Circular Permutation Provides an Evolutionary Link between Two Families of Calcium-dependent Carbohydrate Binding Modules**

Cedric Montanier†, James E. Flint‡, David N. Bolam§, Hefang Xie**, Ziyuan Liu†, Artur Rogowski†, David P. Weiner†, Supriya Ratnaparkhe†, Didier Nurizzo†, Shirley M. Roberts**, Johan P. Turkenburg**, Gideon J. Davies**‡, and Harry J. Gilbert**©

From the †Institute for Cell and Molecular Biosciences, The Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom, ‡Verenium Corp., San Diego, California 92121, the **European Synchrotron Radiation Facility, 6 rue Jules Horowitz, BP 220, F-38043 Grenoble Cedex, France, the ***York Structural Biology Laboratory, Department of Chemistry, The University of York, York YO10 5DD, United Kingdom, and the ††Complex Carbohydrate Research Center, The University of Georgia, Athens, Georgia 30602-4712

The microbial deconstruction of the plant cell wall is a critical biological process, which also provides important substrates for environmentally sustainable industries. Enzymes that hydrolyze the plant cell wall generally contain non-catalytic carbohydrate binding modules (CBMs) that contribute to plant cell wall degradation. Here we report the biochemical properties and crystal structure of a family of CBMs (CBM60) that are located in xylanases. Uniquely, the proteins display broad ligand specificity, targeting xylans, galactans, and cellulose. Some of the CBM60s display enhanced affinity for their ligands through avidity effects mediated by protein dimerization. The crystal structure of vCBM60, displays a β-sandwich with the ligand binding site comprising a broad cleft formed by the loops connecting the two β-sheets. Ligand recognition at site 1 is, exclusively, through hydrophobic interactions, whereas binding at site 2 is conferred by polar interactions between a protein-bound calcium and the O2 and O3 of the sugar. The observation, that ligand recognition at site 2 requires only a β-linked sugar that contains equatorial hydroxyls at C2 and C3, explains the broad ligand specificity displayed by vCBM60. The ligand-binding apparatus of vCBM60 displays remarkable structural conservation with a family 36 CBM (CBM36); however, the residues that contribute to carbohydrate recognition are derived from different regions of the two proteins. Three-dimensional structure-based sequence alignments reveal that CBM36 and CBM60 are related by circular permutation. The biological and evolutionary significance of the mechanism of ligand recognition displayed by family 60 CBMs is discussed.

Plant cell walls are complex macromolecular structures that consist of a diverse repertoire of interlocking polysaccharides (1). The microbial deconstruction of these composite structures, which is mediated by an extensive repertoire of hydrolytic enzymes, is of considerable biological and industrial significance. The release of sugars from the plant cell wall is not only required to maintain microbial ecosystems, but the volatile fatty acids generated by these microbiota are essential nutrients for higher order organisms such as mammalian herbivores (2). Within an industrial context, the microbial enzymes that catalyze this process are integral to the exploitation of lignocellulose as an environmentally sustainable substrate for the biofuel and bioprocessing industries (3, 4).

The physical complexity of plant cell walls limits the access of the hydrolytic enzymes to their target substrates. To overcome the “access problem” glycoside hydrolases, esterases, and lyases that degrade plant structural polysaccharides, in general, have a modular structure in which the catalytic module is appended to non-catalytic carbohydrate binding modules (CBMs; see Ref. 5 for review), which are grouped into sequence-based families within the CAZy database (6). The general function of CBMs is to direct the cognate catalytic modules to their target substrate within the plant cell wall, thereby increasing the efficiency of catalysis (7–9). CBMs, in addition to their family assignment, have also been defined as type A, B, and C modules reflecting their mode of binding (5). Type A CBMs bind to crystalline surfaces such as cellulose and chitin, type B modules recognize the internal regions of single glycan chains, whereas type C proteins typically recognize no more than two sugars and often target the end of glycan chains (5). The majority of CBMs display a β-sandwich fold with the ligand binding site located in either the concave surface presented by one of the β-sheets (a topography that facilitates the targeting of the internal regions of glycan chains (10–16)) or in the loops that connect the two sheets (17–21). This latter binding site can

**This work was supported by grants from the Biotechnology and Biological Sciences Research Council. This work was also supported in part by a grant from the National Science Foundation Plant Genome Program (IOS-0923992) and the Department of Energy (DE-FG02-09ER16076).

The atomic coordinates and structure factors (codes 2xhh, 2xjx, 2xfd, and 2xfe) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The on-line version of this article (available at http://www.jbc.org) contains supplemental Table 1S.

1 Recipient of a fellowship from Newcastle University.
2 Holds a Royal Society Wolfson Research Merit Award.
3 To whom correspondence should be addressed. Tel.: 706-583-0655; Fax: 706-542-4412; E-mail: h.j.gilbert@ccrc.uga.edu.

4 The abbreviations used are: CBM, carbohydrate binding module; CBM36, carbohydrate binding module family 36; CBM60, carbohydrate binding module family 60; vCBM60, a CBM60 member derived from an uncultured bacterium; CjXyn11A, xylanase 11A from C. japonicus; CjXyn108, xylanase 11A from C. japonicus; ITC, isothermal titration calorimetry.

31742 JOURNAL OF BIOLOGICAL CHEMISTRY

© 2010 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
either target the end (17, 19, 21) or, less frequently, the internal regions of glycan chains (18, 20).

Xylan, the major hemicellulose component of most plant cell walls, is one of the most complex carbohydrates targeted by CBMs. The polymer consists of a β-1,4-linked xylopyranose backbone that can be decorated at O2 with 4-O-methyl-α-D-glucuronic acid, or α-D-glucuronic acid, and at O2 and/or O3 with acetate and α-1-arabinofuranose moieties, which can also be linked, through O5, to ferulic acid (1). The extent and nature of these decorations are species-, tissue-, and differentiation-specific (22). CBMs that target the xylan backbone in vitro belong to families 2 (2b subfamily), 4, 6, 15, 22, and 36 (13, 15, 16, 18, 20). The deep clefts presented by CBM4, 6, and 22, however, restrict the capacity of these modules to bind to xylans within plant cell walls, whereas the more open ligand binding sites of CBM2b and 15 enable these proteins to recognize their target ligands in planta (23). The bacterium Cellvibrio japonicus expresses an extensive xylan-degrading system comprising four glycoside hydrolase family (GH)10 and two GH11 endoxylanases, a GH5 enzyme predicted to be a glucuronoxylan-specific xylanase, a GH51 general-acting, and a GH62 xylan-specific, arabinoferaniosidase, a GH67 α-glucuronidase and numerous CE1, CE2, and CE4 xylan acetyl esterases, one of which is appended to a GH11 xylanase (24–31). All the fully secreted xylan-degrading enzymes (not those appended to the outer membrane) contain at least one cellulose-binding CBM and an additional non-catalytic module of unknown function (NCM). Three of the NCMs were recently shown to be CBM35s that target both uronic acids that decorate xylans from rapidly growing cells, and a product released by the cleavage of pectin by pectate lyases (21). Intriguingly, none of these secreted enzymes appear to contain CBMs that target the xylan backbone, although the two GH11 xylanases, CjXyn11A and CjXyn11B, contain highly related NCMs (26, 30) (see Fig. 1).

