Supporting Information

Montainer *et al*. 10.1073/pnas.0808972106

SI Materials and Methods

Cloning of CBM35s Encoding DNA. The region of the genes encoding the CBM35 were amplified by PCR from genomic DNA or, in the case of Pel-CBM35 and Chi-CBM35, from the respective cloned genes, using the thermostable DNA polymerase Pfu Turbo (Stratagene) and primers listed in Table S3. Amplified DNA encoding Chi-CBM35 was cloned into *Nhe*I and *Xho*I restricted pET28a, whereas the other PCR products were cloned into *Nde*I and *Xho*I digested pET22b. All CBM35s contain a *C*-terminal His_6-tag.

Expression and Purification of CBM35s. *Escherichia coli* BL21 (STAR) DE3 cells harboring the Xyl-CBM35, Pel-CBM35, and Rhe-CBM35-encoding recombinant expression vectors were cultured in Luria–Bertani broth at 37 °C to mid-exponential phase ($A_{600\,nm}$ 0.6). Recombinant protein expression was induced by the addition of 1 mM isopropyl 1-thio-β-D-galactopyranoside and the cells were incubated for a further 5 hours at 37 °C. The CBM35s, derived from the plant cell wall degrading enzymes, were purified from lysed cells by immobilized metal ion affinity chromatography (IMAC) using Talon resin (Clontech) and eluted in 20 mM Tris/HCl buffer, pH 8.0, containing 300 mM NaCl and 100 mM imidazole. The CBMs were dialyzed against 20 mM Tris/HCl buffer pH 8.0 (Buffer A) and applied to a Bio-Rad Q12 column. The CBM35s were eluted with a 400 ml 0–500 mM NaCl gradient in Buffer A. The fractions containing the recombinant proteins were concentrated using a 10 kDa MWCO Vivaspin 20 centrifugal concentrator and applied to a Superdex 200 26:60 Hiload column (GE Healthcare) equilibrated with 10 mM Tris-HCl buffer, pH 8.0, containing 150 mM NaCl. The proteins were eluted at a flow rate of 1 $mL{\cdot}min^{-1}$. Both chromatography steps used a Bio-Rad FPLC system. Chi-CBM35 was produced in *E. coli* using the procedure outlined for the other CBM35s; however, Chi-CBM35 was found in the insoluble fraction of *E. coli* lysates. To purify Chi-CBM35, the cell pellet from lysed cells was resuspended in 100 ml 8 M urea and, after removal of any insoluble material, was dialyzed 6 times against 25 vol of 20 mM Tris/HCl buffer, pH 8.0, containing 0.5 M NaCl, 2 mM $CaCl_2$, and 2 mM $MgCl_2$ over a period of 3 days. The protein was purified by IMAC using a 2 ml of Ni^{2+} affinity resin (GE Healthcare) and concentrated in a stirred cell ultrafiltration device with a 5 kDa MWCO membrane (Millipore). SDS/PAGE showed that the 4 CBM35s were >95% pure.

Carbohydrate and Metal Binding Studies. The binding of CBM35s to carbohydrates was determined by isothermal titration calorimetry, which was carried out in 50 mM Hepes-Na buffer, pH 8.0, containing 2 mM $CaCl_2$ unless otherwise specified, at 25 °C. The concentrations of the ligands in the syringe were ≈3 mM and the CBM35s in the reaction cell were at 80–100 μM. The binding of the CBM35s to metal ions was determined as described previously (1). The determined K_A and ΔH values were used to calculate ΔS from the standard thermodynamic equation.

