Molecular Basis for the Selectivity and Specificity of Ligand Recognition by the Family 16 Carbohydrate-binding Modules from *Thermoanaerobacterium polysaccharolyticum* ManA*

Received for publication, August 6, 2007, and in revised form, November 13, 2007. Published, JBC Papers in Press, November 19, 2007, DOI 10.1074/jbc.M706513200

Brian Bae*, Samuel Ohene-Adjei*, Svetlana Kocherginskiya*, Roderick I. Mackie*, M. Ashley Spies*, Isaac K. O. Cann*‡§¶, and Satish K. Nair*‡¶**1

From the Departments of *Biochemistry, ‡Microbiology, and §Animal Sciences, the **Center for Biophysics and Computational Biology, and the †Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Enzymes that hydrolyze complex polysaccharides into simple sugars are modular in architecture and consist of single or multiple catalytic domains fused to targeting modules called carbohydrate-binding modules (CBMs). CBMs bind to their ligands with high affinity and increase the efficiency of the catalytic components by targeting the enzymes to its substrate. Here we utilized a multidisciplinary approach to characterize each of the two family 16 carbohydrate-binding domain components of the highly active mannanase from the thermophile *Thermoanaerobacterium polysaccharolyticum*. These represent the first crystal structures of family 16 CBMs. Calorimetric analysis showed that although these CBMs demonstrate high specificity toward β-1,4-linked sugars, they can engage both cellulo- and manno-polysaccharides. To elucidate the molecular basis for this specificity and selectivity, we have determined high resolution crystal structures of each of the two CBMs, as well as of binary complexes of CBM16-1 bound to either mannopentaose or cellopentaose. These results provide detailed molecular insights into ligand recognition and yield a framework for rational engineering experiments designed to expand the natural repertoire of these targeting modules.

The turnover of photosynthetically fixed carbon through the action of microbial glycoside hydrolases has been estimated to be of the order of 10¹¹ tons annually (1). Consequently, glycoside hydrolases play a key role in the global carbon cycle, and their properties, if well harnessed, will enhance the release of substrates (monomeric sugars) critical to the use of cellulose materials in the biofuel industry. For most glycoside hydrolases, the polypeptides are organized in a modular arrangement that usually consists of a catalytic domain and an associated carbohydrate-binding mod-

---

2 These modules may be joined through linkers that are rich in proline, serine, and threonine (2). The evolutionary rationale that led to these complex molecular architectures is currently unclear (3).

It has been suggested that CBMs attain multivalency through multiplicity (4). Therefore, it could be hypothesized that multiple CBMs act synergistically to bind to their target ligand, leading to an increased accessibility of the catalytic domain to the target polysaccharide. This view is supported by studies on a recombinant double CBM, constructed by fusing family 1 CBMs of two *Trichoderma reesii* cellulohydrolases via a linker peptide (5). A 5–10-fold increase in the affinity for crystalline cellulose was observed for the double CBM compared with the individual modules, and similar results were observed for two CBMs of *Cellulomonas fimi* xylanase 11A (3). Thus, in some arrangements, the affinity for substrates increases with multiplicity of CBMs, and this appears to occur frequently in hyperthermophiles and thermophiles (6). Contrary to these findings, a report on the product of a *manA* gene of *Caldicellulosiruptor* strain RT8B.4 with two N-terminal CBMs showed no relationship between binding affinity and multiplicity of CBMs (7). Similar conclusions were also obtained from studies based on a *Thermotoga neapolitana* xylanase flanked by two different CBMs (7).

An alternative hypothesis that may explain multiplicity of CBMs in glycoside hydrolases is to increase the diversity of their target polysaccharides or substrates, because cellulases and xylanases found in plant cell walls are heterogeneous in nature. The *Paenibacillus* sp. strain W-61 xylanase 5 seems to support this hypothesis. In this polypeptide, there is a C-terminal family 22 and an N-terminal family 9 CBM, respectively (8). The family 22 CBM binds to xylan, whereas the family 9 CBM binds to cellulose (4).

Carbohydrate-binding modules have been identified in over 100 β-1,4-glucanases (9, 10), and aside from increasing the relative concentration of their respective enzymes on the surface of the substrate, they may also modify substrates into amorphous forms and thus make available to the catalytic domain a more easily degradable substrate (11–13). This is perhaps the case for the two CBMs found in the highly active mannanase

---

The abbreviations used are: CBM, carbohydrate-binding module; MES, 4-morpholineethanesulfonic acid; r.m.s.d., root mean square deviation; PDB, Protein Data Bank; SeMet, selenomethionine.

---

1 To whom correspondence should be addressed: Dept. of Biochemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801. Tel.: 217-333-0641; Fax: 217-244-5858; E-mail: snair@uiuc.edu.

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

**This work was supported in part by American Cancer Society Grant 04-37 (to S. K. N.) and in part by the Agricultural Experimental Station Grant ILLU 538-364 (to I. K. O. C.). 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.**
Crystal Structures of Family 16 CBMs

(ManA) of *Thermoanaerobacterium polysaccharolyticum*, because their removal compromises enzyme activity (14). The widespread association of CBMs with glycoside hydrolases suggests that enzyme adsorption play a critical role in the degradation of polysaccharides. Insoluble plant cell wall polysaccharides such as cellulose hold a lot of promise as future substrates for energy production. Understanding the mechanism of CBM-mediated adsorption to β-glucan may therefore be critical to our capacity to fully utilize energy crops as substrates for industrial processes, especially for the biofuels industry.