To test the hypothesis that the NCMs in the two GH11 xylanases comprise novel CBMs that target xylan, the biochemical properties of the module from CjXyn11A were explored. The data show that the NCM is indeed a CBM that targets xylan but is also able to bind cellulose and galactan. Increased affinity for its ligands was conferred by avidity effects caused by protein dimerization through an inter-chain disulfide bond. The three-dimensional structure of a close homolog of the CjXyn11A xylan-binding CBM shows that ligand recognition is primarily conferred through the polar interactions of O2 and O3 of a single sugar with a protein-bound calcium ion. Furthermore, we show that CBM60 modules are evolutionarily related to CBM36 domains through circular permutation in the β-barrel folds. The functional and evolutionary significance for the mechanism of ligand recognition displayed by this CBM is discussed.

**EXPERIMENTAL PROCEDURES**

**Sources of Carbohydrates**—Wheat arabinoxylan (medium viscosity; Ara:Xyl = 41:59), rye arabinoxylan (Ara:Xyl = 49:51), carob galactomannan (high viscosity; Gal:Man = 21:79), xyloglucan (Tamarind, high viscosity; Ara:Gal:Xyl:Glc = 3:16:36:45), konjac glucomannan (high viscosity; Glc:Man = 40:60, acetylated), β-glucan (barley; high viscosity), debranched arabinoxylan (1,5-α-arabinan; Ara:Gal:Rha:GalUA = 88:4:2:6), pectic galactan (potato; Gal:Ara:Rha:GalUA = 78:9:4:9), and rhamnogalacturan (from soybean pectin), xylooligosaccharides were from Megazyme International Ireland Ltd. Carboxymethyl cellulose (medium viscosity), oat spelt xylan (Ara:Xyl = 10:90), monosaccharides, and methyl monosaccharides were from Sigma. Hydroxyethyl cellulose was from Aldrich. Cello-oligosaccharides were from Seikagaku. Bacterial microcrystalline cellulose and acid swollen cellulose were prepared essentially as described by Gilkes et al. (32) and Abou-Hachem et al. (33), respectively. Soluble and insoluble fractions of oat spelt xylan were prepared as described previously (34). To prepare galactobiase and galactotriose, pectic galactan was digested with the C. japonicus galactanase CjGal53A such that di-, tri-, and tetrasaccharides were the major products. The mixture was subjected to size exclusion chromatography using two BioGel P2 columns run in series as described previously (35). Fractions were analyzed by HPLC and those containing only galactobiase or galactotriose were pooled and freeze dried.

**Cloning, Expression, and Purification of CjCBM60A and vCBM60**—The ORF encoding CjCBM60A (UniprotKB accession no. B3P1IN1) and vCBM60 (EMBL accession no. FN908918) were amplified from C. japonicus genome DNA and plasmid pBD7340, respectively, by PCR using forward and reverse primers listed in supplemental Table 1S. The amplified DNA derived from the C. japonicus genome and pBD7340 were cloned into NdeI/XhoI-restricted pET16b (generates pCjCBM60A) and pET22a (generates pVCBM60), respectively. CjCBM60A and vCBM60 contain an N-terminal His8 tag and a C-terminal His10 tag, respectively. pVCM60-VCM60 encodes two copies of vCBM60, separated by the linker sequence KLSVSSSSSSQSSSSSSEF, and contains a C-terminal His10 tag. To generate this plasmid a polylinker sequence (comprising the restriction sites NdeI/BamHI/KpnI/HindIII/EcoRI/SacI/SalI/XhoI) was inserted into the NdeI and XhoI sites of pET22b (Novagen) to generate pFV1. The DNA sequence encoding the serine-rich linker sequence was cloned between the HindIII and EcoRI sites of pFV1 and PCR products encoding the two copies of vCBM60 were inserted, respectively, between the NdeI and HindIII and between the EcoRI and XhoI sites. Escherichia coli BL21 (DE3) (Novagen) cells, harboring pCjCBM60A, were cultured in LB broth containing ampicillin (50 µg/ml) at 30 °C to mid-exponential phase (A490 of 0.6), at which point isopropyl β-D-thiogalactopyranoside was added to a final concentration of 1 mM, and the cultures were incubated for a further 4 h. To produce vCBM60 and vCBM60-vCBM60, E. coli Origami B:pLysS cells containing pVCB60 were cultured to mid-exponential phase at 37 °C, and recombinant gene expression was induced by the addition of isopropyl β-D-thiogalactopyranoside to 0.5 mM and incubation at 30 °C for 4 h. The cells were harvested by centrifugation, and His-tagged recombinant protein was purified from cell-free extracts by immobilized metal ion affinity chromatography using a cobalt-based Talon (Clontech) column deploying standard methodology (36). For biochemical studies vCBM60 was further purified by anion-exchange chromatography using a Q12 anion-exchange column (Bio-Rad) and a 0–500 mM NaCl gradient in 10 mM Tris-HCl buffer, pH 8.0. For crystallographic studies
**Convergent Xylan Recognition**

vCBM60 was purified further by size-exclusion chromatography using an Amersham Biosciences XK16 HiLoad™ 16/60 Superdex™ Prep grade gel-filtration column. The same protocol was used to produce selenomethionine-containing vCBM60, except that the gene was expressed in *E. coli* B834 (Novagen) using growth conditions as described before (13). No reducing agent was included in the buffers used to purify the protein. To generate vCBM60nt, which contains no His tag (for crystallization with ligands), a stop codon was inserted in place of the Leu in the C-terminal motif LEHHHHHH (encoded by pET22a). After producing the protein in Origami B, as described above, 12.6 ml of the resultant cell-free extract was mixed with 1 g of finely sonicated insoluble oat spelt xylan and incubated overnight at 4 °C overnight. After washing the proteins described previously using immunohistochemical methods binding of proteins to tobacco stem sections was determined as 5.5–6.0 at a 1:1 or 3:2 ratio of protein to mother liquor. Selenomethionine-containing protein was crystallized under the same conditions but did not crystallize unless the reducing agents β-mercaptoethanol or DTT, normally included in selenomethionine preparations, were explicitly left out of the buffers during protein preparation and crystallization. Single crystals of all proteins were mounted, using Paratone oil as the cryoprotectant, and flash frozen into liquid N2.

