Specificity and Affinity of Substrate Binding by a Family 17 Carbohydrate-Binding Module from *Clostridium cellulovorans* Cellulase 5A†

Alisdair B. Boraston,‡,§,|| Patrick Chiu,§,|| R. Antony J. Warren,‡,§ and Douglas G. Kilburn*,‡,§,||

Protein Engineering Network of Centres of Excellence, Department of Microbiology and Immunology, and The Biotechnology Laboratory, University of British Columbia, Vancouver, British Columbia, Canada

Received April 4, 2000

ABSTRACT: The C-terminal carbohydrate-binding module (CBM17) from *Clostridium cellulovorans* cellulase 5A is a β-1,4-glucan binding module with a preference for soluble chains. CBM17 binds to phosphoric acid swollen Avicel (PASA) and Avicel with association constants of 2.9 ($\pm$0.2) $\times$ 10^5 and 1.6 ($\pm$0.2) $\times$ 10^5 M^{-1}, respectively. The capacity values for PASA and Avicel were 11.9 and 0.4 μmol/g of cellulose, respectively. CBM17 did not bind to crystalline cellulose. CBM17 bound tightly to soluble barley β-glucan and the derivatized celluloses HEC, EHEC, and CMC. The association constants for binding to barley β-glucan, HEC, and EHEC were approximately 2.0 $\times$ 10^5 M^{-1}. Significant binding affinities were found for cello-oligosaccharides greater than three glucose units in length. The affinities for cellotriose, cellotetraose, cellopentaose, and cellohexaose were 1.2 ($\pm$0.3) $\times$ 10^3, 4.3 ($\pm$0.4) $\times$ 10^3, 3.8 ($\pm$0.1) $\times$ 10^4, and 1.5 ($\pm$0.0) $\times$ 10^5 M^{-1}, respectively. Fluorescence quenching studies and *N*-bromosuccinimide modification indicate the participation of tryptophan residues in ligand binding. The possible architecture of the ligand-binding site is discussed in terms of its binding specificity, affinity, and the participation of tryptophan residues.

Many organisms have evolved diverse polysaccharolytic enzyme systems to take advantage of the abundance of plant biomass as an energy source (*1*). Most of these enzymes are modular, comprising one or more catalytic modules and, often, one to several accessory modules (*2*). One of the best-studied kinds of accessory modules is the carbohydrate-binding modules (CBMs),[1] previously called cellulose-binding domains (CBDs). The CBMs have been classified into ~20 families based on amino acid similarity, with at least 1 member of each family having demonstrated function (*3*−*5*). Examples from families 1−5 have been extensively functionally characterized and their structures determined (*6*−*10*). All bind to β1−4-linked glucose polymers (cellulose), and some bind to chitin (*11*−*15*). The examples from families 1, 2a, 3, and 5 bind most effectively to crystalline cellulose. The N-terminal family 4 CBM from Cellulase 9B of *Cellulomonas fimi*, formerly referred to as CBD_{N1} but now called CBM4-1, binds preferentially to amorphous cellulose and soluble cellooligosaccharides (*14*). It is the only currently characterized CBM to have such a specificity.

Clostridium cellulovorans hydrolyzes cellulose through the action of cellulosomal cellulases and free cellulases (*16*, *17*). One such free cellulase is Cellulase 5A, or Cel5A, formerly known as endoglucanase F, which has maximal activity on soluble glucans (*18*). This enzyme has a C-terminal module with cellulose-binding activity comprising approximately the last 200 amino acids of Cel5A (*18*). This module, which shows no similarity to previously defined CBM families, has recently been classified as a new family of CBM, family 17 (*5*).

A complete understanding of the substrate-binding properties of its CBM provides insight into the target and mode of action of the complete enzyme. In this study, we examine the specificity and affinity of substrate binding by isolated CBM17, the CBM of Cel5A. Based on these results, a putative model of the substrate-binding site of CBM17 is suggested and the possible role of CBM17 in substrate degradation by Cel5A is discussed.

MATERIALS AND METHODS

Carbohydrates and Polysaccharides. Microcrystalline cellulose (Avicel PH101) was obtained from FMC International (Little Island, County Cork, Ireland). Bacterial microcrystalline cellulose (BMCC) was prepared from cultures of *Acetobacter xylinum* (ATCC 23769) as described previously (*12*). Regenerated cellulose (PASA) was obtained by phosphoric acid treatment of Avicel as reported previously (*19*). Mannose (mixed anomers), D-glucose, *N*-acetyl-D-glucosamine, D-galactose, D-xylose, fructose, sucrose, cellobiose, lactose, shrimp chitin, locust-bean gum (galactomannan),

† This work was supported by grants from the Protein Engineering Network of Centres of Excellence, CBD Technologies Inc., and the Natural Sciences and Engineering Research Council of Canada (D.G.K.).

* To whom correspondence should be addressed at Biotechnology Laboratory, #237-6174 University Blvd., University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3. Phone: (604) 822-5115; Fax: (604) 822-2114; e-mail: kilburn@interchange.ubc.ca.

‡ Protein Engineering Network of Centres of Excellence.

§ Department of Microbiology and Immunology.

|| The Biotechnology Laboratory.

[1] Abbreviations: BMCC, bacterial microcrystalline cellulose; CBM, carbohydrate-binding module; CBD, cellulose-binding domain; EHEC, ethylhydroxyethyl-cellulose; HEC, hydroxyethyl-cellulose; IPTG, isopropyl-1-thio-β-D-galactopyranoside; K_a, association constant; MALDI-TOF, matrix-assisted laser-desorption time-of-flight; N_o, binding capacity; NBS, *N*-bromosuccinimide; PASA, phosphoric acid swollen Avicel.

10.1021/bi0007728 CCC: $19.00
Published on Web 08/17/2000

dextran, and arabinogalactan were purchased from Sigma Chemical Co. Hydroxyethyl-cellulose (HEC) was purchased from Aldrich Chemical Co. Maltose was obtained from Baker Chemical Corp. Barley β-glucan and pectic galactan were purchased from MegaZyme Ltd. (North Rocks, N.S.W., Australia). Cello-oligosaccharides were obtained from Seikagaku (Tokyo, Japan). Birchwood Xylan (Roth 7500; MW ~25 000) was obtained from Carl Roth RG (Karlsruhe, Germany). Water-soluble xylan was prepared according to previously described procedures (*20*).