In an effort to understand the influence of the *T. polysaccharolyticum* carbohydrate-binding modules on the enzymatic activity of the parental protein, we have used bioinformatics to delineate its two linker-less CBMs and carried out biochemical and structural studies on the resultant polypeptides. This represents the first comprehensive structure-thermodynamic analysis of family 16 carbohydrate-binding modules. Here we show that the first and second carbohydrate-binding modules from *T. polysaccharolyticum* (TpolCBM16-1 and TpolCBM16-2, respectively) display high affinity for β,1,4-glucose and β,1,4-mannose polymers and provide a quantitative assessment of ligand binding using isothermal titration calorimetry. We also present four different high resolution x-ray crystallographic structures as follows: 1) TpolCBM16-1, determined to resolution of 1.5 Å; 2) TpolCBM16-2, determined to 1.9 Å resolution; 3) the co-crystal structure TpolCBM16-1 bound to cellobiose, determined to 1.2 Å resolution; and 4) the co-crystal structure of TpolCBM16-1 bound to mannotetraose, determined to 2.2 Å resolution. These structures demonstrate that each module contains a single cleft, with two aromatic residues located at the base and along one face and two stretches of polar residues on either side of this cleft, which constitutes the ligand-binding site. The molecular basis for both plasticity in ligand selectivity and specificity by these modules is discussed.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification**—The genes encoding the first and second carbohydrate-binding module from *T. polysaccharolyticum* were cloned by PCR. To amplify the gene segment encoding TpolCBM16-1, a forward primer with the sequence 5′-CATATGtgaacatcttgaacccg, with an engineered NdeI restriction site (shown in capitals) preceding the start codon, and a reverse primer 5′-CTCGAgttaacttcaccaagctaatatatagtggaacag, with an engineered XhoI site (shown in capitals) following the stop codon, were used with the *maNA* gene as the template. The ManA TpolCBM16-2 was also amplified by PCR with forward primer (5′-CATATGtgaacatcttgaacccg) and reverse primer (5′-CTCGAgttaacttcaccaagctaatatatagtggaacag). Note that these two primers also contained engineered NdeI and XhoI sites in the forward and reverse primers, respectively. Each TpolCBM16 encoding sequence was cloned into a TA cloning vector, pGEM-T (Promega), and sequenced to confirm the correctness of the coding sequence. Each gene was then removed and inserted into a modified pET-28 plasmid (Novagen) bearing an N-terminal polyhistidine tag and a thrombin digestion site preceding the inserted gene. The modification in the pET28 vector was the replacement of the gene encoding kanamycin resistance with that encoding ampicillin resistance. The integrity of all constructs was confirmed by sequencing.

*Escherichia coli* expression strain BL21(DE3) was transformed with the appropriate expression vectors, and single colonies of transformed *E. coli* were used to inoculate 5 ml of LB medium supplemented with ampicillin (100 μg/ml). Five hours following inoculation, the small scale culture was added to 1 liter of LB medium containing ampicillin (100 μg/ml) for growth at 37 °C. When the A600nm of the culture reached 0.3, protein expression was induced with 0.1 mM isopropyl β-D-thiogalactopyranoside, and the cells were further grown for 6 h.

Bacterial cells were pelleted by centrifugation (4000 × g, 10 min), resuspended in 50 mM potassium phosphate buffer, pH 7.0, and disrupted by sonication. The cell debris were removed by centrifugation (15 000 × g, 15 min).

Recombinant proteins were purified from the clarified supernatant by virtue of the N-terminal polyhistidine tag using a Talon resin column (Clontech) charged with cobalt chloride. Following elution from the cobalt affinity resin, the cleavable polyhistidine tag was removed using thrombin (1 unit/mg protein; GE Healthcare). The protein was further purified by anion exchange (5 ml HiTrap Q; GE Healthcare) and size exclusion chromatographies (Superdex 75 16/60; GE Healthcare) prior to crystallization. Site-specific variants were purified in essentially the same manner. Selenomethionine-incorporated TpolCBM1 produced was produced by the method of Van Duyne et al. (15) and purified in the manner described above except that 5 mM β-mercaptoethanol was added to all of the buffers. The protein samples were estimated to be greater than 95% pure as judged by SDS-PAGE.

*Affinity Gel Electrophoretic Mobility Assays*—Assays to monitor the binding of TpolCBM16-1 and TpolCBM16-2 to soluble substrates were performed as described by Tomme et al. (16) in the absence of SDS. 2-Mercaptoethanol was excluded from the loading buffer, and the proteins were not heated at 95 °C prior to loading to the gel. Complexes were resolved on a nondenaturing separating 12% polyacrylamide gel utilizing 0.1% of the following soluble polysaccharide substrates: carboxymethyl cellulose, locust bean gum (galactomannan polysaccharide), or xylan. In all experiments the same proteins were run simultaneously in the control gel without incorporated soluble substrates. The electrophoresis was carried out at 100 V at 4 °C for 3 h. Proteins were visualized by staining with Coomassie Brilliant Blue R-250.

*Isothermal Titrations Calorimetry*—Measurements were made at 25 °C using a VP-ITC calorimeter (MicroCal, Inc, Northampton, MA) following the manufacturer’s recommended procedures. All samples were extensively dialyzed against 50 mM potassium phosphate buffer, pH 7.0, and all ligands were dissolved in the same buffer. The protein sample (100–200 μM), in a 1.4399-ml reaction cell, was injected with 28 successive 10-μl aliquots of ligand (1–1.5 mM) at 300-s intervals. Based on the results from our crystallographic experiments, the number of binding sites was presumed to be one. Data were fitted by nonlinear regression using a single site
model (MicroCal Origin), and thermodynamic parameters were calculated using the Gibbs free energy equation \( \Delta G = \Delta H - T \Delta S \), and the relation \( -RT \ln K_p = \Delta G \).

**Crystallization and X-ray Data Collection**—Initial crystallization conditions were established by the sparse-matrix sampling methods using commercial and homemade screens. Refinement of promising conditions yielded large crystals suitable for diffraction analysis. Crystals of TpolCBM16-1, TpolCBM16-2, TpolCBM16-1-cellopentaose, and Tpol-CBM16-1-mannopentaose complexes were all grown using the hanging drop vapor diffusion method. For crystallization of native and SeMet-labeled TpolCBM16-1, 2 \( \mu l \) of protein sample (20 mg/ml, in 100 mM KCl, 10 mM HEPES, pH 7.5) was added to 2 \( \mu l \) of precipitant (3M sodium formate, pH 7, 15% ethylene glycol) and equilibrated over a well containing the precipitant solution at 20 °C. Crystals grew within 3 days and reached a maximum size of 0.3 × 0.4 × 0.6 mm. As the crystallization precipitant proved to be suitable for cryo-protection, single crystals were harvested straight from the crystallization drop and flash-cooled in liquid nitrogen prior to data collection.