The above crystals were mounted on beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble. Data were collected to 1.6 Å using an Area Detectors Systems Corporation Quantum 315 charge-coupled device detector at a wavelength of 1.0716 Å. A total of 360 images of 0.5° each were collected. The program MOSFLM (CCP4 suite (38)) was used to index and integrate the diffraction images. The resulting data were scaled and merged using SCALA (also CCP4 suite). The space group was either P4₁2₁2₁ or P4₄2₂₂ with cell dimensions a = b = 45.4 Å, c = 102.3 Å and one molecule per asymmetric unit. Data statistics are given in Table 1. Data for the selenomethionine form of vCBM60 were collected on beamline ID23-1 at the ESRF, Grenoble, at a wavelength of 0.9796 Å. A total of 180 images of 0.5° each were collected and processed using HKL2000 (39). Selenium positions were determined using SHELXD (recently reviewed in Ref. 40), and initial phase calculations and solvent modification in SHELXE (40) indicated the space group to be P4₁2₁2₁. The resulting phases and the native data were used in model building using the CCP4 implementation of REFMAC (41)/ARP-WARP. Minor manual corrections were performed using the program COOT (42) with maximum likelihood refinement using REFMAC (41). Final data and structure quality statistics for both native and selenomethionine forms of vCBM60 are given in Table 1. Structural figures were prepared with PyMOL.5

Crystals of vCBM60 in complex with cellotriose and galactobiose were obtained by preparing protein:ligand mixtures at 20 mg/ml vCBM60: 10 mM ligand and mixing 1:1 with 2.4 M sodium malonate or 0.01M zinc chloride, 0.1M Tris/HCl, pH 8.0, 20% (w/v) PEG 6000, respectively, and grown at 20 °C. Crystals were mounted in mother liquor supplemented with 25% glycerol and 10 mM appropriate ligand before being cryocooled in liquid N2. Data for these crystals were collected on Diamond IO3 beamline to 1.2 Å for vCBM60-cellotriose and 1.8 Å for vCBM60-galactobiose. Data for the liganded complexes of vCBM60 were integrated and scaled using MOSES and SCALA. The structures of vCBM60-cellotriose and vCBM60-galactobiose were determined by molecular replacement in the CCP4 version of MOLREP using the unliganded vCBM60 as the search model. The starting models for vCBM60-cellotriose and vCBM60-galactobiose were refined by rounds of manual rebuilding in COOT (42) interspersed with restrained refinement in REFMAC (41). Solvent water molecules were added using COOT (42) and checked manually.

5 DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA.
TABLE 1
X-ray data and structure quality statistics for vCBM60 and its complexes

|               | vCBM60 native | rCBM60 SeMet | rCBM60 cellobiose | rCBM60 galactobiose |
|---------------|--------------|-------------|------------------|-------------------|
| Resolution of data (outer shell), Å | 102.3-1.60 (1.09-1.60) | 34.1-1.74 (1.8-1.74) | 40.04-1.19 (1.22-1.19) | 36.71-1.82 (1.87-1.82) |
| Rmerge (outer shell) | 0.07 (0.23) | 0.06 (0.18) | 0.06 (0.14) |
| Rmerge (outer shell) | 29.5 (7.8) | 34.8 (12.3) | 12.9 (6.7) |
| Completeness (outer shell), % | 100.0 (100.0) | 99.9 (99.9) | 98.4 (98.8) |
| Rmerge/REMSD (%) | 12.6/18.6 | 13.0/21.0 | 16.4/19.7 |
| Root mean square deviation, 1–2 bonds (Å) | 0.012 | 0.015 | 0.029 |
| Root mean square deviation, 1–3 bonds (Å) | 1.44 | 1.55 | 2.3 |
| Root mean square deviation, chiral volume (Å³) | 0.07 | 0.08 | 0.15 |
| Root mean square deviation, 1–2 bonds (Å) | 0.012 | 0.015 | 0.029 |
| Root mean square deviation, 1–3 bonds (Å) | 1.44 | 1.55 | 2.3 |
| Root mean square deviation, chiral volume (Å³) | 0.07 | 0.08 | 0.15 |

<sup>*</sup> SeMet, selenomethionine.

RESULTS

Biochemical Properties of CjCBM60A—The two C. japonicus GH11 xylanases, CjXyn11A and CjXyn11B, contain a highly similar 116 residue module of unknown function (26, 30). To explore their functional significance, the CjXyn11A module was expressed in E. coli (attempts to produce the corresponding CjXyn11B module in soluble form were unsuccessful) and, after purifying the protein to electrophoretic homogeneity by IMAC, its biochemical properties were investigated. The module displayed no hydrolytic activity against plant cell wall polysaccharides, including xylans, galactans, mannans, and β-glucans (data not shown). Affinity gel electrophoresis, however, showed that the protein bound to a range of xylans, carboxymethylcellulose, hydroxyethylcellulose, lichenan, barley β-glucan, and lichenan (and β-1,4-galactan (barley β-glucan and lichenan) and β-1,4-galactan (example data in Fig. 1, full data set in Table 2) but did not bind to galactomannan, β-1,3-glucan (laminarin), or a range of α-linked polysaccharides. It is apparent, therefore, that the protein module is a CBM that displays particularly broad ligand specificity and, henceforth, is designated as CjCBM60A, being the founding member of this new family.

ITC was used to quantify the affinity of CjCBM60A for polysaccharide and oligosaccharide ligands. Examples of the titrations are displayed in Fig. 2, whereas the complete data set are reported in Table 3. The data show that CjCBM60A binds tightly to a range of β-1,4-galactans, xylans, and β-glucans (β-1,4-glucan backbone, xyloglucan; mixed-linked β-1,3-β-1,4-glucans, barley, and lichenan). The stoichiometry of binding, assuming a single binding site for each CBM protomer, indicated that, at saturation, each protein molecule occupied ~4–6 tandemly arrayed sugar residues for all polysaccharides. In general the affinity and the coverage, irrespective of the number and type of side chains decorating the backbone, were similar. Thus the ligand binding site of CjCBM60A appears to be able to interact with xylans, glucans, and galactans that are extensively decorated. Typical of carbohydrate-protein recognition (7, 10, 13, 16, 19–21), binding to polysaccharides and oligosaccharides was driven by favorable changes in enthalphy, with generally an unfavorable entropic contribution. It should be noted, however, that the binding of CjCBM60A to glucans is associated with a positive change in entropy, which therefore makes a favorable contribution to affinity (Table 3).