Immunofluorescence Microscopy. ***Bacteria.*** Cells fixed and made permeable on glass slides were incubated for 120 min at room temperature with primary chicken anti-CsxA antibodies and rabbit anti-Chi-CBM35 antibodies or with rabbit anti-Csn antibodies at a 1:1,200 dilution in PBS. After several washes in PBS, primary antibodies were detected by incubation with goat anti-rabbit Alexa Fluor 488 IgG and goat anti-chicken Alexa Fluor 555 IgG at a 1:200 dilution in PBSB for 90 min at room temperature. After several washes in PBSD and in water the DNA was stained for 10 min with Hoechst 33342 at a 1:5,000 dilution. After several washes in water, slides were mounted in Prolong Gold anti-fade reagent. Pictures were taken using an inverted Olympus IX 70 microscope with a Cool SNAP-Pro *cf*. monochrome camera and Image-Pro Plus software or with an Olympus Fluoview FV 300 confocal system.

Plant cell walls. For enzyme treatments before the immunohistochemistry, pectic homogalacturonan hydrolysis was carried out by incubating sections with 10 $\mu g\ ml^{-1}$ of the *Cellvibrio japonicus* pectate lyase *Cj*PelA in 50 mM CAPS, 2 mM $CaCl_2$ buffer pH 10, or with 40 $\mu g\ ml^{-1}$ of the *Aspergillus niger endo*-polygalacturonase M2 (Megazyme) in phosphate–acetate buffer pH 5.0. Treatments were performed for 2 h at room temperature. Due to the CBM35s sensitivity to calcium, a Tris-buffered saline (TBS) containing 5 mM $CaCl_2$ and 5% (wt/vol) BSA (BSA/TBS-Ca) was used for the incubation steps. The CBM35s were used at 100 $\mu g\ ml^{-1}$. The washing steps were performed with the TBS-Ca buffer. Samples were washed, mounted, and observed as described previously (2, 3).

Crystallization. ***Chi-CBM35.*** Chi-CBM35 was treated with thrombin to remove the His_6-tag and the resulting cleaved protein was buffer exchanged into 5 mM Tris/HCl buffer pH 8.0. Crystals were obtained at 18 °C by the vapor diffusion hanging drop method with 20 $mg\ ml^{-1}$ Chi-CBM35 in a mother liquor containing 20% PEG4000, 10% isopropanol, 0.1 M sodium Hepes pH 7.4, at a 1:1 ratio of protein solution/mother liquor. A heavy atom soak was used for phasing, where crystals were soaked in 1 M NaI in mother liquor for 30 min. Crystals were frozen at 113 K directly in the cryostream after a short soak in mother liquor supplemented with 20% ethylene glycol. 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. An atomic resolution data set for native Chi-CBM35 was collected on beamline X8C at the National Synchrotron Light Source (Brookhaven National Laboratories, Upton, NY). Complexes of Chi-CBM35 with GlcA and Δ4,5-GalAα1,4Gal (purified as described previously) (4) were obtained by soaking crystals of Chi-CBM35 in mother liquor supplemented with 10 mM sugar. These crystals were cryoprotected and data collected as described for the native crystals.

Xyl-CBM35, Rhe-CBM35, and Pel-CBM35. Proteins were crystallized by vapor diffusion at 20 °C. Pel-CBM35 was buffer exchanged into water in a 20 ml stirred ultra filtration device using a 10 kDa MWCO membrane and concentrated to 9.4 $mg\ mL^{-1}$. Crystals were grown from 0.25 M sodium phosphate/citrate buffer, pH 4.2, 0.2 M NaCl, and 20% PEG8000 (condition 26 of the Qiagen JCSG derivative screen) in 4 days, and frozen in the mother liquor containing 20% ethylene glycol as cryoprotectant. Data were collected on beamline ID14-4 at the European Synchrotron Radiation Facility (ESRF; Grenoble, France) to 1.7 Å. Xyl-CBM35 was crystallized in 20% (wt/vol) PEG3350, 0.4 M thiocyanate, pH 6.9 (condition 14 of the Qiagen JCSG derivative screen) with protein at 24 $mg\ ml^{-1}$ in 2 mM $CaCl_2$. To obtain crystals of Xyl-CBM35 in complex with Xyl-GlcA the ligand was included in the same crystallization conditions at 2 $mg\ ml^{-1}$. Crystals of the apo protein appeared after 21 days, whereas in the presence of Xyl-GlcA, crystals appeared after ≈100 days. The crystals were cryoprotected with the mother liquor with the addition of 30% (vol/vol) PEG400. Data for the native and complexed protein were collected on beamline IO3 at the Diamond Light Source (Oxfordshire, United Kingdom) to 1.9 Å