DNA Amplification and Cloning. A gene fragment encoding 208 amino acid residues, 204 of which are the C-terminal substrate-binding domain CBM17 from *Clostridium cellulovorans* Cellulase 5A (GenBank accession number U37056), was obtained and amplified by PCR from pEF 32 (*18*). Appropriate restriction sites were introduced at the 5′ and 3′ ends of the *CBM17* gene fragment for cloning into the pET28a expression vector (Novagen, Milwaukee, MI). Each PCR mixture (50 μL total) contained 25−50 ng of pEF 32, 25−50 pmol of primers, 10% DMSO, 0.4 mM 2′-deoxynucleoside 5′-triphosphates, and 1 unit of *Pwo* DNA polymerase in buffer (Boehringer-Mannheim, Laval, Quebec). A protocol of 20 successive cycles of denaturation at 96 °C for 1 min, annealing at 55 °C for 30 s, and primer extension at 72 °C for 1.5 min was followed. An *Nco*I site (underlined) was introduced at the 5′ end of the *CBM17* gene fragment, using the oligonucleotide: 5′-GATATACCATGGGCGCTAGCGGTCAATATGTACGTGCTCG-3′ as primer. A *Hin*dIII (underlined) restriction site was introduced at the 3′ end of the CBM17 sequence using the oligonucleotide: 5′-TGCAGTGCGGCCGCAAGCTTTTATTATTTAGTAAACTTTATA-3′ as primer. The resulting 0.6 kb PCR fragment was digested with *Nco*I and *Hin*dIII and ligated into pET28a previously digested with the same restriction enzymes, to give pET-CBM17. DNA was sequenced by the NAPS Unit (Biotechnology Laboratory, University of British Columbia) using the AmpliTaq dye termination cycle sequencing protocol and an Applied Biosystems Model 377 sequencer.

Protein Purification. Overnight cultures of *E. coli* strain BL21/pET-CBM17 were diluted 100-fold in tryptone−yeast extract−phosphate medium (TYP) (*21*) supplemented with 50 μg of kanamycin/mL, and grown at 30 °C to a cell density ($A_{600\ nm}$) of ~0.3. Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.3 mM. Incubation was continued for a further 6 h at 30 °C. The cells were harvested by centrifugation (8500*g*) for 10 min at 4 °C and resuspended to about 1/50th of the original culture volume by gentle mixing in 50 mM potassium phosphate, pH 7.0. Cells were ruptured by two passages through a French pressure cell (21 000 lb/in^2), and cell debris was removed by centrifugation for 30 min at 27000*g* and 4 °C. CBM17 was purified from the clarified cell extract by cellulose affinity chromatography as described below.

A column (2.5 × 10 cm) was packed with 40 mL of a 50% slurry of Perloza MT 500 resin (Novochemie, Czech Republic) to give a final bed volume of ~ 20 mL. All subsequent operations were done at room temperature and at flow-rates of 5 mL min^{-1} using a BioLogic LP system (BioRad, Mississauga, Ontario). The column was equilibrated with 100 mL of binding buffer (50 mM potassium phosphate, pH 7.0). Fifty milliliters of clarified cell extract, diluted 1:2 in binding buffer, was loaded on the column. Unbound protein was removed by washing the column with 10 column volumes of binding buffer containing 1 M NaCl followed by 10 column volumes of binding buffer. The adsorbed proteins were recovered by desorption with distilled water. Protein fractions were analyzed for purity on 16% SDS−PA gels. Pure CBM17 fractions were pooled and further exchanged into distilled water and concentrated in a stirred ultrafiltration unit (Amicon, Beverly, MA) on a 1K cutoff filter (Filtron, Northborough, MA).

Determination of Protein Concentration. The concentration of purified protein was determined by UV absorbance (280 nm) using a calculated molar extinction coefficient (*22*) of 31 010 M^{-1} cm^{-1}.

Fluorescence Analysis of Substrate Binding. All fluorescence measurements were performed on a Perkin-Elmer LS-50 luminescence spectrometer (Perkin-Elmer, Norwalk, CT) with a four-sample cuvette holder thermostated by a recirculating heating/cooling bath (Forma Scientific, Marieta, OH).

Emission scans were performed using 2 μM CBM17 in the presence or absence of 1 mM ligand in 25 mM Tris-HCl buffer, pH 7.5. The excitation wavelength was 280 nm. Emission intensities were collected over the wavelength range of 300−400 nm. The excitation and emission slit widths were 5 nm. Five scans were averaged.

Quantitative binding experiments were performed by titration of the appropriate carbohydrate in 25 mM Tris-HCl buffer, pH 7.5, into 2.0 mL of CBM17 (2 μM in the same buffer) under continuous stirring. The temperature of the experiment was controlled by adjusting the recirculating water bath, and the temperature of the sample was monitored in the cuvette with a 0.8 mm thermistor temperature probe (Cole-Parmer, Vernon Hills, IL) connected to an external display. Amounts of ligand were delivered from Hamilton syringes by a computer-controlled micrometer driven syringe pump. After each ligand addition, the incubation mixture was stirred for 90 s to allow the reaction to reach equilibrium. The fluorescence intensity was measured using an excitation wavelength of 280 nm and emission wavelengths of 335 and 350 nm. Emission and excitation slit widths were 4 nm. Quadruplicate readings with 2 s integration periods were averaged for each data point. The emission intensities were corrected for background fluorescence caused by buffer and carbohydrates, for dilution, and for inner filter effects (*23*) as required. Plots of relative fluorescence (F/F_o) against total carbohydrate concentration were constructed. The association binding constants (K_a) and the maximal fluorescence change (F_{max}) in the protein upon full complexation with sugar ligand were derived by a nonlinear least-squares fit of the corrected data to a one-site binding model with ligand depletion (*23*) using Origin v.5.0 (Microcal, Northhampton, MA) software.

UV Analysis of Substrate Binding. Absorbance spectra were collected on a Cary 100e UV−Vis spectrophotometer with a thermostated sample block (Varian, Melbourne, Australia). All spectra were collected with integration times of 0.5 s per reading and an acquisition interval of 0.5 nm from 250 to 320 nm. Experiments were conducted at 20 °C in 25 mM Tris-HCl, pH 7.5.

