Family 6 Carbohydrate Binding Modules Recognize the Non-reducing End of β-1,3-Linked Glucans by Presenting a Unique Ligand Binding Surface*

Received for publication, September 2, 2004, and in revised form, October 20, 2004. Published, JBC Papers in Press, October 22, 2004. DOI 10.1074/jbc.M410113200

Alicia Lammerts van Bueren‡, Carl Morland§, Harry J. Gilbert§, and Alisdair B. Boraston‡¶

From the ‡Department of Biochemistry and Microbiology, University of Victoria, P. O. Box 3055 STN CSC, Victoria, British Columbia V8W 3P6, Canada and the §School of Biomedical Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, United Kingdom

Enzymes that hydrolyze insoluble complex polysaccharide structures contain non-catalytic carbohydrate binding modules (CBMs) that play a pivotal role in the action of these enzymes against recalcitrant substrates. Family 6 CBMs (CBM6s) are distinct from other CBM families in that these protein modules contain multiple distinct ligand binding sites, a feature that makes CBM6s particularly appropriate receptors for the β-1,3-glucan laminarin, which displays an extended U-shaped binding mode. The binding cleft in this protein is sealed at one end, which prevents binding of linear polysaccharides such as cellulose, and the orientation of the sugar at this site prevents glycone extension of the ligand and thus conferring specificity for the non-reducing ends of glucans. The high affinity for extended β-1,3-glucooligosaccharides is conferred by interactions with the surface of the protein located between the two binding sites common to CBM6s and thus reveals a third ligand binding site in family 6 CBMs. This study therefore demonstrates how the multiple binding clefts and highly unusual protein surface of family 6 CBMs confers the extensive range of specificities displayed by this protein family. This is in sharp contrast to other families of CBMs where variation in specificity between different members reflects differences in the topology of a single binding site.

The plant cell wall represents the largest reservoir of organic carbon in the biosphere, and thus its degradation by microbial enzymes is pivotal to many biological and industrial processes.

* This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The three-dimensional structure of almost all Type B CBMs determined to date conform to a classic lectin-like β-jelly roll in which a single ligand binding site comprises a shallow cleft on the concave surface of the protein. In family 4 and family 2 CBM variation in ligand recognition between different members of these families is reflected in the structure of this conserved binding site (10). By contrast family 6 CBMs (CBM6s), which also adopt the classic β-jelly roll fold, contain two potential ligand binding sites; one (cleft B) on the concave surface and a second (cleft A) located on one edge of the protein between the loops that connect the inner and outer β-sheets (11). The variation in ligand specificity in CBM6a reflects the functionality of these two binding sites with xylan- and β-(1,4)-1,3 mixed glucan binding modules accommodating these polysaccharides in cleft A and cleft B, respectively (12). One of the features that distinguish Type B CBMs from lectins is the mechanism of ligand recognition. Each binding
site in lectins recognizes one or two sugars through an extensive network of hydrogen bonds, while Type B CBMs generally accommodate four to six sugars, with specificity conferred primarily by the conformation of the ligand, which reflects the topology of the binding site. Although it is well established that the orientation of the aromatic residues in the binding site of CBMs confers specificity for the planar and 3-fold helical conformations of cellulose and xylan, respectively (9), the mechanism by which these protein modules recognize other polysaccharides, which display more elaborate conformations, is unclear. An example of polysaccharides that exhibit complex conformations is provided by the β-1,3-linked glucose polymer laminarin, which adopts an extensive U-shaped conformation, and thus cannot be accommodated in CBMs that contain linear clfts (9). The two potential ligand binding sites in CBM6s may present structural features that make these proteins ideally suited to accommodate polysaccharides with extended U-shaped conformations. To assess this hypothesis we have determined the structure-function relationship of a CBM6, designated BhCBM6, located in an enzyme that displays laminarinnase activity. BhCBM6 does indeed bind to laminarin displaying maximum affinity for ligands with a d.p. >5. Uniquely, the protein displays absolute specificity for the non-reducing end of laminarin chains, and the crystal structure of BhCBM6 in complex with laminarihexaose shows that the ligand extends out of cleft A interacting with the surface of the protein that does not encompass either cleft A or cleft B. These data demonstrate that the remarkable flexibility in ligand specificity displayed by CBM6s reflects variation in the location of the carbohydrate interacting sites on the surface of the protein, and is not exclusively the result of differences in the topology of a single binding site.