The affinity (K_a) of CjCBM60A for xylohexaose, which appears to fully occupy the ligand binding site, was 2.7 × 10⁴ M⁻¹ (data not shown), which was 20- to 30-fold lower than the K_a values obtained for the various xylans (Table 3). The targeting of multivalent polysaccharides, in preference to oligosaccharides, is likely to be mediated through avidity effects (43, 44), which requires that CjCBM60A is capable of oligomerization. This view is supported by size-exclusion studies, which showed that the protein migrated as, predominantly, a dimeric species (data not shown), the biological significance of which is discussed below.
Broad Ligand Specificity Displayed by a Second CBM60 Module—To evaluate whether the broad ligand specificity, observed for CjCBM60A, is a generic feature of CBM60, attempts were made to express other proteins from this family. Out of six homologs of CjCBM60A, provided by Verenium Corp. (San Diego, CA) from their in-house environmental library of GH11 xylanases, only one protein module, designated vCBM60, could be produced in soluble form in E. coli. vCBM60 displays ~55% sequence identity to CjCBM60A, the main difference being a C-terminal extension of ~10 amino acids in the Cellvibrio module. The ligand specificity of vCBM60 (examples of which are displayed in Fig. 1, and the full data set is in Tables 2 and 3) is very similar to CjCBM60A. The protein module binds to galactan, xylans, and β-glucans (that contain β-1,4

FIGURE 2. Representative ITC data of the CBM60s titrated with carbohydrates and calcium. The ligand (0.5–5 mg/ml for the polysaccharide and 3 mM for calcium) in the syringe was titrated into the CBM60 (50–100 μM) in the cell. The galactan was derived from potato, and WAX signifies wheat arabinoxylan. The top half of each titration shows the raw injection heats; the bottom half, the integrated peak areas fitted using a single site model (MicroCal Origin v7.0). ITC was carried out as described under “Experimental Procedures” in 50 mM Na/Hepes, pH 7.5, containing 3 mM CaCl₂, at 25 °C for the carbohydrate ligands. When titrating with calcium the proteins were treated with Chelex-100 to generate the apo forms and calcium was omitted from the Hepes buffer. 

31746 JOURNAL OF BIOLOGICAL CHEMISTRY VOLUME 285 • NUMBER 41 • OCTOBER 8, 2010
TABLE 3
ITC of the binding of CBM60s to carbohydrate ligands
ITC was carried out in 50 mM Na Hepes buffer, pH 7.5, containing 5 mM CaCl$_2$ at 25°C.

| Ligand                          | $K_a \times 10^4$ | $\Delta G$ | $\Delta H$ | $\Delta S$ | $n^a$ |
|--------------------------------|-------------------|------------|------------|------------|-------|
| Soluble oat-spelt xylan        | 7.2 ± 0.08        | 5.3        | -6.6 ± 1.0 | -1.2       | 1.09 ± 0.15 |
| Birchwood xylan                | 5.9 ± 0.1         | 9.2 ± 0.1  | -8.4 ± 1.4 | -4.1       | 1.05 ± 0.0 |
| Wheat arabinoxylan             | 4.3 ± 0.1         | 9.2 ± 0.1  | -8.4 ± 9.0 | -3.5       | 1.10 ± 0.11 |
| Rye arabinoxylan               | 3.8 ± 0.2         | 8.9 ± 0.9  | -8.4 ± 1.4 | -4.2       | 1.04 ± 0.18 |
| Hydroxyethylcellulose          | 6.2 ± 0.9         | 7.5 ± 1.2  | -8.4 ± 1.4 | -3.0       | 1.01 ± 0.11 |
| Xyloglucan                     | 5.5 ± 0.1         | 8.7 ± 0.9  | -8.4 ± 1.0 | -2.5       | 0.85 ± 0.11 |
| Lupin galactan                 | 7.7 ± 1.1         | 8.1 ± 0.4  | -6.9 ± 0.4 | -5.4       | 1.02 ± 0.32 |
| Potato pectic galactan         | 8.4 ± 0.7         | 8.1 ± 0.4  | -7.9 ± 0.8 | -4.6       | 1.05 ± 0.07 |
| Rhamnogalacturonan             | n.b.$^b$          | n.b.       | n.b.       | n.b.       | n.b.  |
| Xylohexose                     | 7.6 ± 0.2         | 5.3        | -7.1 ± 0.9 | -1.7       | 1.01 ± 0.01 |
| Xyloptentaose                  | 1.0 ± 0.8         | 8.5 ± 0.5  | -6.6 ± 0.3 | -1.1       | 1.05 ± 0.07 |
| Xylotriose                     | 5.2 ± 1.2         | 5.7 ± 1.2  | -5.7 ± 0.3 | 0.0        | 1.10 ± 0.06 |
| Xystroside                     | 1.0 ± 0.2         | 5.4        | -5.6 ± 0.3 | -0.2       | 0.96 ± 0.07 |
| Cellohexose                    | 3.2 ± 0.4         | 4.7        | -8.2 ± 2.9 | -3.4       | 1.00 ± 0.32 |
| Celloctoate                    | 5.3 ± 0.1         | 5.1        | -5.3 ± 0.1 | -0.2       | 0.99 ± 0.18 |
| Celloptentaose                 | 6.0 ± 0.2         | 4.8        | -5.3 ± 0.4 | -0.5       | 0.95 ± 0.08 |
| Cellotriose                    | 4.1 ± 0.5         | 4.9        | -6.1 ± 0.12| -1.2       | 1.03 ± 0.36 |
| Galactotriose                  | 13.4 ± 1.3        | 5.6        | -10.5 ± 0.5| -4.9       | 1.00 ± 0.04 |
| Galactobiose                   | 12.7 ± 0.6        | 5.6        | -10.3 ± 0.4| -4.7       | 1.01 ± 0.03 |

$^a$ The ITC data were fitted to a single site binding model for all ligands. For polysaccharide ligands in which the molar concentration of binding sites is unknown, the $n$ value was iteratively fitted to as close as possible to one, by adjusting the molar concentration of the ligand (see Ref. 11).

$^b$ The errors are presented, for the experimentally determined values for typical titrations.

$^c$ n.b., no binding.