Difference spectra were obtained by first collecting a baseline spectrum on 800 μL of 20 μM CBM17. A second scan of 800 μL of 20 μM CBM17 plus 1 mM ligand was acquired. Quantitative binding experiments were performed

by collecting scans of CBM17 after the sequential addition of ligand to 40 μM CBM17 under continuous stirring. The samples were allowed to equilibrate for 1 min after addition of the ligand. The baseline was collected using 25 mM Tris-HCl, pH 7.5. Each scan was corrected for dilution by multiplying with a correction factor [$(\Sigma V_t + V_i)/V_i$, where V_t is the titration volume and V_i is the initial volume]. Scans were normalized by subtraction of the scan at zero ligand concentration. The trough to peak heights at $\Delta A_{290} - \Delta A_{295}$ nm and $\Delta A_{280} - \Delta A_{287}$ nm were plotted versus the total ligand concentration. Association constants were derived by nonlinear regression of the data using a one-site model taking into account ligand depletion as described above.

Stoichiometries of binding were determined by plotting the trough to peak heights at $\Delta A_{290} - \Delta A_{295}$ nm and $\Delta A_{280} - \Delta A_{287}$ nm against the molar ratio of ligand added to total CBM ([L]/[CBM17]). Each titration extended well into the saturation region of the binding curve. Linear regression was used to fit lines to the initial linear portion of the binding curves and the linear portion of the curves at saturation. The value corresponding to the point on the abscissa at the intersection of these lines was taken as the stoichiometric limit of the binding reaction.

N-Bromosuccinimide Modification. The rate of *N*-bromosuccinimide (NBS) modification was monitored by the loss of absorbance at 280 nm using stopped-flow UV absorbance spectroscopy. An S2000 CCD array spectrometer with a 1000 kHz data acquisition card (Ocean Optics, Dunedin, FL) was used for data collection. Stopped-flow experiments were performed with an SFA-20 stopped-flow accessory (Hi-Tech Scientific, Salisbury, U.K.). A molar ratio of 2:1 NBS to CBM17 was used with both reagents in 50 mM sodium acetate, pH 5.2. Experiments were performed without ligand present or with excess cellohexaose (500 μM). Ten data sets were collected and fit simultaneously using nonlinear regression to a two-phase exponential decay function to obtain apparent rate constants (k_{app}).

Solute Quenching. Solute quenching experiments were performed using potassium iodide (KI) and 2 μM CBM17. Ionic strength was kept constant by the addition of NaCl to samples such that the final concentration of salt (total KI and NaCl concentration) was 1 M. Samples were buffered with 25 mM Tris-HCl buffer, pH 7.5. The excitation wavelength was 295 nm. Emission intensities were collected at 350 nm with integration times of 2 s. The excitation and emission slit widths were 5 nm. Four readings were averaged. Data were plotted as $F_o/(F_o - F)$ against 1/[KI] (*24*). The quenching constants (K_{sv}) and fractions of accessible fluorophore (f) were calculated as the inverse of the slopes and ordinate-intercepts, respectively. Samples containing ligand contained 800 μM cellopentaose or 500 μM cellohexaose.

Affinity Electrophoresis. Qualitative binding of CBM17 (5 μg) to soluble polysaccharides was evaluated by affinity electrophoresis (*14*, *25*) in 10% polyacrylamide gels polymerized in the absence or presence of various polysaccharides. Soybean trypsin inhibitor (10 μg) was used as an internal standard. To prevent it from running off the gel, the standard was loaded after 45 min of electrophoresis. Electrophoresis was for 1.5 h on ice, pH 8.8 and 150 V, in a Mini PROTEAN II system (BioRad). After electrophoresis, the gels were stained with Coomassie blue. The migration distances of CBM17 and reference proteins were measured directly on the gels. Interaction of CBM17 with low molecular weight ligands was assessed by competitive affinity electrophoresis (*25*) in gels containing both barley β-glucan (0.25 g/L) and a competing mono- or disaccharide (0.1 M).

Adsorption Assays on Insoluble Polysaccharides. Binding constants were determined by depletion isotherms (*12*, *26*). Duplicate samples of CBM at concentrations ranging from 0.5 to 15 μM were incubated in 1.5 mL tubes with 1−10 mg of insoluble cellulose in a final volume of 1 mL. Duplicate control tubes contained no cellulose. Samples were rotated for 4 h at 4 °C. The cellulose was removed from solution by centrifugation at 13 000 rpm for 20 min at 4 °C in a drum rotor. Supernatants containing unbound material were transferred to a UV-transparent 96 well plate and absorbance measurements taken at 280 and 350 nm in a SpectraMax Plus UV plate reader with path length correction (Molecular Devices, Sunnyvale, CA). The net A_{280} ($A_{280} - A_{350}$; where A_{350} was used to approximate scattering due to particulates) was used to calculate the free CBM concentration. The equilibrium association constants (K_a) were determined by nonlinear regression as described previously (*27*).

N-Terminal Sequencing and MALDI-TOF Mass Spectrometry. N-Terminal sequencing of purified CBM17 was performed by the Nucleic Acid and Protein Services Unit (Biotechnology Laboratory, University of British Columbia). The mass of purified CBM17 was determined by matrix-assisted laser-desorption time-of-flight (MALDI-TOF) mass spectrometry. Purified CBM17, drop-dialyzed overnight into distilled water using a 0.025 μm vs-membrane (Millipore, Bedford, MA), was diluted in distilled water to 40 pmol/μL and mixed 1:1 with a saturated matrix solution of sinapinic acid in 70% acetonitrile/0.1% trifluoroacetic acid. One microliter was spotted onto the MALDI target and allowed to dry. The mass was obtained by positive ion MALDI-TOF mass spectrometry on a SELDI-MassPhoresis system (Ciphergen, Palo-Alto, CA). β-Lactoglobulin (18 363.3 Da) and horseradish peroxidase (42 240.0 Da) were used to calibrate the machine.

RESULTS

Polypeptide Production. CBM17 could be purified to homogeneity from cleared *E. coli* extracts by cellulose affinity chromatography using distilled water to desorb the polypeptide. The polypeptide was judged by SDS−PAGE to be greater than 95% pure (results not shown). MALDI-TOF mass spectrometry showed a homogeneous population of polypeptide with a mass of 20 287.6 daltons (results not shown). N-Terminal sequencing determined that the N-terminus of the polypeptide was truncated beginning with tyrosine 16 (corresponding to residue 1 in Figure 6), giving a polypeptide of 192 amino acids. The mass of CBM17 with this N-terminus is predicted to be 20 364.3 daltons, within 0.4% of the mass determined by mass spectrometry.