**MATERIALS AND METHODS**

**Carbohydrates and Polysaccharides—**Xyloglucan oligosaccharides, laminarioligosaccharides, Konjac glucomannan, wheat arabino-xylan, Tamarind xyloglucan, and oat β-glucan were obtained from Megazyme International Ireland Ltd. (Bray Co., Wicklow, Ireland). All other carbohydrates and glycoproteins were purchased from Sigma.

**Cloning of Catalytic Domain and BhCBM6 of the Laminariase—**The DNA fragment (nucleotides 76–2328) of the laminarinnase gene (see GenBank® Accession No. AP001507; open reading frame BH0236) encoding the catalytic domain of the enzyme was amplified by PCR from *Bacillus halodurans* (C-125) genomic DNA (ATCC BAA-125) using the primers 5′-CACCTCCCCTCATGCGGT-3′ and 5′-CACCCCTCCTGCGTG-GAGC-3′. The amplified product was ligated into pET-151 TOPO (Invitrogen, San Diego, CA) to generate pAB1. The encoded polypeptide contained an N-terminal His6/V5 epitope tag and a TEV protease cleavage site. The DNA fragment encoding BhCBM6 (nucleotides 2367–2775 of the laminarinnase gene) was amplified using the primers 5′-CATGCGTCGA-GATTGGAAAATCTCCCTACAG-3′ (an Nhel site is underlined), and 5′-GCCCAGCCCCCCACGTTTTAAGCGCTTGCCTGGAACC-3′ (a HindIII site is underlined; stop codon in bold) and cloned into Nhel- and HindIII-digested pET28a to give pAB2. The encoded polypeptide contains an N-terminal His6 tag and a thrombin cleavage site.

**Expression and Purification of the Catalytic Domain and BhCBM6 of the Laminariase—**The catalytic domain of the laminarinnase was produced in *Escherichia coli* strain TUNER (Novagen) containing pAB1, and the protein was purified from cell-free extracts by immobilized metal ion affinity chromatography (IMAC) following the method of Freelove et al. (14) except the recombinant protein was eluted with 10 mM imidazole. BhCBM6 was produced in *E. coli* BL21(DE3) containing pAB2 as described previously (13), and the protein was purified by IMAC following the method of Boraston et al. (13). Purified polypeptides were concentrated and exchanged into distilled water in a stirred ultrafiltration unit (Amicon, Beverly, MA) on a 5000 molecular weight cut-off (MWCO) membrane (Filtron, Northborough, MA). Purity, assessed by SDS-PAGE, was greater than 95%.

**Enzyme Activity Assay—**Enzyme reactions were carried out as described previously (15) using 0.2% substrate. The concentration of purified protein was determined by UV absorbance (280 nm) using calculated molar extinction coefficients (16) of 184,060 M⁻¹ cm⁻¹ and 36,130 M⁻¹ cm⁻¹, for the catalytic domain of the laminariase and BhCBM6, respectively.

**UV Difference Titrations—**Automated UV difference titrations were performed as described previously (17, 18) using a Shimadzu UV2600 CCD spectrometer (Ocean Optics, Dunedin, FL) with a diffraction grating providing measurements at 2048 approximately evenly spaced wavelengths between 234 and 395 nm. Difference spectra were examined for peak and trough wavelengths, and values at the appropriate wavelengths extracted for further analysis. The peak-to-trough wavelengths at the highest power wavelengths 289.3/301.0 nm, 289.3/294.5 nm, and 282.8/287.4 nm were calculated by subtraction of the trough values from the peak values, and the dilution-corrected data were plotted against total carbohydrate concentration. Data for the three wavelength pairs were analyzed simultaneously with MicroCal Origin (v.7.0) using a one site binding model accounting for ligand depletions (17). Experiments were performed at 20 °C in 50 mM Tris, pH 7.5. The data reported are the averages and standard errors of the means of three independent titrations.

**Isothermal Titration Calorimetry—**Isothermal titration calorimetry (ITC) was performed as described previously (13) using a VP-ITC (MicroCal, Northampton, MA) in 50 mM potassium phosphate buffer (pH 7.0) at 25 °C using 100–250 mM BhCBM6 in the reaction cell and 1–5 mM oligosaccharides in the syringe, which gave C-values >10. Reverse titrations were performed by adding the oligosaccharides to the enzyme at 0°C. Data were fitted to a single-site binding model using a one site binding model accounting for ligand depletions (17). Experiments were performed at 20 °C in 50 mM Tris, pH 7.5. The data reported are the averages and standard deviation of two or three independent titrations.