Co-crystals of TpolCBM16-1-cellopentaose complex were grown by adding 2 \( \mu l \) of protein sample (20 mg/ml, in 100 mM KCl, 10 mM HEPES, pH 7.5, plus 5 mM cellopentaose) to 2 \( \mu l \) of precipitant (30% polyethylene glycol 3350, 100 mM Tris-HCl, pH 8.3, 200 mM MgCl₂ and equilibrating over a well containing the precipitant solution at 20 °C. Crystals grew within 3 days and reached a maximum size of 0.4 × 0.4 × 0.4 mm. Prior to flash-cooling, crystals were briefly soaked in the precipitant solution supplemented with either 20% methylpentane diol or glycerol.

Co-crystals of the TpolCBM16-1-mannopentaose complex were grown by adding 2 \( \mu l \) of protein sample (20 mg/ml, in 100 mM KCl, 10 mM HEPES, pH 7.5, plus 5 mM mannopentaose) to 2 \( \mu l \) of precipitant (1.6 mM ammonium sulfate, 100 mM MES, pH 6.5, 10% dioxane) and equilibrating the mixture over a well containing the precipitant solution at 20 °C. Crystals grew within 7 days and reached a maximum size of 0.2 × 0.2 × 0.4 mm. Prior to flash-cooling, crystals were briefly soaked in the precipitant solution supplemented with 25% glycerol.

Initial screening with TpolCBM16-2 established sodium formate as a viable precipitant for crystallization. These crystals grew as thin needles and could not be improved upon despite extensive screening. Sequence gazing showed that TpolCBM16-2 contained a cysteine (Cys⁷⁴) at a position occupied by Phe⁶⁷ in TpolCBM16-1. Reasoning that Cys⁷⁴ may be fairly reactive, TpolCBM16-2 was incubated with 1 mM of mercury(II) acetate, to cap this residue, prior to crystallization trials. Large single crystals of CBM16-2 were grown by adding 2 \( \mu l \) of protein sample (20 mg/ml, in 100 mM KCl, 10 mM HEPES, pH 7.5, plus 1 mM Hg₃C₂H₄O₂) to 2 \( \mu l \) of precipitant (3.5 mM sodium formate, pH 7.0) and equilibrating over a well containing the precipitant solution at 20 °C. Prior to flash-cooling, crystals were briefly soaked in the precipitant solution supplemented with 20% ethylene glycol.

Flash-cooled crystals of unliganded TpolCBM16-1 diffracted x-rays to a minimum Bragg spacing of 1.3 Å, using an insertion device x-ray beam line utilizing an ADSC Q4 CCD detector (IMCA-CAT-Sector 17ID, Advanced Photon Source, Argonne, IL). These crystals occupied space group \( P2_12_12_1 \), with unit cell parameters \( a = 57.4 \AA, b = 77.8 \AA, \) and \( c = 77.9 \AA \) and contained two molecules in the crystallographic asymmetric unit. A 6-fold redundant data set was collected to a limiting resolution of 1.4 Å (overall \( R_{merge} = 4.60, I/\sigma(I) = 3.6 \) in the highest resolution shell).

Crystals of TpolCBM16-1-cellopentaose complex diffract x-rays beyond a Bragg spacing of 1 Å at an insertion device synchrotron beam line (SER-CAT Sector 22BM, Advanced Photon Source, Argonne, IL). Crystals occupy space group \( P1 \) with unit cell parameters \( a = 32.5, b = 48.8 \AA, c = 49.4 \AA, \alpha = 62.0°, \beta = 84.9°, \gamma = 85.9° \), with two molecules in the crystallographic asymmetric unit. A 4-fold redundant data set was collected to 1.2 Å resolution (overall \( R_{merge} = 3.6, I/\sigma(I) = 6 \) in the highest resolution shell) utilizing a Mar 225 CCD detector.

A 6-fold redundant data set was collected from crystals of TpolCBM16-1-mannopentaose complex (overall \( R_{merge} = 3.9, I/\sigma(I) = 3.5 \) in the highest resolution shell) to a limiting Bragg spacing of 2.2 Å at an insertion device synchrotron beam line utilizing an ADSC Q4 CCD detector (IMCA-CAT-Sector 17ID, Advanced Photon Source, Argonne, IL). Crystals occupy the trigonal space group \( P3_21 \) with unit cell parameters \( a = b = 74.1, c = 116.1 \AA, \) with two molecules in the crystallographic asymmetric unit.

Flash-cooled crystals of unliganded TpolCBM16-2 occupy space group \( C222_1 \), with unit cell parameters \( a = 90.4, b = 90.5 \AA, c = 139.4 \AA, \) with four molecules in the crystallographic asymmetric unit. The \( a \) and \( b \) unit cell edges are coincidentally close, and the crystals demonstrate tetragonal pseudo-symmetry consistent with space group \( P4_22_2 \), with two molecules in the crystallographic asymmetric unit. However, scaling of data in the higher symmetry space group results in a slight increase in merging statistics, and subsequent crystallographic refinement in the tetragonal setting resulted in a free \( R \) factor that stalls at 40%. Hence, the data were scaled in the orthorhombic setting and structure determination, and refinement was carried out in space group \( C222_1 \). A 6-fold redundant data set was collected to a limiting resolution of 1.9 Å (overall \( R_{merge} = 4.3, I/\sigma(I) = 12 \) in the highest resolution shell) utilizing a Mar 225 CCD detector (SER-CAT-Sector 22BM, Advanced Photon Source, Argonne, IL).

**Phasing and Structure Determination**—Although crystals of selenomethionine-substituted TpolCBM16-1 were grown under similar conditions, SeMet-labeled TpolCBM16-1 occupied a different space group (\( C222_1 \), with one molecule in the crystallographic asymmetric unit), with unit cell parameters \( a = 56.4 \AA, b = 96.8 \AA, c = 73.6 \AA \). A 6-fold redundant data set was collected at a wavelength near the selenium absorption edge, to a limiting resolution of 1.7 Å (overall \( R_{merge} = 5.5, I/\sigma(I) = 7 \) in the highest resolution shell) utilizing a ADSC Q4 CCD detector (IMCA-CAT-Sector 17ID, Advanced Photon Source, Argonne, IL). The structure of SeMet TpolCBM16-1 was solved by single wavelength anomalous diffraction utilizing anomalous scattering from the three selenium-substituted methionine residues per monomer of TpolCBM16-1. Data were indexed and scaled using the HKL2000 package (17, 18). Selenium sites were identified, and heavy atom parameters
were further refined using SOLVE (19, 20) yielding an initial figure of merit of 0.28–1.7 Å resolution (Table 1). Solvent flattening with RESOLVE further improved the quality of the initial map and permitted most of the main chain and 45% of the side chain residues to be automatically built by the software (21). The remainder of the model was fitted using XtalView (22) and further improved by rounds of refinement with REFMAC5 (23) and manual building.