Saccharide Binding. The soluble polysaccharide binding characteristics of CBM17 were initially assessed by affinity electrophoresis. CBM17 bound tightly to barley β-glucan, hydroxyethyl-cellulose (HEC), ethylhydroxyethyl-cellulose (EHEC), and weakly to carboxymethyl-cellulose (results not shown). It did not bind to birchwood xylan, arabinogalactan, pectic-galactan, galactomannan (locust-bean gum), or dextran (T-70).

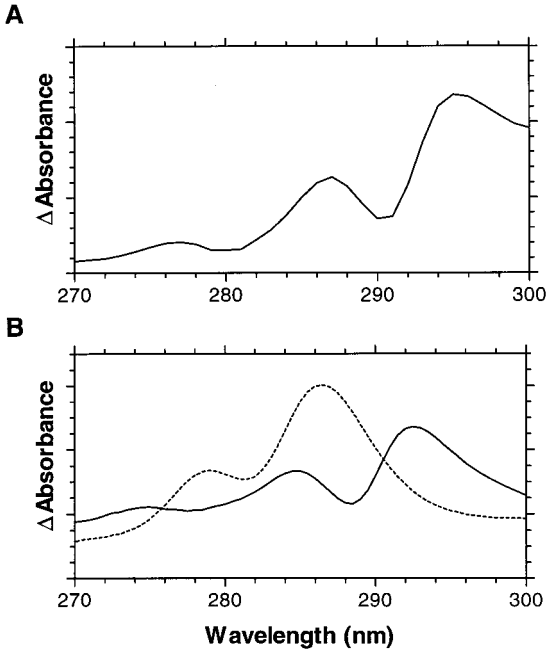

FIGURE 1: UV difference spectrum of CBM17 complexed with cellohexaose. Panel A: difference spectrum of CBM17 with 122 μM cellohexaose relative to uncomplexed CBM17. Panel B: difference spectrum of *N*-acetyltryptophan (solid line) and *N*-acetyltyrosine (dotted line) perturbed in 20% DMSO. Samples containing 50 μM *N*-acetyltryptophan or 100 μM *N*-acetyltyrosine in 25 mM Tris-HCl, pH 7.5, were used to collect baseline spectra. Samples of the same concentrations but containing 20% DMSO were used to collect difference spectra.

Competition affinity electrophoresis using barley β-glucan as the polysaccharide ligand was used to screen the binding of CBM17 to a variety of mono- and disaccharides. None of the sugars tested, which included xylose, glucose, galactose, mannose, *N*-acetylglucosamine, fructose, sucrose, lactose, maltose, and cellobiose, inhibited the interaction of CBM17 with barley β-glucan when included in the affinity gels at concentrations of 0.1 M.

The complexation of oligosaccharides with CBM17 resulted in perturbations in both the intrinsic fluorescence emission spectrum and the UV difference absorption spectrum of CBM17. The binding of cellohexaose to CBM17 caused a slight blue-shift in the fluorescence emission maximum and a reduction in the emission intensity (result not shown). The change in the UV absorption spectrum of CBM17 was quite pronounced upon cellohexaose binding as indicated by large changes in the difference spectrum (relative to uncomplexed CBM17) at 295, 290, 287, and 280 nm (Figure 1).

Both fluorescence and UV spectroscopy could be used to monitor the binding of CBM17 upon titration of soluble sugars (Figure 2). CBM17 bound to cello-oligosaccharides with a 1:1 stoichiometry, indicating a single binding site on CBM17. Subsequently, both the fluorescence and UV data were used to fit a one-site binding model with ligand depletion using nonlinear regression. The association constants increased from 1.2 ($\pm$0.3) $\times$ 10^3 M^{-1} for cellotriose to 1.5 ($\pm$0.0) $\times$ 10^5 M^{-1} for cellohexaose (Table 1). No binding to glucose or cellobiose could be detected. For determining the stoichiometries of binding to polysaccharides, the chain sizes of which are often difficult to obtain, the concentration unit of weight/volume of polysaccharide

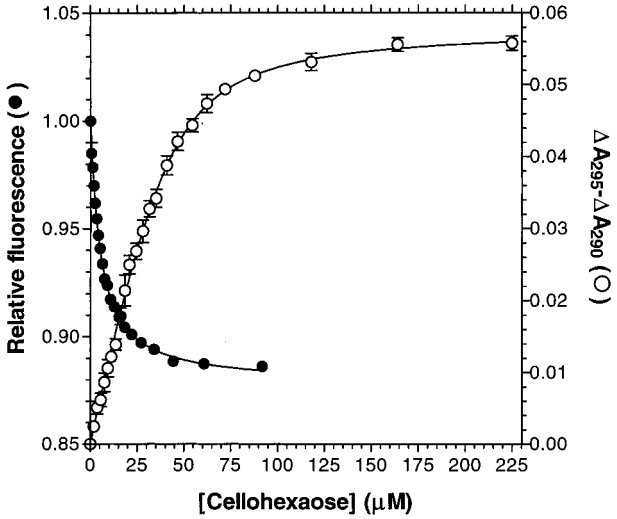

FIGURE 2: Spectrophotometric monitoring of CBM17 titrated with cellohexaose. An example of fluorescence titration (●) and UV difference titration (○) is shown. Solid lines represent best-fit lines of a one-site binding model with ligand depletion. Error bars show the standard deviation of quadriplicate readings.

was used. The stoichiometric limit of the binding reaction was then determined as moles of CBM17 per mass of polysaccharide (i.e., moles of binding site per gram of polysaccharide). The concentration of polysaccharide was converted to reflect this value, and the data were fit to a one-site binding model with ligand depletion using nonlinear regression. The association constants of CBM17 binding to barley β-glucan, HEC, and EHEC were all found to be $\sim$2 $\times$ 10^5 M^{-1} (Table 1). Based on an average barley β-glucan chain length of 95 glucose units (*14*), the stoichiometry of binding was estimated to be 3 mol of CBM17 per mole of barley β-glucan, similar to the 4:1 stoichiometry of CBM4-1 binding to barley β-glucan (*14*).