Convergent Xylan Recognition

linkages), although the actual affinities for these polysaccharides are considerably lower than observed with CjCBM60A. This likely reflects the inability of vCBM60 to oligomerize (data not shown), precluding any benefits derived from avidity effects. Indeed, this view is supported by the observation that the artificial construct, containing two tandem copies of vCBM60 (designated vCBM60-vCBM60), binds to xylan and galactan, respectively, 20- and 40-fold more tightly than the monomeric form of the protein (Table 4). The affinity of vCBM60 for a series of oligosaccharides (Table 3) shows that the protein exhibits similar affinity for the hexasaccharide and trisaccharide of cello and xylooligosaccharides (but not xylose or cellobiose), displaying a preference for the xylo-configured ligands, whereas the $K_a$ values for galactobiase and galactotriose were very similar. It would appear, therefore, that the ligand binding site is optimized to bind trisaccharides of glycan that have a 2- and 3-fold screw axis ($\beta$-1,4-containing glucans and xyans), but to disaccharides of glycans that adopt an extended helical conformation, such as $\beta$-1,4-galactan. These data suggest that the broad specificity displayed by vCBM60 and CjCBM60A may be a general feature of this family, particularly as the ligand binding residues (discussed in detail below) are highly conserved.

vCBM60 Displays Absolute Requirement for Calcium—To investigate whether ligand recognition was metal-dependent,
Convergent Xylan Recognition

The influence of calcium on ligand recognition by CBM60s

| Protein          | Ligand                  | $K_a \times 10^3$ |
|------------------|-------------------------|-------------------|
| vCBM60           | Oat-spelt xylan         | 7.2 ± 0.8         |
| vCBM60 + 10 mM EDTA | Oat-spelt xylan      | n.b.              |
| vCBM60 (apo form) | Oat-spelt xylan      | n.b.              |
| vCBM60 (apo form) | Oat-spelt xylan      | 7.5 ± 0.6         |
| vCBM60 (apo form) | Calcium                 | 212 ± 6.9         |
| CjCBM60A         | Oat-spelt xylan         | 400 ± 40          |
| CjCBM60A + 10 mM EDTA | Oat-spelt xylan      | n.b.              |
| CjCBM60A (apo-form) | Oat-spelt xylan      | n.b.              |
| CjCBM60A (apo-form) | Oat-spelt xylan      | 312 ± 42          |
| CjCBM60A (apo-form) | Calcium                 | 173 ± 18          |

a n.b., no binding.

b CBM60 (apo form): the apo form of the CBM60s were prepared by treatment with Chelex-100 to remove bound divalent metal ion. Titration was performed in 50 mM Na Hepes, pH 8.0, treated with Chelex-100, with 0.5 or 5 mg/ml oat-spelt xylan.

c CBM60 (apo form): titration of the apo-form of the CBM60s with xylan was carried out in 50 mM Na Hepes, pH 7.5, containing 5 mM CaCl₂.

d For calcium, the concentration of calcium in the syringe was 0.5–3 mM.

In Vitro Labeling—The capacity of the tandem vCBM60-vCBM60 protein to bind to ligands in intact cell walls of tobacco stem sections was assessed using immunohistochemistry (Fig. 3). This analysis indicated that the protein bound extensively to secondary cell walls of the xylem elements and phloem fibers. Although only weak recognition of the primary cell walls was observed in cortical parenchyma, pith parenchyma and epidermal tissues (Fig. 3), this likely reflects the thinness of the walls. The thickened vascular secondary walls contain a large amount of xylan, as well as cellulose. In primary cell walls the cellulose microfibrils are embedded in a matrix consisting mainly of pectins that include galactans, albeit at low concentrations (45). The binding profile of vCBM60-vCBM60 to tobacco stems is broadly similar to other xylan-specific CBMs (such as CBM2b-1–2 (23)), although CBM2b-1–2 did not display any binding to primary cell walls. The lack of cellulose recognition (CBMs that bind crystalline cellulose bind extensively to both secondary and primary cell walls (46)) likely reflects the extensive inter- and intra-chain hydrogen bonds between the hydroxyl groups within the microfibrils, which are therefore unable to make polar contacts with vCBM60. The weak binding to primary cells also indicates that either galactan is present at low concentrations in these walls or is not accessible to vCBM60. Thus, although vCBM60 appears to display broad ligand specificity in vitro, in an ex vivo setting this module targets mainly cell walls rich in xylan, consistent with the location of family 60 CBMs predominantly in xylanases (see below).

Crystal Structure of vCBM60—The crystal structure of vCBM60 was solved using the single-wavelength anomalous dispersion method employing selenomethionine-labeled protein and refined using data extending to 1.6-Å resolution (Table 1). The vCBM60 crystals belong to a space group of P4₁2₁2 with cell dimensions $a = b = 45.4$ Å, $c = 102.3$ Å and one molecule per asymmetric unit. Overall, the electron density is of high quality throughout the entire length of the protein, and the final model contains 110 amino acid residues, 175 water molecules, and 2 ions that, based on their coordination sphere, ligands, and $B$-factors have been modeled as Ca$^{2+}$. The structure generally has backbone torsion angles in the most favored region (90.6% of the bonds), and none in the disallowed regions of the Ramachandran plot, according to PROCHECK (47).
vCBM60 adopts a \(\beta\)-sandwich fold in which the two \(\beta\)-sheets, each containing four anti-parallel \(\beta\)-strands, are connected entirely by loops (Fig. 4). The first \(\beta\)-sheet (\(\beta\)-sheet 1) includes \(\beta\)-strands \(\beta\)-4 (Gln\(^{36}\)–Ala\(^{46}\)), \(\beta\)-1 (Ile\(^{4}\)–Gly\(^{19}\)), \(\beta\)-6 (Val\(^{61}\)–Val\(^{68}\)), and \(\beta\)-7 (Gln\(^{71}\)–Gln\(^{78}\)), whereas the second \(\beta\)-sheet (\(\beta\)-sheet 2) consists of strands \(\beta\)-3 (Thr\(^{24}\)–Thr\(^{31}\)), \(\beta\)-2 (Ser\(^{16}\)–Val\(^{21}\)), \(\beta\)-5 (Glu\(^{47}\)–Phe\(^{52}\)), and \(\beta\)-8 (Glu\(^{103}\)–Gly\(^{108}\)). The hydrophobic core of the protein comprises 3 leucine, 3 isoleucine, 2 phenylalanine, 1 tryptophan, and 8 valine residues. The calcium ion (Ca\(_1\)) in vCBM60, which connects \(\beta\)-6 and \(\beta\)-7 with the loop linking \(\beta\)-7 and \(\beta\)-8, is highly conserved in CBMs that display a \(\beta\)-sandwich fold (5). The metal ion makes coordinate bonds with Asp\(^{64}\) O and O\(_3\), and also with Trp\(^{85}\), mainly with the pyrrole component of the aromatic ring system, but does not make any direct polar contacts with the protein. Whereas Gal2 also makes hydrophobic contacts with Trp\(^{85}\), primarily the benzene component of the indole ring, the O2 and O3 of the sugar make extensive interactions with Ca\(_2\), and with the aspartates that form coordinate bonds with the divalent metal ion (Fig. 5). Thus, Ca\(_2\) interacts with the ligand binding site by making coordinate bonds with Arg\(^{59}\) O, His\(^{109}\) O, Asp\(^{55}\) O\(_{\text{H}}\), and Asp\(^{60}\) O\(_{\text{H}}\), while its octahedral coordination is completed through interactions with O2 and O3 of the divalent metal ion (Fig. 5). Electron density for both oligosaccharides were solved to resolutions of 1.2 and 1.8 Å, respectively. Bound ligand is shown in light blue, and the ligand-binding calcium is shown as a magenta sphere. The disulfide bond stabilizing the extended loop between \(\beta\)-7 and \(\beta\)-8, while also contributing to the floor of the ligand binding site (Fig. 4).

