not binding to Avicel, as it has the four cysteines predicted to form the two stabilizing disulphides (Fig. 4). While *Sd*GH5CBM10-3 has a *Kd* for Avicel in the same range as the rest of the characterized CBM10s from the same subgroup in the phylogenetic tree (Fig. 3), it differs by showing ability to bind soluble galactomannan in distinction also from the two other CBM10s from this subgroup tested for galactomannan binding (17).

**Table 4**

<table>
<thead>
<tr>
<th>Protein form</th>
<th><em>Kd</em> (mg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sd</em>GH5_8-CBM10x3</td>
<td>&lt;0.125, a</td>
</tr>
<tr>
<td><em>Sd</em>GH5_8-CBM10x2</td>
<td>No binding</td>
</tr>
<tr>
<td><em>Sd</em>GH5_8-CBM10x1</td>
<td>No binding</td>
</tr>
<tr>
<td><em>Sd</em>GH5_8 (no CBMx4)</td>
<td>No binding</td>
</tr>
<tr>
<td><em>Sd</em>GH5CBM10x3_GFP</td>
<td>&lt;0.125, a</td>
</tr>
<tr>
<td><em>Sd</em>GH5CBM10-1_GFP</td>
<td>3.7</td>
</tr>
<tr>
<td><em>Sd</em>GH5CBM10-2_GFP</td>
<td>0.4</td>
</tr>
<tr>
<td><em>Sd</em>GH5CBM10-3_GFP</td>
<td>0.2</td>
</tr>
<tr>
<td>GFP</td>
<td>No binding</td>
</tr>
</tbody>
</table>

*Saturation was reached already at the lowest in-gel galactomannan concentration.*

**Functional diversity of CBM10s**

Based on the present results, the C-terminal CBM10, *Sd*GH5CBM10-3, is proposed to anchor the secreted enzyme to mannan and to crystalline cellulose, which seems valuable for a free-living marine bacterium as *S. degradans*. *S. degradans* is a so-called super degrader and can utilize complex polysaccharides of algal, higher plant, fungal, and animal origin as sole carbon and energy source (46). Avicel is neither a substrate for *Sd*GH5_8-CBM10x3 not for the *C. japonicus* GH74 endo-xylloglucanase although the CBM10s of both enzymes
bind well to Avicel (19). It is well known that CBMs can interact with polysaccharides like cellulose in the plant cell wall or other biomass, rather than soluble branched polysaccharides such as galactomannan (47). However, $Sd_{GH5}$CBM10-1 and $Sd_{GH5}$CBM10-2 seem to have evolved to become dynamic domains securing flexible searching for an enzyme attack site on a substrate surface or gel for the CD, while being firmly anchored via the terminal $Sd_{GH5}$CBM10-3. Notably, in particular, $Sd_{GH5}$CBM10-2 has reasonable affinity for galactomannan and may potentiate binding to the actual substrate polysaccharide molecule in a structurally complex plant cell wall environment. CBM10 family thus includes members with promiscuous functions binding to a diversity of polysaccharides, such as soluble galactomannan, insoluble crystalline mannans, and cellulose with good affinity (e.g. $Sd_{GH5}$CBM10-3), whereas some bind soluble polysaccharide substrates with good affinity (e.g. $Sd_{GH5}$CBM10-2) or with low affinity (e.g. $Sd_{GH5}$CBM10-1). Members of the same phylogenetic subgroup as $Sd_{GH5}$CBM10-3-like CBM10s probably anchor full-length enzymes firmly to substrate or spatially adjacent polysaccharides, whereas CBM10s from phylogenetically different subgroups may still show binding to certain polysaccharides and enable the CD to flexibly survey the vicinity for susceptible glycosidic bonds while yet others may be protein domains serving as linkers or spacers not engaged in direct binding onto a crystalline cellulose via a classical CBM type A platform. In nature, multimodal enzymes characteristically participate in heterogeneous catalysis on insoluble substrates to which firm binding is mandatory. However, binding is not the same as productive binding, and the density of functional and nonfunctional binding attack sites depends on the material and enzyme as

Figure 6. Qualitative pull-down assay with $Sd_{GH5}$CBM10–GFP fusions of each of the three domains $Sd_{GH5}$CBM10-1, $Sd_{GH5}$CBM10-2, and $Sd_{GH5}$CBM10-3 and all three domains combined $Sd_{GH5}$CBM10x3.GFP by the insoluble polysaccharides microcrystalline cellulose (Avicel) and ivory nut mannann (INM), respectively. SDS-PAGE of samples from supernatants (SN) and pellets (P) as well as the protein stocks used for the pull-down assay. GFP alone is included as a control (right panel). Molecular mass values (kDa) of marker proteins are indicated.