**Crystalization of BhCBM6—**BhCBM6 was treated overnight with thrombin at room temperature, concentrated, and buffer exchanged into water in a 10-mL stirred ultrafiltration device using a 5000 MWCO membrane. Crystals of BhCBM6 (25 mg/ml) were grown at 18 °C using the vapor-phase diffusion technique from hanging drops in 24% polyethylene glycol 2000 monomethylether, 0.2 M sodium citrate, 0.1 M MES, pH 6.5, and 3% glycerol. Crystals of BhCBM6 in complex with xylose were prepared by adding xylose powder directly to hanging drops containing crystals and allowing these to equilibrate for ~72 h. Crystals of BhCBM6 (20 mg/ml) in complex with excess laminariase were grown using the same technique in 10-μl drops with 0.1 M MES, pH 6.5, containing 1.8 M ammonium sulfate as the mother liquor.

**Data Collection, Structure, Solution, and Refinement—**All computing was done using the CCP4 suite (20) unless otherwise stated. Uncomplexed or xylene complexed crystals were frozen at 113 K after a short soak in artificial mother liquor supplemented with glycerol at 20% before seeding. Crystals of BhCBM6 in complex with laminariase were cryoprotected in the same manner with mother liquor containing ethylene glycol at 20% (v/v). Data were collected with a Rigaku R-Axis 4 ++ area detector coupled to a MM-002 x-ray generator with Osmic “blue” optics and an Oxford Cryostream 700. Data were processed using the CrystalClear dtrek (21) software provided with the instrument. In all data sets, five percent of the observations were flagged as free (22) and used to monitor refinement procedures. In the case of the uncomplexed BhCBM6 and xylene complex, the same reflections were flagged as free. Statistics are given in Table I for those crystals and data sets used in the structure solution and refinement.

**Using the data for the triclinic uncomplexed crystals and the coordinates of the CBM6 from the *Clostridium thermocellum* xylanase Xyn10B (PDB ID 1GMM; Ref. 11) as a search model, the program Molrep (23) was able to find two molecular replacement solutions corresponding to the two BhCBM6 molecules in the asymmetric unit. One molecule was corrected by successive rounds of building using XtalView (24) and refinement with REFMAC (25). This corrected model was used to replace the second molecule in the unit cell followed by additional rounds of building and refinement. This model was used directly as a starting point in the building and refinement of the xylene complex. Molecular Replacement using the refined uncomplexed coordinates was used to solve the structure of the laminarihexaose complex. This model was corrected, and the laminarihexaose molecule was built manually in XtalView followed by refinement with REFMAC. Water molecules were added using REFMAC/ARP-WARP and inspected visually prior to deposition. All final model statistics are given in Table I. Figs. 4, 5, and 7 were prepared with PyMOL (see URL pymol.sourceforge.net) and are shown in divergent stereo. The native BhCBM6, xylene complex and laminarihexaose complex have been deposited with the PDB codes of 1W9S, 1W9T, and 1W9W, respectively.
RESULTS AND DISCUSSION

The Modular Architecture and Catalytic Activity of the \textit{B. halodurans} Laminarinase—The alkalophilic bacterium \textit{Bacillus halodurans} contains an open reading frame (BH0236) that encodes a 1020 amino acid of unknown function. Based on PSIBLAST amino acid sequence alignments (26) the encoded protein appears to comprise three modules (Fig. 1). The N-terminal module of this protein shows similarity with family 81 glycoside hydrolases, a family of proteins in which some members display \(\beta\)-1,3-glucanase activity, exemplified by Eng1p from \textit{Saccharomyces cerevisiae}, which is involved in cell separation (27, 28), whereas others, an example of which is the Glycine max xylan binding CBM6, is devoid of catalytic activity but do bind \(\beta\)-1,3-glucans (29). The putative glycoside hydrolase catalytic module from the \textit{B. halodurans} protein is the only bacterial example in this family. The recombinant catalytic module of this enzyme hydrolyzed laminarin with an activity of 14,231 min\(^{-1}\) (14,231 mol of reducing sugar produced per mol of enzyme per min) but displays no detectable activity against oat-splet xylan, wheat arabinoxylan, xyloglucan, lichenan, galactan, arabinan, carob galactomannan, \textit{kun}jac glucomannan, \(\beta\)-glucan, polygalacturonate, pectin, amylase, amylopectin, hydroxyethyl cellulose, or carboxymethyl cellulose. Thus, the \textit{B. halodurans} protein is clearly a \(\beta\)-1,3-glucanase (or laminarinase), similar to its eukaryotic homologues.

The C-terminal \(\sim 100\) amino acid module has no identity to proteins of known function. Separating the C-terminal module and the N-terminal catalytic module is a module of \(\sim 140\) amino acids module having \(36\%\) identity to the xylan binding CBM6 from \textit{C. thermocellum} xylanase 10A (11). Based on its identity with CBM6s and its presence in a functional \(\beta\)-1,3-glucanase, we hypothesized that this module, defined as BhCBM6, is a CBM with \(\beta\)-1,3-glucan binding specificity.