The structure of native TpolCBM16-1 was determined by molecular replacement using the SeMet TpolCBM16-1 structure as a search probe (24). Following rigid body refinement of the initial molecular replacement solution, the atomic model was subject to automatic rebuilding using ARP/wARP (25), resulting in a near complete trace of both molecules in the crystallographic asymmetric unit. The remainder of the model was fitted using XtalView (22) and further improved by rounds of refinement with REFMAC5 (23) and manual building.

The structure of native TpolCBM16-1 was determined by molecular replacement (23) using the refined coordinates of the SeMet TpolCBM16-1 structure as a search probe. Multiple rounds of manual model building were interspersed with refinement using REFMAC5 (23) to complete structure refinement. Cross-validation used 5% of the data in the calculation of the free R factor (26, 27). Despite extensive manual rebuilding and refinement, residues in two loop regions, bridging Tyr<sup>92</sup> through Thr<sup>94</sup> and Thr<sup>111</sup> through Phe<sup>116</sup>, respectively, remain ill-defined and have not been modeled in the structure. Consequently, the free R factor of 30% for this structure is slightly higher than would be expected, given the resolution limit of 1.9 Å. The stereochemistry of the model was routinely monitored throughout the course of refinement using PROCHECK (28).

Crystal parameters, data collection parameters, and refinement statistics for each of the structures are summarized in Table 1. The refined coordinates have been deposited in the PDB with identification numbers 2ZEW (TpolCBM16-1), 2ZEX (TpolCBM16-1-cellobipectose), 2ZEY (TpolCBM16-1-mannopentaose), and 2ZEZ (TpolCBM16-2).

**Computational Procedures**—All calculations were performed with MOE 2006.08 (Chemical Computing Group, Inc.), using either the MMFF94x or OPLS-AA force fields. Starting coordinates for the energy minimization of the TpolCBM16-1-cellobipectose and TpolCBM16-1-mannopentaose complexes (this study) and the cellobipectose-CBM complex from *Cellulomonas fimi* (PDB 1gu3), solved by Davies and co-workers (29), were taken from the coordinates of the crystal structures, in which crystal waters were included. All atoms (solvent, CBM, and ligand) were energy-minimized until an energy gradient of 10<sup>-4</sup> kcal/mol/Å<sup>2</sup> was achieved. Interaction energy of the ligand with CBM and explicit solvent was then calculated using the MMFF94x force field, respectively.

**RESULTS**

**Delineation of the CBM Domain Boundaries**—In a previous report (14), our delineation of TpolCBM16-1 and Tpol-

---

**Table 1.** Data collection, phasing, and refinement statistics

| Data collection | Native CBM16-1 | SeMet CBM16-1 | Cellopectose | Mannopentaose | Native CBM16-2 |
|-----------------|---------------|---------------|--------------|--------------|---------------|
| Space group     | P<sub>2</sub>1<sub>a</sub> | C22<sub>1</sub> | P1           | P3 21        | C22<sub>1</sub> |
| Completeness (%)| 97.8 (87.7)   | 100.0 (100.0) | 94.2 (87.5)  | 98.9 (91.7)  | 93.5 (85.8)  |
| Redundancy      | 5.8 (3.9)     | 8.5 (8.5)     | 3.7 (3.5)    | 5.7 (3.6)    | 5.6 (5.4)    |
| R<sub>f</sub> (%) | 20.0-1.4      | 20.0-1.7      | 20.0-1.2     | 25.0-2.2     | 30.0-1.9     |
| R<sub>free</sub> (%) | 27.0 (26.4)  | 26.0 (26.4)  | 25.0 (26.4)  | 24.0 (26.4)  | 30.0-1.9     |
| R<sub>free</sub>/R<sub>work</sub> | 17.7/19.6    | 19.2/22.4     | 15.6/17.9    | 16.9/21.2    | 25.5/30.0    |
| No. of reflections | 64,126       | 21,336        | 75,010       | 18,027       | 48,583       |
| No. of atoms    | 2290          | 1102          | 2261         | 2246         | 4001         |
| Protein         | 0             | 0             | 112          | 112          | 0            |
| Sugar           | 0             | 0             | 520          | 229          | 268          |
| Water           | 341           | 190           | 520          | 229          | 268          |
| B-factors       | 19.00         | 19.49         | 12.06        | 51.09        | 30.27        |
| Sugar           | NA<sup>a</sup> | NA<sup>a</sup> | 14.44        | 62.8         | NA           |
| Water           | 37.32         | 36.06         | 28.12        | 53.44        | 34.64        |
| r.m.s.d. Bond lengths (Å) | 1.25 | 1.38 | 1.24 | 1.42 | 1.35 |
| Bond angles (%) | 0.009         | 0.011         | 0.008        | 0.011        | 0.013        |

<sup>a</sup> Highest resolution shell is shown in parentheses.

<sup>b</sup> NA means not applicable.
CBM16-2 suggested that the two modules were separated by 23 amino acids; hence each repeat started with a highly conserved N terminus characterized by the seven amino acids SAVPEAA and ended with a highly conserved sequence LVEV (Fig. 1C). We cloned and expressed each module based on this demarcation. However, neither of the products bound to Sigmacell type 50, although a construct with both putative TpolCBM16s included bound to this substrate. A visual inspection of the 23 amino acids occurring between the putative TpolCBM16-1 and putative TpolCBM16-2 showed that the sequence shares conserved residues with the 23 amino acids preceding the putative TpolCBM16-1. Therefore, we made two new constructs that included the 23 amino acids upstream of the previously demarcated N termini of the putative TpolCBM16s. Each of the gene constructs reported here that extend the N-terminal regions of each of TpolCBM16 construct demonstrates tight binding to Sigmacell type 50. Thus, although our original C-terminal delineation of the duplicated TpolCBM16 of ManA was correct, the N-terminal demarcation was erroneous.