The Ligand Binding Site in vCBM60—The ligand binding sites in CBMs that display a \(\beta\)-sandwich fold comprise, typically, the concave surface presented by one of the \(\beta\)-sheets, or at the end of the elliptical protein, within the loops connecting these two structural elements (see Ref. 5 for review). Inspection of the concave surface of vCBM60 does not reveal a cleft-like structure that would accommodate polysaccharide ligands. By contrast, the loops connecting the two \(\beta\)-sheets present a very broad but short cleft (Fig. 4). The floor and one wall of the cleft are formed by the loop connecting \(\beta\)-7 with \(\beta\)-8, whereas the other wall comprises the loop linking \(\beta\)-5 with \(\beta\)-6. Confirmation that the loop-derived cleft consists of the ligand binding site is derived from mutagenesis studies. Substitution of most of the residues on the surface of this cleft with alanine either completely abrogates, or greatly reduces, the affinity of the protein for its carbohydrate ligands (Table 4). Indeed, the observation that these mutations cause a similar reduction in affinity for the gluco-, xylo-, and galacto-configured ligands demonstrates that the protein interacts with these polymers at a common binding site.

The crystal structure of vCBM60 in complex with cellotriose and galactobiose was solved to resolutions of 1.2 and 1.8 Å, respectively. Electron density for both oligosaccharides were evident, and thus the mechanism for the broad ligand specificity displayed by vCBM60 could be explored (Fig. 5). With respect to galactobiose the disaccharide is twisted into a helical structure that would accommodate polysaccharide ligands. By contrast, the loops connecting the two \(\beta\)-sheets present a very broad but short cleft (Fig. 4). The floor and one wall of the cleft are formed by the loop connecting \(\beta\)-7 with \(\beta\)-8, whereas the other wall comprises the loop linking \(\beta\)-5 with \(\beta\)-6. Confirmation that the loop-derived cleft consists of the ligand binding site is derived from mutagenesis studies. Substitution of most of the residues on the surface of this cleft with alanine either completely abrogates, or greatly reduces, the affinity of the protein for its carbohydrate ligands (Table 4). Indeed, the observation that these mutations cause a similar reduction in affinity for the gluco-, xylo-, and galacto-configured ligands demonstrates that the protein interacts with these polymers at a common binding site.

The crystal structure of vCBM60 in complex with cellotriose and galactobiose was solved to resolutions of 1.2 and 1.8 Å, respectively. Electron density for both oligosaccharides were evident, and thus the mechanism for the broad ligand specificity displayed by vCBM60 could be explored (Fig. 5). With respect to galactobiose the disaccharide is twisted into a helical structure that would accommodate polysaccharide ligands. By contrast, the loops connecting the two \(\beta\)-sheets present a very broad but short cleft (Fig. 4). The floor and one wall of the cleft are formed by the loop connecting \(\beta\)-7 with \(\beta\)-8, whereas the other wall comprises the loop linking \(\beta\)-5 with \(\beta\)-6. Confirmation that the loop-derived cleft consists of the ligand binding site is derived from mutagenesis studies. Substitution of most of the residues on the surface of this cleft with alanine either completely abrogates, or greatly reduces, the affinity of the protein for its carbohydrate ligands (Table 4). Indeed, the observation that these mutations cause a similar reduction in affinity for the gluco-, xylo-, and galacto-configured ligands demonstrates that the protein interacts with these polymers at a common binding site.

The crystal structure of vCBM60 in complex with cellotriose and galactobiose was solved to resolutions of 1.2 and 1.8 Å, respectively. Electron density for both oligosaccharides were evident, and thus the mechanism for the broad ligand specificity displayed by vCBM60 could be explored (Fig. 5). With respect to galactobiose the disaccharide is twisted into a helical structure that would accommodate polysaccharide ligands. By contrast, the loops connecting the two \(\beta\)-sheets present a very broad but short cleft (Fig. 4). The floor and one wall of the cleft are formed by the loop connecting \(\beta\)-7 with \(\beta\)-8, whereas the other wall comprises the loop linking \(\beta\)-5 with \(\beta\)-6. Confirmation that the loop-derived cleft consists of the ligand binding site is derived from mutagenesis studies. Substitution of most of the residues on the surface of this cleft with alanine either completely abrogates, or greatly reduces, the affinity of the protein for its carbohydrate ligands (Table 4). Indeed, the observation that these mutations cause a similar reduction in affinity for the gluco-, xylo-, and galacto-configured ligands demonstrates that the protein interacts with these polymers at a common binding site.
Convergent Xylan Recognition

The ligand binding site of vCBM60. Top shows the three-dimensional position of amino acids of vCBM60 that contribute to ligand recognition either directly, or through calcium (slate blue spheres). The amino acids (carbons in green) and carbohydrate ligands (carbon in silver) are shown. The maximum likelihood/orientation/weighted 2Fobs − Fcalc electron density map for the carbohydrate ligands is shown in cyan mesh and contoured at 1σ (0.61 e−/Å3). Only the backbone carbonyl of Arg95 and His100 is shown. Bottom is a schematic of A. The dotted lines show polar interactions, and calcium is shown as a gray sphere.

Gal2 and a water molecule. In addition to interacting with Ca2, O2 and O3 of Gal2 make hydrogen bonds with Asp60 O62 and Asp55 O62, respectively, and also interact with the backbone carbonyl of His100. It is also possible that O2 and O3 make indirect contact with the CBM through the water that interacts with Ca2.