\textbf{Analysis of BhCBM6 Binding Specificity}—Using a standard depletion binding analysis (30) no binding to the insoluble polysaccharides regenerated cellulose or pachyman (an insoluble \(\beta\)-1,3-glucan from \textit{Poria cocos}) was evident (data not shown). Similarly, native affinity gel electrophoresis did not reveal significant binding to soluble preparations of wheat arabinoxylan, amylopectin, oat \(\beta\)-glucan, or lichenan, the latter two of which are commonly recognized by \(\beta\)-1,3-glucan binding CBMs (17, 31, 32). Using a modification of the macroarray method of McCartney \textit{et al.} (33) BhCBM6 showed weak binding to wheat arabino-xylan, birchwood glucurono-xylan, and pectic galactan, but did not interact with polysaccharides containing \(\beta\)-(1,3)(1,4) linked glucan, \textit{i.e.} oat \(\beta\)-glucan (data not shown).

\textbf{UV Difference Studies of BhCBM6 Binding}—Xylose, xyloglucosaccharides, and O-methyl-\(\beta\)-d-xylo induced relatively large changes in the UV absorbance difference spectrum indicating binding to these sugars. The addition of glucose, sophorose (\(\beta\)-1,2-glucobiose), and \(\beta\)-1,3-glucosaccharides (laminarioligosaccharides) also gave UV difference signals (see Fig. 2A for a representative UV difference spectrum), apparently with the inability to detect binding to glucose-based polymers by affinity electrophoresis and macroarray assays. No perturbation of the UV absorbance spectrum was observed when cellulose, cellobiose, sucrose, mannose, galactose, fucose, \textit{N}-acetylglucosamine, or \textit{N}-acetylgalactosamine was added to BhCBM6.

Three wavelength pairs in the UV difference spectra were used to monitor the dependence of the UV absorbance spectra on ligand concentration and quantify binding to carbohydrates (Fig. 2B). BhCBM6 bound to glucose and xylose with association constants \(K_a\) of \(\sim 8 \times 10^9\) M\(^{-1}\) (Table II). Xylose was bound \(2\times\) more tightly indicating a small dependence of binding on ligand length. No additional gains in affinity were observed when the ligand length was increased to xylotriose (Table II). Gentibiose (\(\beta\)-1,2-glucobiose) and sophorose (\(\beta\)-1,2-glucobiose) were bound with affinities similar to that for xylose (Table II). The \(\beta\)-1,3-glucosaccharides laminariobiose, laminaritetraose, and laminarihexaose were bound with affinities similar to that for xylose (Table II). The \(\beta\)-1,3-glucosaccharides and the dependence of affinity on sugar length. Thus, despite the inability to detect binding to polysaccharides containing \(\beta\)-1,3-linked glucose residues by depletion to wheat arabino-xylan, birchwood glucurono-xylan, and pectic galactan, but did not interact with polysaccharides containing \(\beta\)-(1,3)(1,4) linked glucan, \textit{i.e.} oat \(\beta\)-glucan (data not shown).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
 & Uncomplexed & Xylobiose & Laminarihexaose \\
\hline
Spacegroup & P1 & P1 & P4,22 \\
Cell dimensions & \(a = 30.8\ A, b = 40.8\ A, c = 56.1\ A, \alpha = 108.9^\circ, \beta = 105.9^\circ, \gamma = 90.0^\circ\) & \(a = 30.7\ A, b = 41.0\ A, c = 56.0\ A, \alpha = 108.9^\circ, \beta = 105.9^\circ, \gamma = 90.0^\circ\) & \(a = 61.6\ A, c = 121.7\ A; \alpha = \beta = \gamma = 90^\circ\) \\
Max. resolution of data (\(\AA\)) & 20–1.59 (1.65–1.59) & 20–1.62 (1.65–1.62) & 20–2.10 (2.17–2.10) \\
Completeness & 90.1 (98.5) & 91.9 (84.4) & 95.6 (93.5) \\
\(I/\sigma\) & 24.9 (5.7) & 14.9 (4.5) & 17.8 (4.4) \\
\(R_{merge}\) (\%) & 3.0 (16.8) & 4.8 (23.7) & 5.5 (37.2) \\
Multiplicity & 3.7 (3.2) & 3.6 (3.2) & 5.8 (6.0) \\
\(R_{free}\) (\%) & 12.1 & 12.1 & 20.9 \\
\(B\)-factor model & Anisotropic & Anisotropic & Isotropic \\
Mean \(B\)-factor protein (\(\AA^2\)) & 15.2 & 16.7 & 50.8 \\
Mean \(B\)-value ligand(s) (\(\AA^2\)) & 16.7 (Na and glycerol) & 27.7 (Na and sugar) & 49.0 (Na and sugar) \\
Number of residues & 266 & 266 & 134 \\
Number of waters & 372 & 372 & 121 \\
Number of sugar atoms & 72 (in 4 xylobose) & 67 (in 1 laminarihexaose) & 1 \\
PDB code & 1W9S & 1W9T & 1W9W \\
\hline
\end{tabular}
\caption{Data collection and structure statistics}
\end{table}
binding, affinity electrophoresis, and macroarray experiments, BhCBM6 does indeed appear to be primarily a \( \beta \)-1,3-glucan-specific CBM, consistent with its presence in a \( \beta \)-1,3-glucanase.