and on a Rigaku 007 MicroMax copper rotating anode using an RAXIS IV++ imaging plate detector to 1.6 Å, respectively. Native crystals of Rhe-CBM35 were crystallized from 0.1 M Hepes sodium pH 7.5, 0.8 M sodium phosphate monobasic, 0.8 M potassium phosphate monobasic with protein at 20 mg ml^{-1}, and crystals of Rhe-CBM35 in complex with Δ4,5-GalAα1,4GalA from 1.6 M sodium citrate with protein at 20 mg ml^{-1} supplemented with 1% (wt/vol) pectate lyase-digested polygalacturonic acid and 2 mM calcium chloride before mixing with mother liquor. The crystals were cryoprotected with mother liquor supplemented with 2 mM calcium chloride, 1% (wt/vol) digested polygalacturonic acid. Data for the native protein were collected on beamline ID29 at the ESRF to 1.5 Å, and the complex collected on a Rigaku 007 MicroMax copper rotating anode using an RAXIS IV++ imaging plate detector to 1.4 Å

Data Processing and Structure Determination. Data for Chi-CBM35 were processed using Crystal Clear/d*trek (5) and all other computing was carried out with the CCP4 suite (6) unless otherwise stated. ShelxD (7) was used to locate 5 iodide ions in the asymmetric unit using only the anomalous signal of the iodide derivative. Initial phasing to 3.0 Å was done by SAD using SHARP (8). Density modification and phase extension was done with DM (9). ARP/wARP (10) was used for automatic model building followed by manual correction using COOT (11). Refinement was carried out with REFMAC (12). The native model was used to solve the complex structures by molecular replacement using MOLREP (13). All data sets contained 2 molecules in the asymmetric unit. Water molecules were added using the REFMAC implementation of ARP/wARP and inspected visually before deposition. In all data sets, 5 percent of the observations were flagged as "free" (14) and used to monitor refinement procedures.

Data for the native Rhe-CBM35 were integrated and scaled using DENZO and SCALEPACK (15) and for Pel-CBM35, Rhe-CBM35-Δ4,5-GalAα1,4Gal, Xyl-CBM35 (native and complexed with GlcA) using MOSFLM and SCALA (16). The structures of Rhe-CBM35 and Xyl-CBM35 were determined by molecular replacement in MOLREP (13) and for Pel-CBM35 using AMoRe (17) using Chi-CBM35 as the search model. The starting models for Rhe-CBM35, Pel-CBM35, and Xyl-CBM35 were built automatically using ARP/wARP (10), which was followed by rounds of manual rebuilding in COOT (11) interspersed with restrained refinement in REFMAC (12). Solvent water molecules for Rhe-CBM35 and Pel-CBM35 were added using COOT, and checked manually. Xyl-CBM35–GlcA and Rge-CBM35-Δ4,5-GalAα1,4GalA complexes were solved by molecular replacement using MOLREP with the coordinates for the native proteins used as search models and refined similarly to the native crystals. Water molecules were added using Arp_waters, and checked manually using COOT. All final model statistics are given in Table S4. Coordinates and structure factors have been deposited with the PDB codes of Pel-CBM35, 2W3J; Rhe-CBM35, 2W1W; Rhe-CBM35+Δ4,5-GalAα1,4Gal, 2W47; Xyl-CBM35, 2W46; Xyl-CBM35+GlcA, 2W87; Chi-CBM35, 2VZP; Chi-CBM35+GlcA, 2VZR; Chi-CBM35+Δ4,5-GalAα1,4Gal, 2VZQ for the 8 structures.