Figure 7. Quantitative pull-down assay with microcrystalline cellulose (Avicel). Binding isotherms of the C-terminal domain, $Sd_{GH5}$CBM10-3.GFP (A) and all three domains together, $Sd_{GH5}$CBM10x3.GFP (B). The filled and open symbols represent replicates from different days. No significant binding was observed of the first and the middle domains, $Sd_{GH5}$CBM10-1.GFP, $Sd_{GH5}$CBM10-2.GFP, or of GFP alone (data not shown).
illustrated in details for cellulose degradation by cellulases (43, 44).

**Conclusion and perspectives**

The multimodular endo-β-mannanase of family GH5.8 secreted by *S. degradans* is able to bind to polysaccharide substrates and substrate containing biological material by aid of three consecutive C-terminal CBM10 domains. Clearly, the most C-terminal CBM10-3 secures the attachment both to crystalline cellulose (Avicel), galactomannan, and also insoluble seed storage mannan. However, the middle domain CBM10-2 shows also affinity with a 10-fold lower *Kd* than the *Km* for the good substrate galactomannan and may facilitate its interactions with the CD although the truncated enzyme lacking the C-terminal CBM10 lost ability to bind to the galactomannan. The CBM10-2 therefore may also serve as a linker that provides flexibility in polysaccharide interaction for the C-terminal domain and hence for the CD to interact productively with substrate. The role of the innermost CBM10-1 is more enigmatic, but at low enzyme and substrate concentrations, it was able to secure highest activity of the CD productively with substrate. The role of the innermost CBM10-1 is more enigmatic, but at low enzyme and substrate concentrations, it was able to secure highest activity of the CD toward the two poor substrates (INM and guar GG) also relative to the activity of the CD alone or with two, respectively, three CBM10s. The CBM10-1 is in a small separate phylogenetic subgroup for which a possible hallmark and biological role, beyond this mentioned stabilization, remains to be identified.

**Experimental procedures**

**Bioinformatics**

CBM10 sequences were retrieved from the CAZy database ([www.cazy.org; (10)](www.cazy.org)) and subjected to multiple sequence alignment using mULTiple Sequence Comparison by Log-Expectation tool (48). A phylogenetic tree was constructed using BIONJ (49) and visualized using Dendroscope, version 3.6.3 ([https://uni-tuebingen.de/fakultaeten/mathematisch-naturwissenschaftliche-fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/dendroscope/](https://uni-tuebingen.de/fakultaeten/mathematisch-naturwissenschaftliche-fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/dendroscope/)) (50). A structure-based alignment of selected characterized CBM10s was generated using the PROMALS3D Web server (51). Homology models of the three GH5CBM10s were made using HHpred ([https://toolkit.tuebingen.mpg.de/tools/hhpred](https://toolkit.tuebingen.mpg.de/tools/hhpred)) (52) with the structure of *Cj*GH5CBM10x3 fused with GFP, and GFP (AGT98536) was purchased (GeneArt; Thermo Fisher Scientific), and cloned into the pET28a(+) vector using the XhoI and NheI restriction sites resulting in an N-terminal cleavable His tag. The three C-terminally truncated forms *Sd*GH5_8-CBM10x2, *Sd*GH5_8-CBM10x1, and *Sd*GH5_8 (Fig. 2) were constructed by introducing stop codons using Q5 Site-Directed Mutagenesis Kit (New England Biolabs), pET28a-*Sd*GH5_8-CBM10x3 as template and mutagenesis primers (Table S1) designed according to the manufacturers’ instructions. Each of the three *Sd*GH5CBM10s were obtained as *Sd*GH5CBM10 to 1_GFP, *Sd*GH5CBM10 to 2_GFP, and *Sd*GH5CBM10-3_GFP by mutagenesis using Q5 Site-Directed Mutagenesis Kit (New England Biolabs), pET28a-*Sd*GH5CBM10x3_GFP as template and mutagenesis primers. To get *Sd*GH5CBM10-1_GFP, both *Sd*GH5CBM10-2 and *Sd*GH5CBM10-3 were deleted in one step. One-step deletion of *Sd*GH5CBM10-1 and *Sd*GH5CBM10-2 was done to get *Sd*GH5CBM10-3_GFP. Finally, *Sd*GH5CBM10-2_GFP was obtained in two steps; first deleting CBM10-1 and then CBM10-3. The gene constructs were confirmed by sequencing.