The structure of the vCBM60-cellobiore complex (Figs. 4 and 5), reveals only two of the three Glc residues. The disaccharide adopts a 2-fold screw axis in which the two Glc residues (defined as cellobiose) are orientated at 180° with respect to each other (galactobiose adopts a more helical conformation). The interactions between vCBM60 and cellobiose are very similar to the vCBM60-galactobiose complex. Trp85 makes planar hydrophobic contacts with Glc1 and Glc2, however, the poor electron density displayed by Glc1 indicates that these apolar interactions are weak. The O2 and O3 of Glc2 make polar contacts with Ca2, the side chains of Asp55 and Asp60, the backbone carbonyl of His100 and indirect interactions through the water that coordinates with Ca2. In addition, the equatorial O4 appears to make a polar contact with Asp55 O62, which is not mirrored by the axial O4 of Gal2. We were unable to obtain a complex of vCBM60 with xylo-oligosaccharides. However, as there are no interactions with O6 of either galactobiose or cellobiose, it is highly likely that xylo-configured ligands will make similar, if not identical, interactions with the protein to cellobiose. It should be noted, however, that maximal binding to galactan is achieved through interactions with only two sugars, whereas the protein binds to tri- but not disaccharides of cellulose and xylan. The reason(s) for this subtle difference in specificity are not entirely clear. It is possible that the weak hydrophobic interaction between Glc1 (and by inference Xyl1) with Trp85 (compared with Gal1, which makes more extensive apolar contacts with the tryptophan) is maximized by presenting a closed pyranose ring to the protein, which occurs when Glc1 is the central sugar in a trisaccharide. Indeed, previous studies have shown that oligosaccharides that extend beyond the ligand binding site of CBMs display higher affinity than oligosaccharides that have a degree of polymerization that matches the number of available sugar binding sites.

Mutagenesis studies described above (Table 4), are entirely consistent with the structure of the complexes; there is a complete loss in binding in protein variants where Asp55, Asp60, or Trp85 are replaced with alanine. Furthermore the interaction of only the backbone carbonyl of His100 with calcium explains the lack of a significant effect on ligand binding when this residue is substituted with alanine (Table 4). Interestingly, Trp85 is the only amino acid that interacts with Glc/Gal1, comprising the second sugar binding site. Support for the importance of this second binding site in driving ligand recognition is provided by the observation that Gal, Glc and Xyl, or their β-methyl glycosides, do not bind to vCBM60 (data not shown). Generally, the mutagenesis data are consistent with the view that xylo-configured ligands bind through a similar mechanism to either galactobiose or cellobiose.

The role of calcium in the function of CBMs is restricted to those modules that display a β-sandwich fold, where the ligand binding site comprises the loops that link the two β-sheets (49). The metal ion, currently, only contributes to carbohydrate recognition in those CBMs that interact with one or, at most, two sugars, but not in those modules that bind to more extensive ligands, suggesting that calcium can mediate particularly tight binding to sugars. This feature of metal ion requirement is well illustrated in CBMs that bind to the internal regions of polysaccharides. CBMs that bind to four or more sugars in these polymers do not utilize calcium in ligand recognition, however, in modules, such as CBM36 (20) and CBM60, which interact with only two or three internal sugars, the metal ion plays a dominant role in carbohydrate binding, consistent with the view that charged dipole-dipole interactions are stronger than charged ones (50).
of vCBM60 (Fig. 7). Furthermore, there is perfect alignment of the amino acids that plays a key role in ligand recognition by coordinating the calcium or in making hydrophobic interactions with the second sugar. The structural and sequence similarity of non-contiguous regions of these two CBM families proteins is strongly indicative of “circular permutation.”

In circular permutation genetic (or post-translational processing in the case of concanavalin A (52)) events lead to the ligation of the N and C termini of the protein and subsequent cleavage at another site generating new N and C termini (53). There are many natural examples of circular permutation, including some carbohydrate-active proteins and enzymes, transaldolases, glutathione synthetases, methyltransferases, and proteinase inhibitors (see Ref. 54 for review). In general proteins that undergo circular permutation display either a β-barrel or β-sandwich fold and require the N and C termini to be in close proximity. Indeed, in virtually all CBMs that display a β-sandwich fold the two termini are in close proximity (5), suggesting that these proteins are candidates for circular permutation events.

The evolutionary rationale for the circular permutation of CBM60 (or an ancestral sequence of this family) leading to the generation of family 60 CBMs (or vice versa; see “Discussion”) is not readily apparent. It is remarkable that the reorganization of the β-sandwich fold did not disrupt the topology of the ligand binding site. Interestingly, similar results have been obtained through synthetic circular permutations, although the stability of these proteins was compromised (55). Intuitively, such events, which retain overall three-dimensional structure, may introduce subtle structural changes that alter the function of the protein. Indeed, several CBM60s have acquired a C-terminal extension that likely mediates oligomerization and hence increased affinity. It is possible that, in some way, the genetic events that lead to circular permutation also generate these oligomerization sequences.

**CBM Family 60—**Analysis of sequence-based relatives of vCBM60, using BLAST, revealed 15 proteins that contain a module with significant similarity to vCBM60 (including CjCBM60A and -B). All of the proteins were from bacteria, and all but one of the proteins contained xylanase catalytic modules from either GH11 (12 of 15) or GH10 (2 of 15), suggesting that the primary target for this CBM is xylan. Several proteins also contained a carbohydrate esterase from either family 4 (CE4, 5 of 15) or family 6 (CE6, 1 of 15), with both families known to contain xylan-specific esterases. In addition to the CBM60 module, 7 of 15 proteins contain a CBM10, a CBM family that targets crystalline cellulose (56), and 2 contain a CBM5 (one of these also has a CBM10), a family that targets chitin and cellulose. One protein, B3PKS4 from *C. japonicus*, has neither xylanase nor esterase catalytic modules, although, in addition to the CBM60, it does contain a CBM10 and a module of ~400 residues. It is possible that this large module displays a novel catalytic function related to xylan deconstruction. The residues that coordinate Ca2 in vCBM60 (Asp55, Arg59, Asp60, and His100) are invariant in the other CBM60 members, whereas Trp85, which comprises the second sugar binding site, is replaced with tyrosine and is thus functionally conserved within the family (Fig. 8). These data strongly suggest that the ligand specificity is
likely to be extremely similar, if not identical, across the CBM60 landscape. Affinity, however, is variable in CBM60, reflecting avidity effects in some proteins (CjCBM60A), but not in others (vCBM60). Inspection of the CjCBM60A sequence reveals a 10-residue C-terminal extension that is absent in vCBM60, which may contribute to protein dimerization. Indeed this sequence is present in 10 of 17 CBM60 members (Fig. 8), suggesting that increased affinity through avidity effects may be a common feature in this CBM family. It should also be noted that one enzyme (Q21G61), derived from the marine bacterium Saccharophagus degradans, contains a tandem copy of CBM60 that has the C-terminal extension (57). It is possible that this protein will assemble four copies of a CBM60 into a protein dimer and thus bind particularly tightly to multivalent ligands, such as xylan and galactan. Multiple CBMs are a common feature in extracellular glycoside hydrolases expressed by S. degradans, and it has been proposed that this might reflect the marine environment. Thus, in such a dilute ecosystem, if secreted enzymes are not tethered to the plant cell wall via CBMs with high affinity, they will rapidly disperse, and their benefit to the host organism will be lost (57).