All structural figures in the main article and Figs. S1 and S2 were produced using Bobscript (18).

1. Bolam DN, *et al*. (2004) X4 modules represent a new family of carbohydrate-binding modules that display novel properties. *J Biol Chem* 279:22953–22963.
2. McCartney L, *et al*. (2006) Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. *Proc Natl Acad Sci USA* 103:4765–4770.
3. McCartney L, Gilbert HJ, Bolam DN, Boraston AB, Knox JP (2004) Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. *Anal Biochem* 326:49–54.
4. Abbott DW, Boraston AB (2007) Specific recognition of saturated and 4,5-unsaturated hexuronate sugars by a periplasmic binding protein involved in pectin catabolism. *J Mol Biol* 369:759–770.
5. Pflugrath JW (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr D* 55:1718–1725.
6. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
7. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D* 58:1772–1779.
8. Evans G, Bricogne G (2002) Triiodide derivatization and combinatorial counter-ion replacement: Two methods for enhancing phasing signal using laboratory Cu K alpha X-ray equipment. *Acta Crystallogr D* 58:976–991.
9. Cowtan K (1994) 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31:34–38.
10. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
11. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
12. Lamzin VS, Wilson KS (1997) Automated refinement for protein crystallography. *Methods Enzymol* 277:269–305.
13. Vagin A, Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr D* 56:1622–1624.
14. Brunger AT (1992) Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475.
15. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
16. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 26.
17. Navaza J (1994) *AMoRe*: An automated package for molecular replacement. *Acta Crystallogr D* 50:157–163.
18. Esnouf RM (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J Mol Graphics Model* 15:132–134.

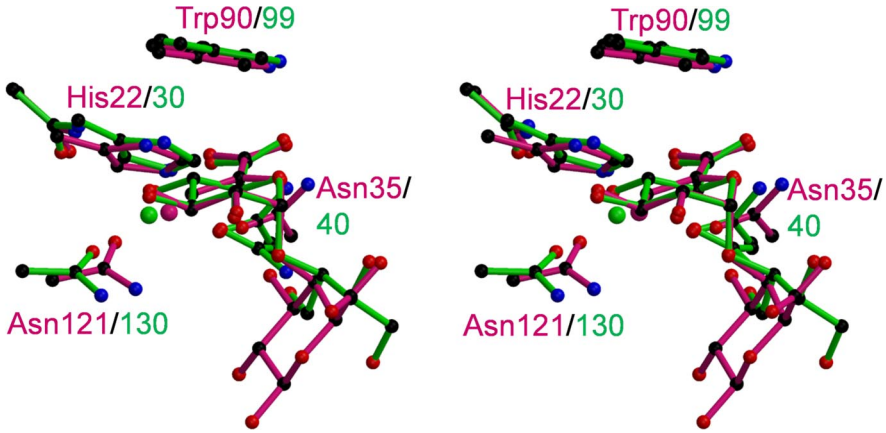

Fig. S1. Comparison of the Chi-CBM35 and Xyl-CBM35 binding sites. An overlay of the Chi-CBM35 binding site (pink) with Xyl-CBM35 (green) is shown in divergent stereo. Chi-CBM35 is shown in complex with Δ4,5-GalAα1,4Gal and Xyl-CBM35 in complex with a GlcA containing disaccharide. Calcium atoms are shown as spheres.