Depletion binding isotherms were used to characterize the binding of CBM17 to insoluble β1$-$4 glucans. CBM17 did not bind to highly crystalline cellulose (bacterial microcrystalline cellulose, BMCC) or to chitin. CBM17 had relatively high affinities for amorphous cellulose (PASA) and microcrystalline cellulose (Avicel) (Table 1), with capacities (N_o) of 11.9 and 0.4 μmol/g of cellulose, respectively. Binding to these substrates was reversible (results not shown).

The change in Gibbs free energy (ΔG) varied linearly, becoming more favorable, with polysaccharide chain length increasing up to six glucose units (Figure 3). The ΔG did not change with the longer soluble or insoluble polysaccharides.

Involvement of Aromatic Residues in Substrate Binding. Solute quenching studies of CBM17 were performed using an excitation wavelength of 295 nm to selectively excite tryptophan residues. Plots of $F_o/(F_o - F)$ versus 1/[KI] were linear (Figure 4). The quenching constants (K_{sv}) were 3.9, 3.0, and 2.7 M^{-1} for buffer-only, plus cellopentaose, and plus cellohexaose, respectively. The fractions of KI-accessible fluorophore (f) also decreased in the presence of cello-oligosaccharide from 0.9 for buffer-only to 0.7 and 0.7 for plus cellopentaose and plus cellohexaose, respectively. The errors for K_{sv} and f values were less than 2%. Because of the large size of hydrated KI, it is limited to quenching exposed aromatic residues (*24*). The reductions in the K_{sv}

Table 1: Specificity and Affinity of Sugar Binding by CBM17[a]

	method			
	fluorescence		UV	
substrate	K_a (M^{-1})	ΔG (kJ/mol)	K_a (M^{-1})	ΔG (kJ/mol)
cellotriose	$1.2 (\pm 0.3) \times 10^3$	−17.3	−	−
cellotetraose	$4.3 (\pm 0.4) \times 10^3$	−20.4	−	−
cellopentaose	$3.8 (\pm 0.1) \times 10^4$	−25.7	$3.4 (\pm 0.2) \times 10^4$	−25.4
cellohexaose	$1.5 (\pm 0.0) \times 10^5$	−29.0	$1.6 (\pm 0.1) \times 10^5$	−29.2
barley β-glucan	$1.8 (\pm 0.0) \times 10^5$	−29.5	$1.6 (\pm 0.1) \times 10^5$	−29.2
HEC	$0.8 (\pm 0.0) \times 10^5$	−27.5	$1.0 (\pm 0.1) \times 10^5$	−28.1
EHEC	$0.6 (\pm 0.0) \times 10^5$	−26.8	$2.1 (\pm 0.4) \times 10^5$	−29.9
PASA	−	−	$2.9 (\pm 0.2) \times 10^5$ [b]	−30.7
Avicel	−	−	$1.6 (\pm 0.2) \times 10^5$ [b]	−29.2

[a] All binding experiments were performed at 25 °C (298 K). [b] Determined by depletion isotherms.

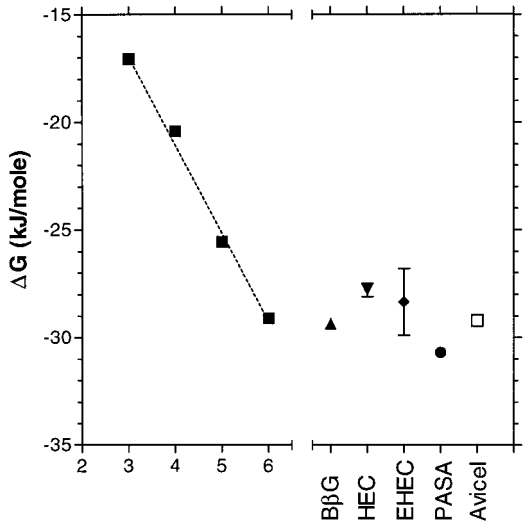

FIGURE 3: Change in Gibb's free energy (ΔG) with substrate length. The left portion of the separated *X*-axis shows the number of glucose units in the cello-oligosaccharide (■). The dotted line shows the linear trend in ΔG with oligosaccharide length. The right portion of the *X*-axis shows substrates with larger and undefined chain lengths. BβG, barely β-glucan (▲); HEC, hydroxyethyl-cellulose (▼); EHEC, ethylhydroxyethyl-cellulose (◆); PASA, phosphoric acid swollen Avicel (●); Avicel, microcrystalline cellulose (□). Error bars show the standard error of values determined by fluorescence and UV for substrates where both techniques were used.

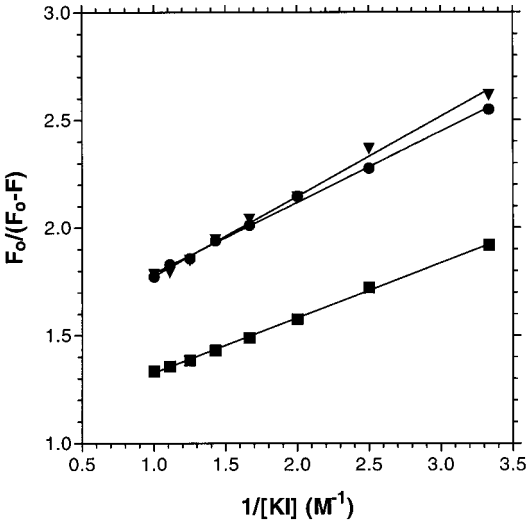

FIGURE 4: Solute quenching of CBM17 in the absence (■) and presence of cellopentaose (800 μM) (●) or cellohexaose (500 μM) (▼). Fluorescence emission intensities at 350 nm, with an excitation wavelength of 295 nm, were collected for 2 μM CBM17 in increasing concentrations of KI. Straight lines show best-fit lines determined by linear regression of the data. The inverse of the slopes and ordinate-intercepts give the K_{sv} and *f* values, respectively (see text).

values and *f* values indicate that exposed tryptophan residues become less accessible to quencher in the presence of cello-oligosaccharide.

The rate of tryptophan modification by *N*-bromosuccinimide (NBS) was monitored by stopped-flow spectroscopy in the presence and absence of cellohexaose. A two-phase exponential decay function was required to adequately fit the reaction traces, indicating at least two classes of tryptophan with different reactivity toward NBS. The two apparent rate constants (k_{app}, s^{-1}) in the presence of cellohexaose were half of the values calculated in the absence of cellohexaose (Table 2). This shows that the presence of ligand slows the reactivity of one or more tryptophans toward NBS. These results suggest that one or more tryptophans is (are) shielded by the bound ligand, which is consistent with the fluorescence quenching results.