**Carbohydrates**

CGM-Lv, INM, and KGM were from Megazyme; high-viscosity CGM (locust bean gum), GG, microcrystalline cellulose (Avicel), and HEC were from Sigma–Aldrich. See Figure 1 for the polysaccharide structures.

**Gene constructs**

The genes encoding the *S. degradans* β-mannanase *Sd*GH5_8-CBM10x3 (GenBank accession no.: ABD79918) without the predicted signal peptide (amino acids 1–21) and the three CBM10s *Sd*GH5CBM10x3 fused with GFP, and GFP (AGT98536) were purchased (GeneArt; Thermo Fisher Scientific), and cloned into the pET28a(+) vector using the XhoI and NheI restriction sites resulting in an N-terminal cleavable His tag. The three C-terminally truncated forms *Sd*GH5_8-CBM10x2, *Sd*GH5_8-CBM10x1, and *Sd*GH5_8 (Fig. 2) were constructed by introducing stop codons using Q5 Site-Directed Mutagenesis Kit (New England Biolabs), pET28a-*Sd*GH5_8-CBM10x3 as template and mutagenesis primers (Table S1) designed according to the manufacturers’ instructions. Each of the three *Sd*GH5CBM10s were obtained as *Sd*GH5CBM10 to 1_GFP, *Sd*GH5CBM10 to 2_GFP, and *Sd*GH5CBM10-3_GFP by mutagenesis using Q5 Site-Directed Mutagenesis Kit (New England Biolabs), pET28a-*Sd*GH5CBM10x3_GFP as template and mutagenesis primers. To get *Sd*GH5CBM10-1_GFP, both *Sd*GH5CBM10-2 and *Sd*GH5CBM10-3 were deleted in one step. One-step deletion of *Sd*GH5CBM10-1 and *Sd*GH5CBM10-2 was done to get *Sd*GH5CBM10-3_GFP. Finally, *Sd*GH5CBM10-2_GFP was obtained in two steps; first deleting CBM10-1 and then CBM10-3. The gene constructs were confirmed by sequencing.

**Protein production**

All full-length and variant-encoding plasmids were transformed into *E. coli* BL21, screened on LB agar including 50 µg/ml kanamycin, and starter cultures (10 ml) were made by inoculating LB medium, 50 µg/ml kanamycin with a single colony and incubated (37 °C, overnight), which are used to inoculate 750 ml LB medium containing 10 mM glucose and 50 µg/ml kanamycin in shake flasks. Cultures were propagated (30 °C, 160 rpm) to an absorbance of 0.6 at 600 nm followed by decreasing the temperature to 16 °C and induction of expression by a final concentration of 0.1 mM IPTG. Cells were harvested (4000g, 20 min, 4 °C) after approximately 20 h and stored at −20 °C until protein purification.