**Thermodynamics of \( \beta \)-1,3-Glucan Recognition**—Isothermal titration calorimetry revealed the interaction of BhCBM6 with \( \beta \)-1,3-glucooligosacharides to be enthalpically favorable and entropically unfavorable at 25 °C (Table III), like most protein-carbohydrate interactions studied to date. The near unitary stoichiometries indicated the formation of a simple 1:1 protein to carbohydrate macromolecular complex. Typically, CBMs that bind to extended glycan chains (Type B CBMs, Ref. 9) show a consistent increase in binding affinity with increasing sugar length. BhCBM6 showed a decrease in affinity when comparing laminaribiose with laminaritriose though the affinity consistently increased when moving from laminaritriose up to laminarihexaose (Table III). The loss of affinity when binding laminaritriose versus laminaribiose appeared to be caused by an enthalpic penalty that more than offset a favorable gain in entropy. The reason for this is unclear, even in light of the structure of BhCBM6 in complex with laminarihexaose (see below), which revealed no obvious peculiarity in how this third sugar interacts with the protein.

BhCBM6s binding to laminarin was assessed by ITC in a mode where the protein was titrated into the polysaccharide. The resulting binding isotherms, which were highly reproducible, were not consistent with a single class of binding interaction (Fig. 2C). At least two binding phases were visually evident: one at low molar ratios (up to ~0.2) and another at higher molar ratios. A Scatchard analysis of these data, which were non-linear (not shown), confirmed the complexity of the isotherms. These results suggest either multiple classes of non-equivalent binding sites present in the laminarin population or cooperativity in the binding. We cannot conclusively comment on this and because of the apparently complex nature of laminarin recognition the kinetic and thermodynamic parameters of binding for the individual interactions could not be accurately deconvoluted. However, estimates from the analysis of the cumulative heats indicated association constants in the range of \( 10^9 \text{ M}^{-1} \), consistent with the affinities for laminari-oligosaccharides. The total number of binding sites was approximated by graphical determination of the stoichiometric limit of the interaction in the isotherms plotted in derivative form (standard presentation of ITC data) and cumulative form (not shown). This gave an \( n \) value of \( \approx 0.8 \) (i.e. \( \sim 1 \) CBM molecule per laminarin chain), which is consistent with BhCBM6 binding only to the ends of laminarin (see below). The binding to xylan and pectic galactan was too weak to quantify by this method.

**Structure of BhCBM6**—In order to gain insight into the mechanism of carbohydrate recognition by BhCBM6 we solved its three-dimensional structure by x-ray crystallography (see “Materials and Methods”). The final model of BhCBM6 in the absence of ligand consisted of two BhCBM6 molecules (133 amino acids each), six sodium atoms, two glycerol molecules,

---

**Table II**

Affinity of BhCBM6 for sugars determined by UV difference titrations at 20 °C in 50 mM Tris, pH 7.5

| Carbohydrate          | \( K_a \) (M\(^{-1}\)) |
|-----------------------|-------------------------|
| \( \beta \)-D-Glucose | \( >10^9 \)          |
| Laminaribiose         | 0.1 (±0.0)             |
| Laminartetraose       | 0.9 (±0.1)             |
| Laminarihexaose       | 1.8 (±0.4)             |
| Gentiobiose            | 10.2 (±2.9)            |
| Sorbitol              | 0.3 (±0.1)             |
| \( \beta \)-D-Xylose  | 0.6 (±0.2)             |
| O-Methyl-\( \beta \)-D-Xylose | 0.1 (±0.0) |
| \( \beta \)-1,4-Xylobiose | 0.3 (±0.0) |
| \( \beta \)-1,4-Xylotriose | 0.3 (±0.1) |