**Thermodynamic Analysis of Ligand Specificity**—The binding specificity of the purified TpolCBM16-1 and TpolCBM16-2 was determined by native electrophoretic mobility assays using various polysaccharide substrates. These studies did not show retardation of each of these modules in the presence of xylans or β1,3-linked glucose polymers. In contrast, tight affinity was observed between both TpolCBM16-1 or TpolCBM16-2 and β1,4-linked glucose polymers such as cellulose. The affinity of these polypeptides for cellulose was such that once the individual module was bound to cellulose, the complex could only be dissociated by denaturing the protein.

The binding characteristics of each of these carbohydrate-binding modules were further quantified by isothermal titration calorimetric analysis using oligosaccharides of defined chain length. The binding isotherms for the interaction of TpolCBM16-1 or TpolCBM16-2 with each oligosaccharide were fit by nonlinear regression with a simple bimolecular interaction model (Table 2). Calorimetric analysis using long chain polymers hydroxyethylcellulose and glucomannan demonstrates that TpolCBM16-1 interacts with glucomannan with a near 9-fold greater affinity than for hydroxyethylcellulose. Glucomannan is an abundant water-soluble polysaccharide that consists of mixed β-1,4-linked polymers containing glucose and mannose. Consistent with our thermodynamic data, glucomannan is likely the natural substrate for TpolCBM16-1 and TpolCBM16-2.

For the interactions with synthetic short chain oligosaccharides, the binding stoichiometries are consistent with a 1:1 interaction. The measured binding affinities increase as a function of chain length with an optimum consistent with a binding site that can accommodate a polymer of five sugar molecules. As is typical of protein-carbohydrate interactions, the thermodynamics of the interaction of TpolCBM16-1 and Tpol-
Crystal Structures of Family 16 CBMs

TABLE 2
Isothermal calorimetric analysis of CBM16-1 and CBM16-2 ligand complexes
NB means not calculated.

| Complex            | $K_a$ (X 10$^3$ M$^{-1}$) | $\Delta G$ (kJ/mol) | $\Delta H$ (kJ/mol) | $T\Delta S$ (kJ/mol) |
|--------------------|---------------------------|---------------------|---------------------|----------------------|
| CBM16-1-Celopentaose| 406 ± 15                  | -32.07 ± 0.24       | -78.69 ± 0.97       | -46.62 ± 0.71        |
| CBM16-1-Cellotetrose| 68.2 ± 2.0                | -27.63 ± 0.10       | -53.12 ± 0.44       | -25.49 ± 0.32        |
| CBM16-1-Cellobiose  | 2 ± 4                     | NB                  | NB                  | NB                   |
| CBM16-2-Celopentaose| 170 ± 5.2                 | -29.91 ± 0.41       | -81.79 ± 0.83       | -51.88 ± 0.93        |
| CBM16-1-Mannopentaose| 153 ± 8                   | -29.56 ± 0.10       | -55.55 ± 0.21       | -25.99 ± 0.47        |
| CBM16-1-HEC         | 28.7 ± 3.3                | -25.55 ± 0.38       | -67.81 ± 2.7        | -42.26 ± 0.73        |
| CBM16-1-glucomannan | 253 ± 15.6                | -30.88 ± 0.20       | -76.10 ± 0.51       | -45.22 ± 0.28        |
| CBM16-2-glucomannan | 152 ± 24.8                | -29.60 ± 0.48       | -80.37 ± 1.65       | -50.77 ± 0.48        |
| CBM16 1+2-glucomannan| 245 ± 25.6                | -30.82 ± 0.53       | -103.7 ± 1.49       | -72.87 ± 0.55        |

![TpolCBM16-1-celopentaose isotherm](image1)

![TpolCBM16-1-mannopentaose isotherm](image2)
CBM-16 with β1,4-linked polymers are enthalpically driven, with an unfavorable entropic contribution.

**Overall Structures of CBM1 and CBM2**—Crystals of CBM1, grown using sodium formate as a precipitant, occupied space group $P2_12_12_1$, with two molecules in the crystallographic asymmetric unit and diffracted synchrotron X-radiation to a Bragg limit of 1.5 Å. To obtain crystallographic phases, crystals of TpolCBM16-1 were grown from selenomethionine-labeled protein. However, crystals of SeMet-labeled TpolCBM16-1 occupied a different space group ($C22_12_1$, with one molecule in the crystallographic asymmetric unit), and the structure of SeMet TpolCBM16-1 was solved to a resolution of 1.7 Å by single wavelength anomalous diffraction methods. The structure of native TpolCBM16-1 was solved to a resolution of 1.4 Å by molecular replacement using the refined coordinates of the structure of SeMet-labeled TpolCBM1. The two structures provide three independent views of the molecule and allow for detailed analysis of the structure, independent of crystal packing artifacts. The structure of TpolCBM16-2 was solved to a resolution of 2.1 Å by molecular replacement utilizing the final refined coordinates of TpolCBM1-1.

Both TpolCBM16-1 and TpolCBM16-2 fold into a compact domain typical of a β-sandwich fold consisting of two layers of five antiparallel β-sheets. Given the structural similarity between TpolCBM16-1 and TpolCBM16-2, for brevity, only the topology of TpolCBM16-1 will be described in detail. The bottom layer of the β-sandwich is composed of sheets β2 (Met$^{28}$-Val$^{107}$), β3 (Gly$^{37}$-Ile$^{46}$), β10 (Leu$^{134}$-Glu$^{143}$), β5 (Thr$^{59}$-Phe$^{68}$), and β8 (Thr$^{103}$-Thr$^{111}$). The upper layer of the sandwich is composed of sheets β1 (Gln$^{18}$-Asp$^{19}$), β4 (Ala$^{46}$-Asp$^{51}$), β9 (Gln$^{121}$-Lys$^{126}$), β6 (Phe$^{76}$-His$^{83}$), and β7 (Tyr$^{91}$-Phe$^{98}$). A single structural metal ion, presumed to be calcium by analogy to the structures of other carbohydrate-binding modules, is flanked by the N-terminal loop residues, the loop region between sheet β2 and β3, and the C-terminal region of β4. The ion has a coordination geometry characteristic of calcium and is engaged by the O-e2 of Glu$^{13}$, a bidentate interaction with Asp$^{139}$, and the main chain carbonyls of Gly$^{11}$, Asn$^{35}$, and Leu$^{38}$, and a water molecule. A diagram of the overall polypeptide fold of TpolCBM16-1 and TpolCBM16-2 is shown in Fig. 1, A and B, respectively, and a structure-based secondary structure assignment is shown in Fig. 1C.