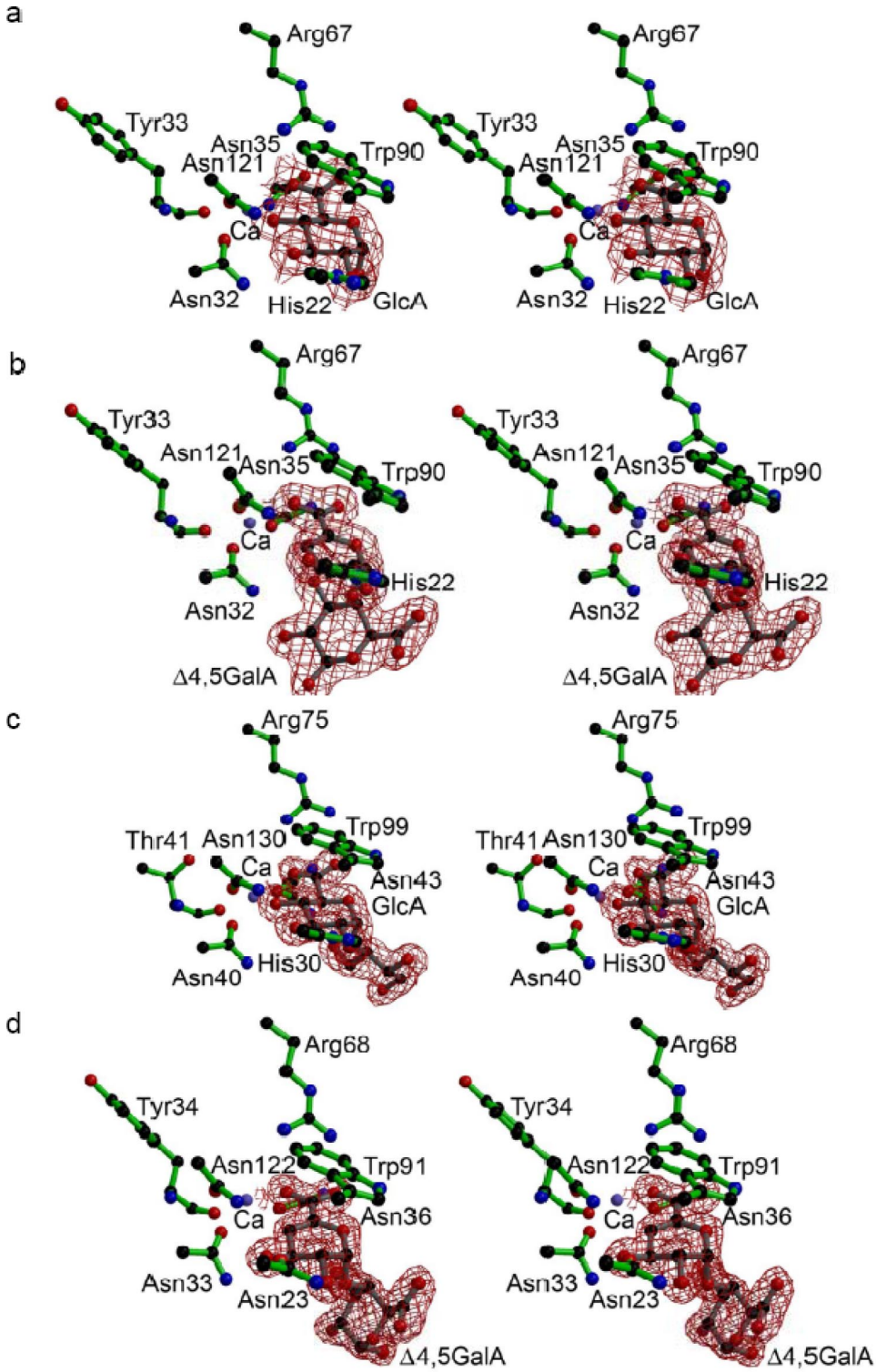

Fig. S2. Stereo views of the CBM35 binding sites. (*a*) The binding site of Chi-CBM35 in complex with GlcA. (*b*) The Chi-CBM35 binding site in complex with Δ4,5-Galα1,4Gal. (*c*) The binding site of Xyl-CBM35 in complex with a GlcA containing disaccharide. (*d*) The binding site of Rhe-CBM35 in complex with Δ4,5-Galα1,4Gal. The electron density maps are shown as maximum likelihood weighted $2F_{obs}$-F_{calc} maps contoured at 1σ.