---

**Figure 2** UV difference and ITC analysis of BhCBM6 binding. Panel A, UV difference spectra collected with the indicated concentrations of added laminarihexaose. Peak and trough wavelengths are shown. Panel B, isotherm of laminarihexaose titrated into BhCBM6. The curves show the data at the wavelength pairs of 289.3/301.0 (open circles), 289.3/294.5 nm (closed squares), and 282.8/287.4 nm (closed circles). Solid lines show the global fits to a one-site binding model. Error bars represent the standard errors of 3 measurements. Panel C, isotherm of BhCBM6 binding to laminarin obtained by isothermal titration calorimetry (see “Materials and Methods” for experimental details). The lower panel shows the raw calorimetric data. The upper panel shows the raw calorimetric data. The lower panel shows the integrated data (closed circles) and the results of a heat of dilution experiment (open circles).
and 350 water molecules (repetition statistics are given in Table I). Like other CBM6s whose structures have been determined, BhCBM6 adopts a β-sheet fold with a 5-stranded β-sheet opposing a 4-stranded β-sheet (Fig. 3). This fold was highly similar to the xylan-binding CBM6 from a *C. thermocellum* xylanase (root mean-square-deviations (r.m.s.d.) of 0.81 Å² over 95 matched Ca) (11); the two xylan-binding CBM6s from a *Clostridium stercorarium* xylanase (r.m.s.d. of 0.71 Å² over 107 matched Ca and 0.83 Å² over 108 matched Ca, respectively) (34, 35) and the glucon binding CBM6 from a *Cellulibiovix mixtus*, endoglucanase (r.m.s.d. of 1.04 Å² over 95 matched Ca) (12).

Each BhCBM6 monomer coordinated 3 metal ions (Fig. 3). The first is coordinated by the side chain oxygens of Gin38, Glu18, and Asn134. The coordination is completed by the backbone carbonyl oxygens of Asn134, Asp90, and a single water molecule. The placement of this metal ion is conserved with the four other CBM6 structures, where these metal ions were modeled as calcium atoms. However, in the case of BhCBM6, the B-factor refined to an unreasonably high value when this atom was modeled as calcium, suggesting a metal with fewer electrons. Based on the presence of a relatively high concentration of sodium atoms in the conditions used to crystallize BhCBM6 the electron density corresponding to the metal ion was modeled as sodium. The second metal ion, also modeled as sodium, was coordinated by the side-chain oxygens of Asn134, Asp90, and Asp47. The backbone carbonyl oxygens of Trp42, Gly25, and Thr84 (of a separate, neighboring molecule) also participate in binding this atom. The third bound sodium ion is bound to the protein by one side chain oxygen of Asp90 and four water molecules and, thus, is somewhat tenuously associated. Indeed, while the first bound sodium is likely structurally significant, as its position appears to be conserved in this protein family, the significance of the other two is unclear.

**BhCBM6 in Complex with Xylobiose—**A complex of BhCBM6 was obtained by soaking unliganded P1 crystals in excess xylobiose. Electron density for two xylobiose molecules bound to each of the two monomers of BhCBM6 in the unit cell was clearly evident. The secondary binding site contained a somewhat disordered sugar molecule (not shown). The xylobiose molecule in this site made few interactions with the protein, with one being a potential hydrogen bond with a second neighboring BhCBM6 molecule. Thus, the biological significance of this binding site is uncertain and will not be discussed further. The other binding site, which accommodated a glycerol in the unliganded structure, is the conserved binding site among xylan-binding CBM6s. The electron density for the complete xylobiose molecule bound to this site (in both of the BhCBM6 monomers) allowed modeling of all of the atoms in this xylobiose molecule (Fig. 4). However, because of the inability to discriminate between the positions of C-5 and O-5, the direction of the sugar (i.e. reducing end versus non-reducing end) was somewhat ambiguous. This was resolved to some extent by examining the B-factors of C-5 and O-5; one orientation resulted in large B-factor discrepancies whereas in the other they were approximately equal. Furthermore, two well ordered water molecules were observed to be properly positioned to hydrogen bond to what was assumed to be the O-5 atoms in each of the xylose residues when oriented based on the B-factors. Thus, the carbohydrate was modeled as having the non-reducing sugar sandwiched between tryptophans 42 and 99 (Fig. 4), while the hydrogen bonding schematic reveals five potential direct hydrogen bonds (Fig. 5). Three potential water-mediated hydrogen bonds are present with one making numerous potential interactions and is structurally conserved with water molecules in other CBM6s (Figs. 4, 5, and 6A). The ability to bind in this orientation is supported by the capacity of BhCBM6s to interact with O-methyl-β-D-xylose (Table II). This sugar, with a blocked reducing end, bound better than its unmodified counterpart, indicating that BhCBM6 must be able to accommodate the non-reducing end of the sugar. Despite this, we currently have no evidence that BhCBM6 cannot also bind the reducing end of xylobiose. It is apparent that the non-reducing terminus of the sugar is oriented in the binding pocket such that the oligosaccharide chain cannot extend past Gin38 (i.e. over a wall of the binding pocket) without substantial distortions to the sugar. In contrast, the reducing end appears free to extend into solvent.