**Protein purification**

The proteins were purified in two steps. Cells were resuspended in HisTrap equilibration buffer (10 mM Hepes, pH 7.4, 10 mM imidazole, 0.5 M NaCl, and 10% glycerol), added two cOmplete Mini, EDTA-free, protease inhibitor tablets (Roche Diagnostics GmbH), lysed using a high-pressure homogenizer at 1 bar, added 3 µl Benzonase Nuclease (Sigma–Aldrich), and centrifuged (40,000g; 4 °C, 30 min). The filtered (0.45 µm) supernatant was loaded onto a 5 ml HisTrap HP column (GE Healthcare) pre-equilibrated with HisTrap equilibration buffer at a flow of 3 ml/min and eluted by a linear gradient from 2.5% to 100% HisTrap elution buffer (10 mM Hepes, pH 7.4, 400 mM imidazole, 0.5 M NaCl, and 10% glycerol) in 20 column volumes. Protein-containing fractions (based on absorbance at 280 nm) were pooled, concentrated to 4 ml (10 kDa Amicon Ultra centrifugal filters; Merck–Millipore) and further purified by gel filtration (Superdex 16/60 75 equilibrated with 10 mM Hepes, pH 7.0, 150 mM NaCl, and 10% glycerol) at a flow rate of 0.5 ml/min. Fractions containing recombinant protein of correct size as...
based on SDS-PAGE were pooled and concentrated as above to 2 to 3 mg/ml. The absorbance at 280 nm was measured, and the protein concentration was determined using theoretical extinction coefficients.

**Substrate specificity**

The specific activity was determined essentially as previously described (18, 53) using a reducing end 3,5-dinitrosalicylic acid (DNS) assay. Briefly, full-length and truncated enzyme forms (0.4–9.1 nM) were incubated with low- and high-viscosity galactomannans (2.5 mg/ml, 10 min), KGM (2.5 mg/ml, 10 min), GG (2.5 mg/ml, 2.5 h), INM (5 mg/ml, 45 min), or Avicel (10 mg/ml, 3 h) under standard activity assay conditions (400 μl assay volume, 40 mM sodium phosphate citrate pH 6.0, 0.005% [v/v] Triton X-100, 37 °C) in triplicates. The reactions were stopped by addition of 600 μl DNS reagents followed by heat treatment at 95 °C for 15 min. After 15 min in ice water, the samples were centrifuged at 20,000g for 12 min. Finally, the absorbance was measured at 540 nm.

**Kinetic analysis**

The kinetics on CGM-lv, GG, and INM were determined using the aforementioned DNS assay with a range of substrate concentrations (0.45–9 mg/ml CGM-lv; 1.25–7.5 mg/ml GG; and 2.5–20 mg/ml INM) and enzyme (full-length and truncated forms: 1.5–2.8 nM for CGM-lv, 13–25 nM for GG, and 99–179 nM for INM) in 2000 μl and withdrawing 400 μl aliquots (200 μl in the case of INM) at 3, 6, 9, and 12 min. The kinetic assays were done in duplicates. In the case of the CGM-lv and INM data, the Michaelis–Menten model was fitted to the initial velocity data using GraphPad Prism 6 (GraphPad Software Inc).

**Pull-down assay**

Qualitative screening of binding of the individual CBM10s separately and the triple CBM10 to microcrystalline cellulose (Avicel) and insoluble crystalline INM was done by a pull-down assay, where the samples were analyzed by SDS-PAGE: 10 mg polysaccharide (prewashed three times) was mixed with 200 μl 0.5 mg/ml protein in assay buffer (50 mM phosphate buffer, pH 7.0). GFP was used as a control in the assay. The samples were incubated at 4 °C for 1 h with gentle agitation and then centrifuged (20,000g, 10 min, 4 °C). Supernatants were transferred to fresh tubes and centrifuged (20,000g, 10 min, 4 °C), before 4 μl supernatant was heat treated in the presence of SDS-loading buffer and applied on the SDS-PAGE. The pellets from the pull-down assay were washed with 250 μl assay buffer and pelleted again as aforementioned, resuspended in 200 μl assay buffer, added 50 μl SDS-loading buffer, boiled for 10 min, and applied (4 μl) on the SDS-PAGE. For each protein, a sample of the 0.5 mg/ml protein stock used for the pull-down assay was treated as the supernatant sample and included on the gel.