The UV difference spectrum of CBM17 complexed with cellohexaose resembles the UV difference spectrum of solvent-perturbed *N*-acetyltryptophan much more closely than that of solvent-perturbed *N*-acetyltyrosine (Figure 1). The

Table 2: Kinetics of CBM17 Tryptophan Modification by NBS

	$k1_{app}$ (s^{-1})	$t1_{1/2}$ (s)	$k2_{app}$ (s^{-1})	$t2_{1/2}$ (s)
CBM17	16.92 ± 1.09	0.04	0.85 ± 0.03	0.81
CBM17 + 200 μM G6[a]	7.27 ± 1.24	0.09	0.45 ± 0.03	1.53

[a] G6, cellohexaose.

solvent exposure of tyrosine residues, as examined by second-derivative UV spectroscopy (*28*), was unchanged upon ligand binding (results not shown). These results implicate the involvement of tryptophans in cello-oligosaccharide binding by CBM17 with likely little or no involvement of tyrosines. Furthermore, the maximum change in fluorescence intensity (equal to $100 \times [(1 - F_\infty)/F_o]$, where F_∞ is the relative fluorescence at saturating ligand concentration) in the presence of ligand increased with chain length: 2% for cellotriose, 9% for cellotetraose, 12% for cellopentaose, and 14% for cellohexaose, barley β-glucan, EHEC, and HEC. This indicates greater involvement of aromatic residues in the binding of CBM17 to longer oligosaccharides.

Effects of Salt on Oligosaccharide Binding. The binding of CBM17 to cellopentaose and cellohexaose was enhanced by the presence of NaCl, but binding to cellotetraose was

Table 3: Effects of NaCl Concentration on Oligosaccharide Binding by CBM17[a]

	buffer only		0.5 M NaCl		1.0 M NaCl		2.0 M NaCl	
	$K_a \times 10^{-3}$ (M^{-1})	ΔG (kJ/mol)	$K_a \times 10^{-3}$ (M^{-1})	ΔG (kJ/mol)	$K_a \times 10^{-3}$ (M^{-1})	ΔG (kJ/mol)	$K_a \times 10^{-3}$ (M^{-1})	ΔG (kJ/mol)
cellotetraose	3.3 (±0.1)	−20.1	ND	ND	ND	ND	4.3 (±0.2)	−20.8
cellopentaose	84.8 (±0.4)	−28.1	119.8 (±0.5)	−29	148.2 (±0.9)	−29.5	200.8 (±1.0)	−30.3
cellohexaose	113.2 (±0.2)	−28.8	143.2 (±0.3)	−29.4	189.7 (±0.5)	−30.1	228.8 (±0.5)	−30.6

[a] All samples were buffered in 25 mM Tris-HCl, pH 7.5. Experiments were performed at 25 °C (298 K). ND, not determined.

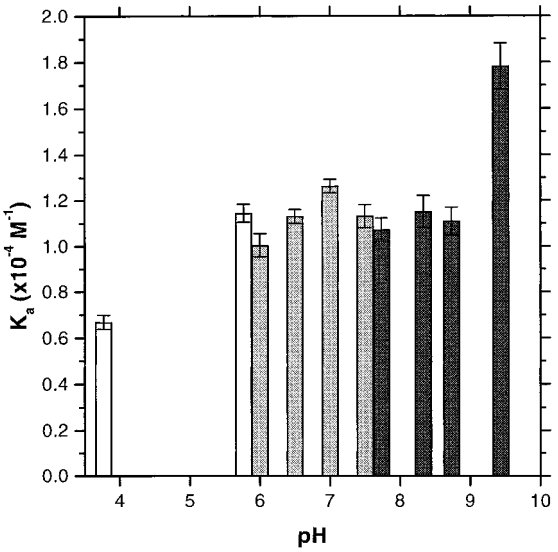

FIGURE 5: pH dependence of CBM17 binding to cellohexaose. Fluorescence titration binding experiments were performed in 25 mM citrate (white bars), 25 mM potassium phosphate (light gray bars), and 25 mM Tris (dark gray bars).

less affected (Table 3). This is consistent with tryptophan residues contributing a hydrophobic component to substrate binding although a counterion effect cannot be ruled out. This shift in the equilibrium at lower ionic strength may explain the ability to desorb CBM17 from cellulose with water.

Effects of pH on Oligosaccharide Binding. The K_as for cellohexaose were relatively constant between pH 6.0 and 8.5 (Figure 5). No binding was detected at pHs below 4.0. The K_a improved by ~50% at pH 9.2. The K_a values for titrations performed in the two buffer systems at the transition regions were nearly identical, indicating no effects on binding due to the buffer composition. This ability to function at relatively high pH is consistent with the frequent occurrence of family 17 CBMs in alkalophilic *Bacillus* species (Figure 6).

DISCUSSION

The participation of aromatic amino acid residues in substrate binding by lectins is nearly universal (*29*). In particular, tryptophans are most common. Structural studies have shown these residues to interact directly with the pyranose ring in what has been described as a 'stacking' interaction, although the relative angles of the aromatic residue to the pyranose ring vary from 17° to 52°, rather than parallel (*29*). Aromatic amino acid residues have also been shown to be important in the interactions of CBMs with cellulose (*30*, *31*) and xylans (*32*). Typically, these residues are conserved among family members (*3*). It appears that tryptophan residues are also involved in cello-oligosaccharide binding by CBM17. This was indicated by the shielding of tryptophans from KI fluorescence quenching and NBS modification in the presence of ligand and by the tryptophan domination of the UV difference signal upon ligand binding. Of the five tryptophans in CBM17, four are conserved in family 17 (Figure 6) and are thus the likely candidates to participate in binding. Only two tyrosine residues are conserved, and there is no evidence for their involvement in substrate binding.