**BhCBM6 in Complex with Laminarihexaose—**BhCBM6 co-crystallized with laminarihexaose in the space group P41212 with a single protein and sugar molecule in the asymmetric unit. All six glucose residues of the laminarihexaose molecule could be unambiguously modeled (Fig. 4). The terminal glucose residue at the non-reducing end of this oligosaccharide sandwiched between tryptophans 42 and 99, as was modeled for xylobiose (Fig. 4).

The complex of BhCBM6 with laminarihexaose revealed six binding subsites (Fig. 5). The extended nature of this binding site and BhCBM6s preference for oligosaccharides with a degree of polymerization > 4 classify this CBM as a Type B CBM (9). Relatively few direct potential hydrogen bonds (≈ 10) were distributed throughout these subsites, consistent with the relatively low density of direct hydrogen bonds observed with other Type B CBMs (9). Numerous additional potential hydrogen bonds were mediated by five water molecules (Fig. 5).

The bound laminarihexaose molecule adopts a U-shape very similar to that observed in the x-ray crystal structure of the laminarin-binding CBM from a *Thermotoga maritima* laminarinase (Fig. 6 and Ref. 10) and similar to that predicted for β1,3-glucans in solution (36). Similar to the laminarin-binding CBM4, the constellation of interactions between BhCBM6 and laminarihexaose are unique to the conformation of this sugar and leverage against high affinity for other sugars. Unlike the family 4 CBM, which binds β1,3-glucans in a deep binding groove (Fig. 6C), the BhCBM6 binding site begins with a small “slot” only sufficient to accommodate the glucose residue at the reducing end of the oligosaccharide (Fig. 6B). The remainder of the oligosaccharide curls around the CBM to form a crown (Fig. 6B). The U shaped, or open helical, conformation of laminarin enables the ligand to curl over the walls that form the slot-like binding site and maintain interactions with the protein. In contrast, polysaccharides that have a 2-fold or 3-fold linear axis, such as cellulose and xylan, respectively, would simply extend out into solvent and not make any further direct interactions with the protein surface, explaining why xylose dis-
plays the same affinity for the protein as xylooligosaccharides and no increase in affinity is observed for xylose polymers. The sugar at the non-reducing end of laminarihexaose interacts with the terminal region of the CBM6 binding site such that O-3 is pointing directly at the protein surface sterically occluding extension of the sugar polymer and thus conferring specificity for the non-reducing end of the polysaccharide.

Comparison of CBM6 Structures That Lead to the Differences in Specificity—Comparison of the structures of CBM6 modules that recognize xylan and mixed linked \((1,4)/(1,3)\) glucans, respectively, with \(Bh\) CBM6, which binds to laminarin, provides novel insights into the structural basis for the extensive range of ligand specificities displayed by this family of proteins. In the xylan binding CBM6s cleft A is open at both ends explaining why these proteins are able to bind to the internal regions of the xylose polymers, with the central sugar in this site sandwiched between two aromatic residues (Fig. 7). The surface of the cleft is likely to clash with the C-6 hydroxymethyl group of the pyranose ring at subsites 4 and 5, legislating against tight binding to glucose-containing ligands. By contrast, the primary ligand binding site in the \(C.\ mixtus\) CBM6 that recognizes \((1,4)/(1,3)\) mixed linked glucans is in cleft B, which lacks the extended loop that occludes this binding site in the other CBM6 modules. Cleft A in the \(C.\ mixtus\) CBM does display weak affinity for the terminal sugars of both xylose and glucose-containing oligosaccharides, but not the internal regions of the respective polysaccharides. The monosaccharide is again sandwiched between the parallel aromatic residues but is orientated 90° relative to the position of the sugar in the xylan binding modules and thus C-1 is pointing at the surface.
Fig. 5. A schematic showing the interactions of BhCBM6 with xylobiose (A) and laminarihexaose (B). Binding subsites referred to in the text are shown above the schematics with brackets and are numbered in accordance with IUPAC nomenclature. The water molecule conserved in the cleft A binding site of CBM6s is indicated with an arrow.