The core architecture of TpolCBM16-1 and TpolCBM16-2 is similar to that observed in the three-dimensional structures of other families of carbohydrate-binding modules. A DALI search against other structures deposited in the Protein Data Bank shows the closest structural homologs to be the β-1,3-xylan-binding domain from CBM22 (30) (PDB code 1DYO) (r.m.s.d. of 2.9 Å over 138 aligned residues; Z-score = 14.8), the β-1,3-binding laminarinase of CBM4-2 (29) (PDB code 1GUIL) (r.m.s.d. of 2.1 Å over 123 aligned residues), and the sugar-binding domain of the F-box ubiquitin ligase (PDB code 1UMH) (r.m.s.d. of 3.1 Å over 135 aligned residues) (31). As in the structures of TpolCBM16-1 and TpolCBM16-2 presented here, a calcium ion is located at an equivalent position in the above noted structures of CBM22 and CBM4, where it contributes to the stability of the polypeptide (4, 29, 30).

Within the β-sandwich fold of TpolCBM16-1, the top layer of the sandwich is curved so as to create a cleft that runs along the resultant concave face, perpendicular to the direction of the β-sheets (Fig. 2A). This cleft spans about 25 Å across the face of the β-sandwich and constitutes the ligand-binding site (Fig. 2B). Two solvent-exposed tryptophan residues (Trp$^{20}$ and Trp$^{125}$) lie at the base and along one face of this cleft, respectively. A number of polar residues surround the periphery of this cleft, including Gln$^{21}$, Asp$^{77}$, and Asn$^{97}$ on one side and Gln$^{81}$, Gln$^{93}$, and Gln$^{121}$ on the other side, where they are poised to hydrogen bond with the sugar ligand.

Given the sequence identity of 62% over 144 amino acids between TpolCBM16-1 and TpolCBM16-2, it is not surprising that the two structures are nearly identical. The main chain atoms from final refined coordinates of the two structures can be superimposed with an r.m.s.d. of 0.8 Å over 120 atoms. Although the two polypeptides share the same overall fold, and contain similar structural features, such as the single structural calcium ion, significant differences exist at the ligand-binding site. A superposition of the two structures reveals that the largest deviations exist in the loop region encompassing residues 18–28 that includes the solvent-exposed Trp$^{77}$. In addition, several of the polar residues that surround the ligand-binding
Crystal Structures of Family 16 CBMs

FIGURE 3. A, difference Fourier electron density maps (contoured at 3σ over background) calculated with coefficients F_{obs} - F_{calc} at the ligand-binding site for the 1.2 Å resolution TpolCBM16-1-cellopentaose complex. B, direct hydrogen bonding interactions between residues of TpolCBM16-1 and cellopentaose. C, difference Fourier electron density maps (contoured at 3σ over background) calculated with coefficients F_{obs} - F_{calc} at the ligand-binding site for the 2.2 Å resolution TpolCBM16-1-mannopentaose complex. D, direct hydrogen bonding interactions between residues of TpolCBM16-1 and mannopentaose.

Site of TpolCBM16-1 are not conserved in TpolCBM16-2, including Gln^{21} (Tyr^{21} in CBM2) and Asn^{97} (Arg^{97} in TpolCBM16-2). These notable differences may account for the measured 2-fold lower affinity of TpolCBM16-2 for cellopentaose, relative to that for TpolCBM16-1 (see Table 2).

Molecular Basis for Ligand Specificity and Selectivity by CBMs—To determine the molecular basis for ligand specificity, we determined the crystal structure of TpolCBM16-1 bound to cellopentaose and to mannopentaose. The co-crystal structures of the TpolCBM16-1-oligosaccharide complexes were determined by molecular replacement using the refined coordinates of unliganded TpolCBM16-1 as a search model. Following rigid body refinement of the molecular replacement solutions, clear and continuous electron density, corresponding to each of the five β-1,4-linked sugars, could be observed in initial electron density maps (Fig. 3, A and C). Following further refinement and addition of solvent molecules, the pentasaccharide sugars were built into difference Fourier maps, and the binary complexes were refined without any constraints.

In the 1.2 Å resolution co-crystal structure of the TpolCBM16-1-cellopentaose complex, there are two identical copies of the complex in the asymmetric unit and the structural descriptions are apt for either molecule. Cellopentaose interacts with TpolCBM16-1 in an extended conformation in a 25 Å cleft located perpendicular to the face of the β-sandwich (Fig. 2). As in other co-crystal structures of other carbohydrate-binding modules, two solvent-exposed aromatic residues (Trp^{20} and Trp^{125}), which lie at the base and along one face of the ligand-binding cleft, engage in hydrophobic interactions with the last three terminal pyranosides in the reducing end of the oligosaccharide (Fig. 3B). The tryptophan residues are separated by roughly 8.5 Å, and their indole side chains are co-planar with the D-configured sugars to provide extensive hydrophobic contact. Mutational analyses show that replacement of either of these tryptophan residues compromised ligand recognition and demonstrate the importance of the hydrophobic platform created by these two residues in ligand binding.

In addition to these hydrophobic contacts, an extensive set of hydrogen bond interactions stabilizes the interaction between the protein and the remaining two sugars at the nonreducing end of the pentasaccharide ligand (Fig. 3D). Sets of polar residues flank either side of the ligand-binding cleft where they interact with the hydroxyl groups of the cellopentaose. Asn^{97}, Asp^{77}, and Gln^{21}, located along one face of the cleft, make contact with the glucose residues at sites 1, 2, and 4, respectively. Gln^{93}, Gln^{81}, and Gln^{121}, located along the opposite face of the cleft, engage the sugar residues at sites 3 and 4, respectively. Additional contacts are observed between the C-6 hydroxymethyl group of glucose molecules at the reducing end of the oligosaccharide and residues Gln^{21} and the backbone carbonyl of Trp^{20}. The interaction between TpolCBM16-1 and cellopentaose buries a total of 499.6 Å^2 of total surface area.