CBM17 and CBM4-1 have very similar binding specificities. Both bind cello-oligosaccharides, derivatized cellulose, mixed β1-3,4 glucans (barley β-glucan), and insoluble cellulose with amorphous content (Avicel and PASA). Like CBM4-1, which also does not bind BMCC (highly crystalline cellulose), it appears that CBM17 binds only the amorphous regions of Avicel, a heterogeneous allomorph of cellulose. The binding capacity of Avicel for CBM17, which is low in comparison to the binding capacity of family 2a CBMs that bind highly crystalline cellulose (*12*, *26*), also suggests that it binds the amorphous regions of Avicel. Despite the similarities in binding specificity, the association constants of CBM17 binding to cello-oligosaccharides are almost an order of magnitude greater than those for CBM4-1 (*14*, *33*). CBM4-1 binds substrate in a shallow cleft, which could accommodate a glucan chain five sugars in length (*9*). The full length of this cleft is lined with hydrophobic amino acid residues and flanked by polar residues. As predicted by the physical length of the cleft (~25 Å), cellopentaose is the optimal ligand for CBM4-1 (*33*). Cellohexaose was the optimal ligand for CBM17, suggesting a binding site on CBM17 that best accommodates a cellooligosaccharide of six glucose units length. Cellohexaose is approximately 31 Å long, implying that the binding site of CBM17 is approximately this size.

Cellotriose was the minimum ligand for CBM17. The decrease in ΔG with oligosaccharide length increasing from cellotriose to cellohexaose was linear, changing by −4.2 (±0.4) kJ mol^{-1} (glucose unit)$^{-1}$. The maximum change in the spectroscopic signal upon binding did not vary linearly with increasing length but was more profound with the longer ligands. Based on these results, a schematic model of the CBM17 binding site is shown in Figure 7. The presence of NaCl at 2 M improved the binding of CBM17 to cellopentaose and cellohexaose by approximately −2.0 kJ/mol. In contrast, binding to cellotetraose was improved by only −0.6 kJ/mol. The increased interaction of longer cello-oligosaccharides with aromatic amino acids (e.g., tryptophan) may lend a more hydrophobic character to the association of these ligands. Thus, dehydration effects are predicted to play a more important role in the association of cellopentaose and cellohexaose than they do for cellotriose and possibly

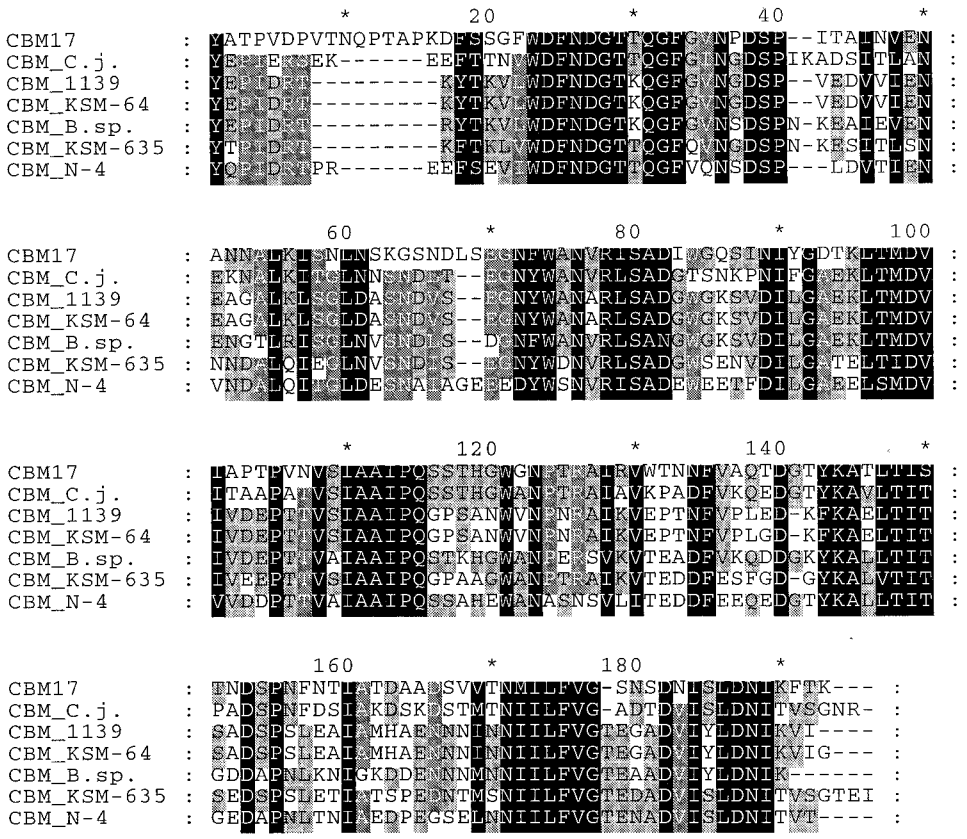

FIGURE 6: Amino acid alignment of the family 17 CBMs. Abbreviations are as follows: CBM17, from *C. cellulovorans* Cel5A (GenBank U37056); CBM_C.j., from *Clostridium josui* Cel5A (GenBank D85526); CBM_1139, from *Bacillus sp*., strain 1139, Cel5A (GenBank D00066); CBM_KSM-64, from *Bacillus sp*., strain KSM-64, Cel5A (GenBank M84963); CBM_B.sp., from *Bacillus sp*. Cel5A (GenBank D85236); CBM_KSM-635, from *Bacillus sp*., strain KSM-635, Cel5A (GenBank M27420): CBM_N-4, from *Bacillus sp*., strain N-4, Cel5A (GenBank M25500).

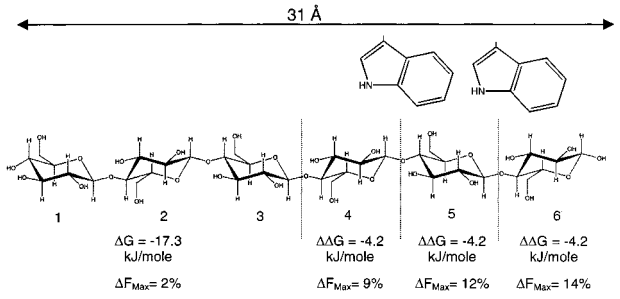

FIGURE 7: Schematic representation of the CBM17 substrate binding site. ΔG is for cellotriose. The $\Delta\Delta G$ values are relative to the next smallest cello-oligosaccharides (e.g., $\Delta G_{\text{cellotetraose}} - \Delta G_{\text{cellotriose}}$). ΔF_{Max} are the maximum relative changes in fluorescence at saturation. The depiction of two tryptophan residues is for the purposes of the schematic only. Up to five tryptophans may be present; however, it was estimated that two tryptophan residues would be sufficient to span the length of the last three glucose units of cellohexaose, which gives the largest change in the spectroscopic signal upon binding.

cellotetraose. Structural and mutagenic studies are currently underway to further investigate the architecture of the CBM17 binding site.