Fig. 6. Solvent accessible surface of BhCBM6 complexed with xylobiose (A) and laminarihexaose (B) and the family 4 CBM from T. maritima, TmCBM4–2, in complex with laminarihexaose (C). Purple regions indicate the surface contributed by the binding site apolar amino acid side chains. The sugar molecules are shown in blue and red licorice representations. The surfaces in panels A and B reveal the pocket where a well ordered water molecule (red sphere) that is conserved in the cleft A binding site of CBM6s and bridges multiple interactions between the ligand and protein.

Fig. 7. Overlap of cleft A region of BhCBM6 in complex with laminarihexaose (blue), CmCBM6 from C. mixtus LicA in complex with cellobiose (beige, Ref. 12), and CsCBM6–1 from C. stercorarium in complex with xylotetraose (green, Ref. 35). Relevant residues are labeled as follows: (a) Trp39, Trp36, and Trp34; (b) Trp56, Tyr35, and Tyr36; (c) Gln29, Gln39, and Ile37 in BhCBM6, CmCBM6, and CsCBM6–1, respectively. The label d, which indicates Tyr128 in BhCBM6, also shows the loop comprising residues 124–129 that is discussed in the text.
of the protein and sugars attached to O-4 extend into solvent and thus do not interact with the protein (Fig. 7). The topology of cleft A in BhCBM6 is most similar to the C. mixtus CBM (Fig. 7). A glutamine residue, which at this position is typically an isoleucine or phenylalanine in xylan binding CBM6s, blocks off one end of the binding site in both the C. mixtus CBM6 and BhCBM6. However, while in the C. mixtus CBM this residue hydrogen bonds with either terminus of the sugar, in BhCBM6 it interacts specifically with the non-reducing sugar. The unique feature of the BhCBM6 binding site is an extended loop comprising residues 124–129. This loop, and most notably the side chain of Tyr^{128} creates a raised platform that follows the U-shaped curvature of the laminarioligosaccharide up and out of cleft A and along a surface that is distinct from the usual Type B CBMs for laminarin is conferred by a binding surface that is not only different from cleft A and cleft B, but has a convex shape, while the ligand binding site of all other Type B CBMs conform to concave clefts. These data therefore reveal a third ligand binding site in the CBM6 family of proteins that exhibits a unique topology. These data demonstrate how the multiple binding clefts and highly unusual protein surface of CBM6s confers the extensive range of specificities displayed by this protein family. This is in sharp contrast to other families of CBMs where variation in specificity between different members is conferred by differences in the topology of a single binding site, while the range of ligand recognition observed in CBM6 is the result of variation in the location of the ligand binding site in different members of this family.

Conclusions—The biological rationale for the targeting of BhCBM6 to the non-reducing ends of β-1,3-glucan chains is intriguing and rather counterintuitive as the molar concentration of available binding sites will be considerably less than for the majority of Type B CBMs, which bind to the internal regions of polysaccharides. Similar targeting, but to the reducing end termini of plant structural polysaccharides is mediated by TmCBM9–2 from the T. maritima xylanase10A (13, 37). While it is possible that localization of the B. halodurans laminarinase to the ends of polysaccharide chains may reflect an exo-mode of action by the enzyme, the reaction products generated by its catalytic domain are consistent with a typical endo-mode of action; the enzyme releases oligosaccharides that display a range of different sizes (data not shown), whereas exo-acting glycoside hydrolases produce a single reaction product. The targeting of B. halodurans laminarinase to the ends of laminarin may reflect the complexity of the macromolecular structure that contains this polysaccharide, which, as a consequence, is recalcitrant to enzymatic attack. It is possible that disrupted regions of the plant cell wall, through either mechanical damage or the action of other enzymes, will contain a relatively large number of polysaccharide termini and be susceptible to laminarinase attack. BhCBM6, by targeting the enzyme to these susceptible regions may potentiate its catalytic activity.
Family 6 Carbohydrate Binding Modules Recognize the Non-reducing End of β-1,3-Linked Glucans by Presenting a Unique Ligand Binding Surface
Alicia Lammerts van Bueren, Carl Morland, Harry J. Gilbert and Alisdair B. Boraston

J. Biol. Chem. 2005, 280:530-537.
doi: 10.1074/jbc.M410113200 originally published online October 22, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M410113200

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

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

This article cites 34 references, 7 of which can be accessed free at
http://www.jbc.org/content/280/1/530.full.html#ref-list-1