In the 2.2 Å resolution co-crystal structure of TpolCBM16-1 bound to mannopentaose, the pentasaccharide is likewise buried in an extended conformation within the 25 Å cleft located perpendicular to the face of the β-sandwich. As in the TpolCBM16-1-cellopentaose complex, hydrophobic residues Trp^{20} and Trp^{125} interact with the three pyranosides at the reducing end (Fig. 3D). Hydrogen bond interactions between sugars to sites 1 and 4, and protein residues Asn^{97} and Gln^{21} reside along one face of the ligand-binding cleft. Along the opposite face of the cleft, residues Gln^{93}, Gln^{81}, and Gln^{121} interact with the pyranosides at sites 3 and 4. The mannoses at sites 1 and 5 are twisted away from the ligand-binding cleft to accommodate the axial location of the C-2 hydroxyls of the mannoses (Fig. 3D). As a result, the TpolCBM16-1 forms fewer hydrogen bond interactions with the pyranosides of mannopentaose, relative to cellopentaose, and buries slightly less total surface area (430 Å^2) upon complex formation.

Comparison of the two co-crystal structures provides insights into the molecular basis for promiscuous ligand binding by TpolCBM16-1. Of particular interest is how the protein scaffold can accommodate the C-2 hydroxyl in an equatorial configuration in the cellopentaose complex, in contrast to how the axial C-2 hydroxyl is accommodated in the mannopentaose complex. The C-2 hydroxyls of the carbohydrates at sites 2 and 5 in the pentasaccharide complexes do not directly interact with the D-configured sugars to provide extensive hydrophobic contact. Mutational analyses show that replacement of either of these tryptophan residues compromised ligand recognition and demonstrate the importance of the hydrophobic platform created by these two residues in ligand binding.
In Silico Analysis of Ligand Specificity—To examine the nature of the TpolCBM16 specificity and affinity for ligands, the crystallographic complexes with cellobiotaose and manno- 

pentaose were compared with energy-minimized TpolCBM16-polysaccharide complexes. Calculation of interaction energies (i.e. the sum of nonbonding energies) between oligosaccharide and TpolCBM16-1 showed a clear correlation with experimental ΔH values determined from isothermal titration calorimetry (Fig. 4). When structural and thermodynamic data from a similar CBM-ligand complex (CBM-cellobiotaose complex from C. fimi, PDB code 1gu3) (29) are also included with the TpolCBM16-1 data from this work, a strong correlation is seen between CBM-ligand interaction energy and experimental ΔH (Fig. 4), with a slope of 6 ± 1, and r = 0.96. These calculations were all carried out using the MMFF94x force field. These energy calculations were also repeated using a different force field, the OPLS-AA, which yielded an almost identical trend in energy calculations were also repeated using a different force field.

In this study we have characterized each of the two carbohydrate-binding modules from T. polysaccharolyticum ManA using a combination of biochemical, thermodynamic, computational, and structural biological approaches. We demonstrate that each of the TpolCBM16s can bind both β-1,4-glucose and mannose polymers and quantify the binding affinity for such ligands using isothermal calorimetry. Our thermodynamic analysis suggests that the natural substrate for CBM16s is likely glucomannan, a mixed β-1,4-linked polymer containing both glucose and mannose that comprises nearly half the dry weight of roots and softwoods. Hence, this scaffold has been adapted for plasticity in substrate recognition. In addition, we present four crystal structures, including those of ligand-free TpolCBM16-1 and TpolCBM16-2 and of the TpolCBM16-1-cello-
Crystal Structures of Family 16 CBMs

pentose and TpolCBM16-1-mannopentose complexes, which provide a molecular rationale for understanding the ligand promiscuity and specificity by the TpolCBM16s.

Despite the binding promiscuity toward both mannann and glucan polymers, the TpolCBM16s are not able to interact with other oligosaccharides, including chitin, oligoxylans, and β-1,3-glucan species. Our co-crystal structures provide the chemical bases for understanding this binding specificity. Chitin is an unbranched polymer of N-acetyl-D-glucosamine that differs from cellulose by the replacement of the C-2 hydroxyl with an acetamido group. The lack of interaction between the TpolCBM16s and chitin may be explained by the steric clashes that would result between the acetamido groups and the protein scaffold at sugar residues at positions 3 and 4 of the polysaccharide. Although the chemical structures of xylan and cellulose are similar (xylans lack the C-6 hydroxymethyl group), their tertiary structures are quite distinct. Cellulose and mannan form flat polymers with a 180° rotation between consecutive sugars, whereas xylan forms helical structures distinguished by a 120° rotation between consecutive units (32).

Hence, TpolCBM16s cannot engage the helical orientation of xylan polysaccharides within a ligand binding cleft that has evolved to accommodate planar polymers such as those of cellulose and mannan. Similarly, β-1,3-glucan polymers are shown to form helical structures in solution. Modeling studies of laminarins demonstrate an energetically favorable helical configuration with φ and ψ angles of −72° and 108°, respectively (33). The binding preference for planar β-1,4-glucan polymers over helical β-1,3-glucans may likewise be dictated by the rigid nature of the binding cleft that is optimized to accommodate planar polymers.

The flexibility in ligand binding demonstrated by the TpolCBMs is reminiscent of plasticity observed for the family 29 carbohydrate-binding modules from the Piromyces equi noncatalytic protein NCP1 (34). These two CBM29-1 and CBM29-2 bind to a range of β-1,4-linked polysaccharides, including cellulose, mannan, xylan, and glucosmanan. Structural studies of CBM29-2 bound to cellohexaose and to mannohexaose reveal very few direct hydrogen bonds, particularly between the protein and the C-2 hydroxyl of the ligand (35, 36). Thus, β-1,4-linked polysaccharides with different configurations of the C-2 hydroxyl can be accommodated at the binding site of CBM29s as these hydroxyl groups do not significantly contribute to binding. Our co-crystal structures of TopCBM16-1 bound to polysaccharides illustrate an alternative mechanism for promiscuous ligand binding, namely compensatory changes in the hydrogen bonding between the C-2 hydroxyl and the protein to accommodate differences in configuration of the carbohydrate. Thus, whereas both CBM29 and TpolCBM16 demonstrate promiscuity in ligand specificity, each utilizes a distinct strategy for achieving this plasticity.