Quantitative pull-down assay was done with microcrystalline cellulose (Avicel) utilizing the presence of GFP for quantification (19). The total assay volume (750 μl) contained 10 mg/ml Avicel (washed with assay buffer thrice) and the different SdGH5CBM10-GFP forms in ten different concentrations (10–100 μg/ml) in 50 mM phosphate pH 7.0 and 0.1 mg/ml bovine serum albumin. GFP (10–150 μg/ml) served as a negative control. The mixtures were rotated end over end (4 °C, 1 h) and centrifuged (20,000g, 10 min, 4 °C). Supernatants were transferred to fresh tubes, centrifuged (20,000g, 10 min, 4 °C), and the concentration of unbound protein was determined by fluorescence (FP-8500 Spectrophotometer equipped with an FMP-825 plate reader; JASCO Corporation) with excitation at 450 and emission at 510 nm using a standard curve made with each protein form (1–1250 nM) in assay buffer. The amount of bound protein was calculated by subtracting the unbound from the total protein, and the dissociation constant, \( K_d \), was obtained by fitting Equation 1 to binding isotherm data (GraphPad Prism, version 6).

\[
[PC] = \frac{[FP][PC]_{\text{max}}}{K_d} \tag{1}
\]

PC and FP are the bound and unbound protein concentrations, respectively. The pull-down assays were done in duplicates on two different days.

**AGE**

Qualitative screening of binding of all protein forms and GFP (as control) was done by AGE with 2.5 mg/ml CGM-lv or HEC in 12% acrylamide gels (54). Protein (10 μl 0.4 mg/ml) in loading buffer was applied to polysaccharide-containing and control (without polysaccharide) gels, respectively. NativeMark unstained protein standard (Invitrogen) was included on all gels for normalization. The gels were run (20 h at 4 °C and 45 V) and stained using InstantBlue (Expedeon). Quantitative protein binding analysis to CGM-lv was done from migration and retardation in gels without (control) or with 0.125, 0.25, 0.5, 1, 1.5, and 2.5 mg/ml CGM-lv, respectively. \( K_d \) was determined from retardation distances according to the method (54) using a Takeo–Nakamura plot (55) (Equation 2):

\[
\frac{1}{R_{mi}} = \frac{1}{R_{mo}} \left(1 + \frac{c}{K_d} \right) \tag{2}
\]

\( R_{mi} \) and \( R_{mo} \) are migration distances of sample relative to a reference protein in the presence and absence of polysaccharide, respectively; \( c \) is the polysaccharide concentration. The relative migration (Equation 3) defines the retardation of migration by the polysaccharide compared with the control.

\[
R_m = \frac{R_{mi}}{R_{mo}} \tag{3}
\]

**Data availability**

All the data are contained within the article.

**Supporting information**—This article contains supporting information.
Acknowledgments—We thank Karina Jansen (Technical University of Denmark) for technical assistance. We also thank Prof Jens Preben Morth (Technical University of Denmark) for introduction to the JASCO FP-8500 Fluorescence Spectrometer. This work was supported by Independent Research Fund Denmark|Natural Sciences (FNU) by a grant to the project “Roles and structural determinants of low-affinity carbohydrate-protein interactions” (grant no. 6108-00476B).

Author contributions—M. S. M. and B. S. conceptualized the work and designed the experiments. M. S. M. and B. H. performed phylogenetic analysis. M. S. M. and S. E. B. performed the experiments. M. S. M. and B. S. analyzed the data. M. S. M. wrote the article with input from B. S. B. S. acquired the funding.

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

Abbreviations—The abbreviations used are: AGE, affinity gel electrophoresis; CBMCC, bacterial microcrystalline cellulose; CBM, carbohydrate-binding module; CBM10, CBM family 10; CD, catalytic domain; CGM-Iv, low-viscosity carob galactomannan; DNS, 3,5-dinitrosalicylic acid; GG, guar gum; GH, glycoside hydrolase; domain; CGM-lv, low-viscosity carob galactomannan; DNS, 3,5-dinitrosalicylic acid; CBM10, CBM family 10; CD, catalytic domain; CGM-lv, low-viscosity carob galactomannan; DNS, 3,5-dinitrosalicylic acid; CBM64C, CBM64C.; Tailored catalysts for breaching the intestinal mucus barrier.

References

Functional diversity of CBM10s