Carbohydrate-binding modules appear to serve the primary function of increasing the local concentration of enzyme on an insoluble substrate, thereby allowing more efficient substrate hydrolysis (*34*−*36*). Cellulose exists in several forms (*1*); however, they may be generalized to amorphous cellulose, cellulose with disordered regions, and crystalline cellulose, cellulose with highly organized chains. Many cellulases exhibit preferences for either amorphous or crystalline cellulose. Cel9A (formerly CenB) and Cel5A (formerly CenD) from *C. fimi* are efficient at hydrolyzing crystalline cellulose, and both have family 2a CBMs, which have high affinities for crystalline cellulose (*12*, *26*). Cel9B (formerly CenC) from *C. fimi* is most efficient at hydrolyzing amorphous or soluble cellulose and has tandem family 4 CBMs that bind strictly to amorphous or soluble cellulose (*14*). We have shown previously that by substituting the N-terminal family 4 CBM for the family 2a CBM of Cel6A (formerly CenA) from *C. fimi* its preference for amorphous cellulose is increased (*37*). It would appear that CBM17, which is shown here to bind only amorphous and soluble cellulose, serves the same purpose of targeting *C. cellulovorans* Cel5A to amorphous cellulose, the substrate on which it is most active. Thus, in some cases, the specificity of the CBM contained within an enzyme may aid in targeting the enzyme to regions of cellulose that are most efficiently hydrolyzed by the catalytic module.

ACKNOWLEDGMENT

We thank Dr. Roy H. Doi for providing the pEF 32 plasmid from which the CBM17 gene fragment was amplified. We also thank Angela Das and Jeanette Beatty for technical assistance.

REFERENCES

1. Tomme, P., Warren, R. A., and Gilkes, N. R. (1995) *Adv. Microb. Physiol. 37*, 1−81.
2. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. J., and Warren, R. A. (1991) *Microbiol. Rev. 55*, 303−315.

3. Tomme, P., Warren, R. A., Miller, R. C., Jr., Kilburn, D. G., and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Polysaccharides* (Saddler, J. N., and Penner, M., Eds.) pp 142−163, American Chemical Society, Washington, DC.
4. Coutinho, P. M., and Henrissat, B. (1999) in *Recent advances in carbohydrate bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B., and Svensson, B., Eds.) pp 3−12, Royal Society of Chemistry, Cambridge, U.K.
5. Boraston, A. B., McLean, B. W., Kormos, J. M., Alam, M., Gilkes, N. R., Haynes, C. A., Tomme, P., Kilburn, D. G., and Warren, R. A. J. (1999) in *Recent advances in carbohydrate bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B., and Svensson, B., Eds.) pp 202−211, Royal Society of Chemistry, Cambridge. U.K.
6. Mattinen, M. L., Kontteli, M., Kerovuo, J., Linder, M., Annila, A., Lindeberg, G., Reinikainen, T., and Drakenberg, T. (1997) *Protein Sci. 6*, 294−303.
7. Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.
8. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
9. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry 35*, 14381−14394.
10. Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) *Biochemistry 36*, 16074−16086.
11. Linder, M., Mattinen, M. L., Kontteli, M., Lindeberg, G., Stahlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G., and Annila, A. (1995) *Protein Sci. 4*, 1056−1064.
12. Gilkes, N. R., Jervis, E., Henrissat, B., Tekant, B., Miller, R. C. J., Warren, R. A., and Kilburn, D. G. (1992) *J. Biol. Chem. 267*, 6743−6749.
13. Goldstein, M. A., Takagi, M., Hashida, S., Shoseyov, O., Doi, R. H., and Segel, I. H. (1993) *J. Bacteriol. 175*, 5762−5768.
14. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
15. Brun, E., Gans, P., Marion, D., and Barras, F. (1995) *Eur. J. Biochem. 231*, 142−148.
16. Doi, R. H., Goldstein, M., Hashida, S., Park, J. S., and Takagi, M. (1994) *Crit. Rev. Microbiol. 20*, 87−93.
17. Doi, R. H., Park, J. S., Liu, C. C., Malburg, L. M., Tamaru, Y., Ichiishi, A., and Ibrahim, A. (1998) *Extremophiles 2*, 53−60.
18. Ishi, A., Sheweita, S., and Doi, R. H. (1998) *Appl. Environ. Microbiol. 64*, 1086−1090.
19. Coutinho, J. B., Gilkes, N. R., Warren, R. A., Kilburn, D. G., and Miller, R. C. J. (1992) *Mol. Microbiol. 6*, 1243−1252.
20. Blake, J. D.. and Richards, G. N. (1971) *Carbohydr. Res. 17*, 253−268.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY.
22. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) *Anal. Biochem. 200*, 74−80.
23. Eftink, M. R. (1997) *Methods Enzymol. 278*, 221−257.
24. Eftink, M. R., and Ghiron, C. A. (1981) *Anal. Biochem 114*, 199−227.
25. Takeo, K. (1985) *Electrophoresis 5*, 187−195.
26. Ong, E., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A., and Kilburn, D. G. (1993) *Biotechnol. Bioeng. 42*, 401−409.
27. 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.
28. Ragone, R., Colonna, G., Balestrieri, C., Servillo, L., and Irace, G. (1984) *Biochemistry 23*, 1871−1875.
29. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem. 65*, 441−473.
30. Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J., and Orosz, L. (1998) *FEBS Lett. 429*, 312−316.
31. Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C. J., Warren, R. A., and Kilburn, D. G. (1994) *Mol. Microbiol. 11*, 747−755.
32. Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J., and Williamson, M. P. (1999) *Struct. Fold. Des. 7*, 853−864.
33. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
34. Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988) *Eur. J. Biochem. 170*, 575−581.
35. Reinikainen, T., Teleman, O., and Teeri, T. T. (1995) *Proteins: Struct., Funct., Genet. 22*, 392−403.
36. Hall, J., Black, G. W., Ferreira, L. M., Millward-Sadler, S. J., Ali, B. R., Hazlewood, G. P., and Gilbert, H. J. (1995) *Biochem. J. 309*, 749−756.
37. Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A., and Miller, R. C., Jr. (1993) *FEMS Microbiol. Lett. 113*, 211−218.

BI0007728