The ubiquitous nature of carbohydrate-binding modules that share an overall common fold suggests that this scaffold is one that effectively optimized polysaccharide binding. Subtle differences in the ligand-binding site have resulted in the evolution of a number of different polypeptides capable of engaging a wide diversity of polysaccharides utilizing common core elements. Many such carbohydrate-binding modules are teth-ered to a catalytic domain, and this modular architecture serves to enhance the efficacy by targeting the catalytic components toward their substrate through interactions mediated by the carbohydrate-binding domain. The modular nature of these enzymes suggests that the combinatorial assembly of different catalytic and carbohydrate-binding domains can result in the production of novel activities. Our work provides a foundation for understanding the ligand specificity of the TpolCBM16s, and these insights may be utilized in altering specificity through minor changes in the protein scaffold. This rational redesign approach can be utilized to expand the natural repertoire of enzymes that can degrade complex polysaccharides into simple sugars, and such an approach will have a broad impact in the biofuels industry.

Acknowledgments—We thank Lisa Keefe and the staff at IMCA-CAT (17-ID on Argonne National Laboratories, Advanced Photon Source) and John Chrzas and staff at SER-CAT (22-BM at Argonne National Laboratories, Advanced Photon Source) for facilitating data collection. We also thank Joseph Brunzel at LS-CAT (23-ID at Argonne National Laboratories, Advanced Photon Source) for assistance with data collection.

REFERENCES

1. Coughlan, M. P. (1985) Biochem. Soc. Trans. 13, 405–406
2. Ahsan, M. M., Kimura, T., Karita, S., Sakka, K., and Ohmiya, K. (1996) J. Bacteriol. 178, 5732–5740
3. Bolam, D. N., Xie, H., White, P., Simpson, P. J., Hancock, S. M., Williamson, M. P., and Gilbert, H. J. (2001) Biochemistry 40, 2468–2477
4. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Biochem. J. 382, 769–781
5. Linder, M., Salovuo, I., Ruohonен, L., and Teeri, T. T. (1996) J. Biol. Chem. 271, 21268–21272
6. Boraston, A. B., Kwan, E., Chiu, P., Warren, R. A., and Kilburn, D. G. (2003) J. Biol. Chem. 278, 6120–6127
7. Sunna, A., Gibbs, M. D., and Bergquist, P. L. (2001) Biochem. J. 356, 791–798
8. Ito, Y., Tomita, T., Roy, N., Nakano, A., Sugawara-Tomita, N., Watanabe, S., Okai, N., Abe, N., and Kamio, Y. (2003) Appl. Environ. Microbiol. 69, 6969–6978
9. Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998) Biochem. J. 331, 775–781
10. Walden, H., Taylor, G. L., Lorentzen, E., Pohl, E., Lilie, H., Schramm, A., Knura, T., Stubbe, K., Tjaden, B., and Hensel, R. (2004) J. Mol. Biol. 342, 861–875
11. Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Theodorou, M. K., Morris, P., and Gilbert, H. J. (1996) Biochem. J. 319, 515–520
12. Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A., and Kilburn, D. G. (1996) Mol. Microbiol. 11, 747–755
13. Southall, S. M., Simpson, P. J., Gilbert, H. J., Williamson, G., and Williamson, M. P. (1999) FEBS Lett. 447, 58–60
14. Cann, J. K., Kocherginskaya, S., King, M. R., White, B. A., and Mackie, R. I. (1999) J. Bacteriol. 181, 1643–1651
15. Van Duyne, G. D., Standeart, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) J. Mol. Biol. 229, 105–124
16. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1991) Biochemistry 30, 2085–2092
17. Otwinowski, Z., Borek, D., Majoros, W., and Minor, W. (2003) Acta Crystallogr. Sect. A 59, 228–234
18. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
19. Terwilliger, T. C. (2003) Methods Enzymol. 374, 22–37
20. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D Biol.
Crystallogr. 55, 849–861
21. Terwilliger, T. C. (2002) Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 1937–1940
22. McRee, D. E. (1999) J. Struct. Biol. 125, 156–165
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 53, 240–255
24. Vagin, A., and Teplyakov, A. (2000) Acta Crystallogr. Sect. D Biol. Crystallogr. 56, 1622–1624
25. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 1445–1450
26. Brunger, A. T. (1993) Acta Crystallogr. Sect. D Biol. Crystallogr. 49, 24–36
27. Kleywegt, G. J., and Brunger, A. T. (1996) Structure (Lond.) 4, 897–904
28. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
29. 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
30. Charnock, S. J., Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M., Davies, G. J., and Fontes, C. M. (2000) Biochemistry 39, 5013–5021
31. Mizushima, T., Hiarto, T., Yoshida, Y., Lee, S. J., Chiba, T., Iwai, K., Yamaguchi, Y., Kato, K., Tsukihara, T., and Tanaka, K. (2004) Nat. Struct. Mol. Biol. 11, 365–370
32. Young, S. H., Dong, W. J., and Jacobs, R. R. (2000) J. Biol. Chem. 275, 11874–11879
33. Saito, H., Yoshioka, Y., Uehara, N., Aketagawa, J., Tanaka, S., and Shibata, Y. (1991) Carbohydr. Res. 217, 181–190
34. Freelove, A. C. J., Bolam, D. N., White, P., Hazlewood, G. P., and Gilbert, H. J. (2001) J. Biol. Chem. 276, 43010–43017
35. Charnock, S. J., Bolam, D. N., Nurizzo, D., Szabo, L., McKie, V. A., Gilbert, H. J., and Davies, G. J. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 14077–14082
36. Flint, J., Bolam, D. N., Nurizzo, D., Taylor, E. J., Williamson, M. P., Walters, C., Davies, G. J., and Gilbert, H. J. (2005) J. Biol. Chem. 280, 23718–23726