Table S1. Assessment of the binding of CBM35s to sugars

Protein	Sugar	Binding
Xyl-CBM35	4-Methyl-glucuonic acid	No binding detected
Pel-CBM35	4-Methyl-glucuonic acid	No binding detected
Rhe-CBM35	4-Methyl-glucuonic acid	No binding detected
Xyl-CBM35	Galacturonic acid	No binding detected
Pel-CBM35	Galacturonic acid	No binding detected
Rhe-CBM35	Galacturonic acid	No binding detected
Xyl-CBM35	Mannose	No binding detected
Xyl-CBM35	Mannohexaose	No binding detected
Xyl-CBM35	Galactomannan	No binding detected
Xyl-CBM35	Glucomannan	No binding detected
Xyl-CBM35	Glucose	No binding detected
Xyl-CBM35	Cellohexaose	No binding detected
Xyl-CBM35	Laminarin	No binding detected
Xyl-CBM35	Glucomannan	No binding detected
Xyl-CBM35	Galactose	No binding detected
Xyl-CBM35	Galactan	No binding detected

The binding of the CBM35s to the sugars listed was by isothermal titration calorimetry using the sugars or oligosaccharides at 10 mM and the polysaccharides at 2mg/ml. The protein, which is in the cell, was at 50 μ M. Titrations were carried out in the presence of 2 mM calcium in 50 mM sodium Hepes buffer, pH 7.5.

Table S2. Evaluation of the binding of Xyl-CBM35 to *Cellvibrio japonicas* cells

Cellular fraction	Detection of Xyl-CBM35 added to a cell suspension of *Cellvibrio japonicus*
Cells recovered by centrifugation	No Xyl-CBM35 detected
Supernatant	Xyl-CBM35 detected

Cells of Xyl-CBM35 derived from a 100 mL overnight culture, after recovery by centrifugation, were resuspended in 1/10 volume of PBS and incubated with Xyl-CBM35 added to a final concentration of 10 μM for 2 h. The cells were removed by centrifugation and proteins in both the recovered cells and in the supernatant were subjected to Western analysis using an anti-His_6 antibody, which recognizes the His-tagged Xyl-CBM35, as the probe.

Table S3. Primers used to mutate Pel-CBM35

Residues	End of primer	Primers	End of primer
K21A-F	5′	CGTGATTGATAAC**GCG**CATGCTGGCTATACC	3′
K21A-R	3′	GGTATAGCCAGCATG**CGC**GTTATCAATCACG	5′
K21N-F	5′	CGTGATTGATAAC**AAT**CATGCTGGCTATACC	3′
K21N-R	3′	GGTATAGCCAGCATG**ATT**GTTATCAATCACG	5′
H22A-F	5′	GATTGATAACAAA**GCG**GCTGGCTATACC	3′
H22A-R	3′	GGTATAGCCAGC**CGC**TTTGTTATCAATC	5′
Y25A-F	5′	GATAACAAACATGCTGGC**GCG**ACCGGTACAGGATTTATCG	3′
Y25A-R	3′	CGATAAATCCTGTACCGGT**CGC**GCCAGCATGTTTGTTATC	5′
R67A-F	5′	GGCACAAGCGCA**GCG**CGTGCAACAGTAG	3′
R67A-R	3′	CTACTGTTGCACG**CGC**TGCGCTTGTGCC	5′
W89A-F	5′	CACCAACAGTAAT**GCG**ACCCAATGGCAAAC	3′
W89A-R	3′	GTTTGCCATTGGGT**CGC**ATTACTGTTGGTG	5′
D116A-F	5′	GCCGAGACTGCA**GCG**GGTTTAGCC	3′
D116A-R	3′	GGCTAAACC**CGC**TGCAGTCTCGGC	5′
D116N-F	5′	GCCGAGACTGCA**AAT**GGTTTAGCC	3′
D116N-R	3′	GGCTAAACC**ATT**TGCAGTCTCGGC	5′
N120A-F	5′	GATGGTTTAGCC**GCG**ATTGATAGTATTCGC	3′
N120A-R	3′	GCGAATACTATCAAT**CGC**GGCTAAACCATC	5′

Bold type indicates the nucleotides coding for the mutation. F, forward; R, reverse.