**DISCUSSION**

This study reports a new CBM family, CBM60, and the structural basis for the specificity displayed by this family. CBM60 proteins appear
to display unusually broad ligand specificities, recognizing a range of β-linked polysaccharides that adopt very different conformations. In general CBMs display tight specificity for their target ligand (16, 17, 19), and thus it is particularly unusual for a module to bind xylan, β-1,4-galactan, and β-1,4-containing glucans. This broad specificity reflects the targeting of the equatorial O2 and O3 of pyranose sugars as the primary specificity determinant, which is a feature of many carbohydrates. The key geometric signature of such sugars is the O3-C3-C2-O2 torsion angle of +60°. In comparison, the O3-C3-C2-O2 torsion angle of a mannoside is −60°, which explains why CBM60 is unable to bind homopolymers of mannan. Indeed, for the notable occasions where Ca2⁺ has been observed to bind mannosides (it is a feature of GH47 and GH92 mannosidases, for example), the Ca2⁺ coordination is actually used to help distort O3 and O2 away from the normal geometry of a mannoside ring, contracting the torsion to −0° (58, 59). Furthermore, in the case of CBM60 the axial O2 will not interact with the calcium in site 1 and will make a steric clash with the tryptophan in site 2. It is likely that the biologically significant ligand recognized by CBM60 modules is xylan, because they are located predominantly in xylanases (26, 30). The capacity of this family to bind to galactan and β-glucans probably reflects the limited hydroxyls targeted by the proteins.

Although this lack of specificity may appear counterintuitive, xylans are complex carbohydrates in which the backbone β-1,4-polymer is decorated by an array of structures at the O2 and/or O3 positions (1). Thus, by targeting the polar groups of only a single xylose, coupled to the approximate location of the adjacent moiety, the protein can bind to myriad decorated xylans, because side chains are not appended to every backbone sugar in these polymers (1). Indeed, the observation that xylan esterases are often located in CBM60-containing xylanases suggests that these multimodular enzymes target acetylated xylans, reinforcing the importance of CBM60 binding to only a very limited region of the polymer. It has previously been shown that other xylan-specific CBMs have evolved to accommodate decorated xylans by targeting the pyranose rings, exposing the hydroxyls to solvent (15, 16). Nevertheless, because the ligand binding sites of these CBMs often accommodate six tandem xylose units, a significant proportion of the hydroxyls is not accessible to the solvent, and thus these modules could not accommodate ligands with a high degree of decoration (15, 16).

The discovery of the CBM60 family completes the modular assignment of the enzymes that comprise the complex xylan degrading system of C. japonicus. Thus, a subset of three enzymes (xylanase CjXyn10B, esterase CjCE1, and the xylan-specific arabinofuranosidase, CjAbf62A) contains a cellulose-specific CBM2a and a CBM35, which targets uronic acids that decorate xylans from rapidly growing cells and pectins, which are attacked by pectate lyases (21). The bacterium also synthesizes a GH10 (CjXyn10A) and a GH5 (CjXyn5A) xylanase that contain only cellulose-specific CBMs (CBM2a and CBM10 in the GH10; only CBM2a in the GH5), whereas the two GH11 xylanases both contain a CBM60 (CjXyn11A also contains a cellulose-specific CBM10 and a CE4 xylan-specific esterase) (25). Thus, cellulose appears to function as a primary receptor for the xylan-degrading apparatus of C. japonicus, first proposed by Kellett and colleagues (29). Furthermore, CBM2as, and likely other type A CBMs, are able to slide over cellulose surfaces (60). An attractive hypothesis is that the xylan-degrading enzymes are initially locked onto the plant cell wall through their cellulose binding CBMs and are then able to locate regions rich in xylans. It is possible that CjXyn5A and CjXyn10A target xylans that are in intimate contact with cellulose microfibrils, and that CjXyn11B, and its associated accessory enzymes, attack regions of xylan that are in close proximity with pectins (or are present in rapidly growing cells), while the two CBM60 modules lock the two GH11 xylanases, CjXyn11A and CjXyn11B, onto xylans, which, because of their extensive acetyl decorations, are not in intimate contact with cellulose or other polysaccharides. The removal of the acetate groups by the xylan esterase in CjXyn11A would enable the two CBM60-containing enzymes to subsequently hydrolyze the exposed xylan backbone.

To conclude, this report describes a new family of xylan-binding CBMs that, by targeting, primarily, a single backbone xylose residue, likely directs the cognate enzymes to highly decorated regions of the polysaccharide. The ligand-binding apparatus of CBM60 and CBM36 display remarkable similarity; however, the binding residues are located in very different regions of the two protein families. It is evident that families CBM60 and CBM36 arose through a circular permutation event and, remarkably, the reorganization of the jelly roll fold did not disrupt the integrity of the ligand binding site. It should be noted that the strand order in CBM36 is conserved in the other β-sandwich CBM families. It is likely, therefore, that CBM60 arose through the circular permutation of CBM36 rather than vice versa. The flexibility in xylan recognition displayed by CBM60 underscores the utility of this CBM family in directing enzymes to different substrates of this major hemicellulosic polysaccharide, illustrating its utility in the toolbox of biocatalysts required to deconstruct the plant cell wall.

REFERENCES
1. Brett, C. T., and Waldren, K. (1996) *Physiology and Biochemistry of Plant Cell Walls. Topics in Plant Functional Biology*, Chapman and Hall, London
2. Hooper, I. V., Midveldt, T., and Gordon, J. I. (2002) *Annu. Rev. Nutr.* **22**, 283–307
3. Himmel, M. E., and Bayer, E. A. (2009) *Curr. Opin. Biotechnol.* **20**, 316–317
4. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) *Biochem. J.* **378**, 769–781
5. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2007) *Nucleic Acids Res.* **35**, D117–D124
6. Boraston, A. B., Kwan, E., Chiu, P., Warren, R. A., and Kilburn, D. G. (2003) *J. Biol. Chem.* **278**, 6119–6127
7. Hall, J., Black, G. W., Ferreira, L. M., Millward-Sadler, S. J., Ali, B. R., Hazelwood, G. P., and Gilbert, H. J. (1995) *Biochem. J.* **309**, 749–756
8. Tomme, P., Van Tilburg, H., Pettersson, G., Van Damme, J., Vandenkocke, J., Knowles, J., Teeri, T., and Claeyssens, M. (2008) *Eur. J. Biochem.* **170**, 575–581
9. Boraston, A. B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D. R., Kilburn, D. G., and Davies, G. J. (2002) *J. Mol. Biol.* **319**, 1143–1156
10. Carvalho, A. L., Goyal, A., Prates, J. A., Bolam, D. N., Gilbert, H. J., Pires,