Table S4. Data processing and refinement for the 4 CBM35s and their complexes

Data process	Chi-CBM35 native	Chi-CBM35 + GlcA	Chi-CBM35 + Δ4,5-GalAα1,4Gal	Pel-CBM35 native	Xyl-CBM35 native	Xyl-CBM35 + GlcA	Rhe-CBM35 native	Rhe-CBM35 + Δ4,5-GalAα1,4Gal
Data collection beamline, source	X8C, National Synchrotron Light Source*	X8C, National Synchrotron Light Source*	X8C, National Synchrotron Light Source*	ESRF, ID14–4	Rigaku 007, MicroMax copper rotating anode	Diamond, IO3	ESRF, ID-29	Rigaku 007, MicroMax copper rotating anode
Resolution: outer shell, Å	15.10–1.05 (1.11–1.05)	19.71–1.95 (2.02–1.95)	20.00–1.70 (1.76–1.70)	42.0–1.70 (1.79–1.70)	23.88–1.9 (1.95–1.90)	24.1–1.58 (1.67–1.58)	15.0–1.55 (1.61–1.55)	74.54–1.4 (1.43–1.40)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$H3$	$P1$	P1	$P3_121$	$P6_522$
Unit cell parameters	a=58.6 Å, b=57.1 Å, c=81.7 Å	a=54.6 Å, b=57.0 Å, c=81.3 Å	a=28.9 Å, b=54.8 Å, c=81.3 Å; β=100.1°	$a=b$=83.9 Å, c=79.5 Å	a=28.2 Å b=46.1 Å, c=49.0 Å α=71.6° β=89.7° γ= 82.2 °	a=27.9 Å, b= 45.5 Å, c=48.9 Å; α=71.8° β=89.7° γ=81.3°	$a=b$=46.4 Å, c=204.5 Å	$a=b$=86.0 Å, c=75.9 Å
R_{merge}, outer shell	0.08 (0.46)	0.09 (0.27)	0.10 (0.39)	0.07 (0.48)	0.12 (0.29)	0.11 (0.22)	0.10 (0.20)	0.08 (0.64)
Mean I/σI, outer shell	13.3 (3.2)	8.9 (3.8)	9.2 (3.5)	15.7 (2.5)	10.6 (2.4)	11.7 (3.4)	14.3 (9.0)	24.1 (3.7)
Completeness: outer shell, %	98.9 (98.0)	97.7 (97.9)	98.8 (97.0)	99.7 (99.7)	94.1 (90.8)	91.9 (80.8)	99.0 (100.0)	99.7 (98.3)
Multiplicity, outer shell	6.2 (5.0)	4.3 (4.2)	6.6 (6.0)	5.3 (3.6)	1.7 (1.6)	3.1 (3.1)	7.7 (7.2)	13 (12.5)
No. unique reflections	117896	18625	27317	22914	16341	30820	37854	31544
R_{cryst}	0.13	0.18	0.17	0.16	0.14	0.14	0.14	0.20
R_{free}	0.15	0.25	0.22	0.19	0.22	0.19	0.18	0.25
RMSD bonds, Å	0.018	0.011	0.019	0.013	0.017	0.017	0.016	0.007
RMSD angles, deg	1.93	1.27	1.79	1.35	1.56	1.37	1.58	1.19
PDB code	2VZP	2VZR	2VZQ	2W3J	2W46	2W87	2W1W	2W47

*Brookhaven National Laboratories.