Binding Specificity and Thermodynamics of a Family 9 Carbohydrate-Binding Module from *Thermotoga maritima* Xylanase 10A†

Alisdair B. Boraston,[‡,§,||] A. Louise Creagh,[‡,||,⊥] Md. Mustafa Alam,[‡,§,#] Jeffery M. Kormos,[||] Peter Tomme,[‡,§,||,○] Charles A. Haynes,[‡,||,⊥] R. Antony J. Warren,[‡,§] and Douglas G. Kilburn*[,‡,§,||]

Protein Engineering Network of Centres of Excellence, Department of Microbiology and Immunology, Department of Chemical and Biological Engineering, and The Biotechnology Laboratory, University of British Columbia, Vancouver, British Columbia, V6T 1Z3 Canada

Received January 26, 2001

ABSTRACT: The C-terminal family 9 carbohydrate-binding module of xylanase 10A from *Thermotoga maritima* (CBM9-2) binds to amorphous cellulose, crystalline cellulose, and the insoluble fraction of oat spelt xylan. The association constants (K_a) for adsorption to insoluble polysaccharides are 1×10^5 to 3×10^5 M^{-1}. Of the soluble polysaccharides tested, CBM9-2 binds to barley β-glucan, xyloglucan, and xylan. CBM9-2 binds specifically to the reducing ends of cellulose and soluble polysaccharides, a property that is currently unique to this CBM. CBM9-2 also binds glucose, xylose, galactose, arabinose, cellooligosaccharides, xylooligosaccharides, maltose, and lactose, with affinities ranging from 10^3 M^{-1} for monosaccharides to 10^6 M^{-1} for disaccharides and oligosaccharides. Cellooligosaccharides longer than two glucose units do not bind with improved affinity, indicating that cellobiose is sufficient to occupy the entire binding site. In general, the binding reaction is dominated by favorable changes in enthalpy, which are partially compensated by unfavorable entropy changes.

Many organisms have evolved diverse polysaccharolytic enzyme systems to tap the abundant carbohydrate energy source found in plant biomass. A large proportion of these enzymes have modular structures comprising catalytic modules, carbohydrate-binding modules (CBMs),[1] modules that mediate protein−protein interactions, modules of unknown function, and modules that appear to serve only as linkers between modules (*1*). CBMs, previously called cellulose-binding domains (CBDs), are grouped into families of related amino acid sequences, of which there are 13 at present (*2*, *3*).

CBMs within families and from different families can have different binding specificities. The N-terminal family 4 CBM, CBM4-1, of endoglucanase Cel9B from the mesophilic bacterium *Cellulomonas fimi*, binds to cellooligosaccharides and to amorphous cellulose but not to crystalline cellulose (*4*, *5*). The family 2a CBM, CBM2a, of xylanase 10A from *C. fimi*, binds irreversibly to crystalline cellulose and to crystalline regions of amorphous cellulose preparations but not to soluble cellulose or cellooligosaccharides (*6*−*9*). Thermodynamic studies by microcalorimetry show that the binding mechanisms of CBM4-1 and CBM2a are distinctly different. The binding of CBM4-1 is dominated by a favorable change in enthalpy, which is partly compensated by unfavorable entropy changes. Binding is also characterized by a small negative change in heat capacity (*4*). This indicates that the interaction between CBM4-1 and cellooligosaccharides involves both H-bonding and van der Waals interactions. The solution structure of CBM4-1 (*10*) shows that its binding site is a groove lined with hydrophobic amino acid residues and flanked with planar polar amino acid residues, suggesting an interaction with cellulose consistent with the calorimetry results. In contrast, the interaction of CBM2a with cellulose is dominated by a favorable change in entropy, a favorable, but very small, change in enthalpy, and a large negative change in heat capacity (*8*). The entropic driving force was interpreted to arise from ordered water molecules returned to the bulk solution by the dehydration of the polypeptide and the cellulose surface. The hydrophobic nature of the CBM2a binding site supports this (*9*, *11*), indicating that thermodynamic analysis of CBM−ligand interactions thus provides some understanding of the mechanisms as well as the energetics of binding. However, such a study has not yet been applied to CBMs from hyperthermophilic sources, which may have evolved novel characteristics to operate at elevated temperatures.

† 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: The Biotechnology Laboratory, University of British Columbia, 237-6174 University Blvd., Vancouver, BC, Canada V6T 1Z3. Tel: (604) 822-4182. 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.

⊥ Department of Chemical and Biological Engineering.

Current address: University of Alberta, 315 Heritage Medical Research Centre, Edmonton, Alberta, Canada.

○ Current address: Tibotec Generall de Wittelaan, L-11, 2800 Mechelen, Belgium.

[1] Abbreviations: BMCC, bacterial microcrystalline cellulose; CBM, carbohydrate-binding module; CBD, cellulose-binding domain; IPTG, isopropyl 1-thio-β-D-galactopyranoside; ITC, isothermal titration calorimetry: K_a, association constant; N_o, binding capacity; PASA, phosphoric acid swollen cellulose.

10.1021/bi0101695 CCC: $20.00
Published on Web 05/04/2001

Xylanase 10A (Xyn10A) from *Thermotoga maritima* is a highly thermostable 120 kDa modular xylanase. It comprises tandem N-terminal CBM22 modules, a family 10 catalytic module, and tandem ~170 amino acid C-terminal CBM9 modules, formerly referred to as C1 and C2 (*12*, *13*). The C-terminal-most C2 module, recently classified as a family 9 CBM (*3*, *14*), was initially shown to have cellulose-binding properties (*13*). In this work, we report the novel binding specificity and thermodynamics for this module, formally named Xyn10ACBM9-2 but referred to as CBM9-2 for brevity. The three-dimensional structure of this module determined by X-ray crystallography in ligand-bound forms is presented in the accompanying paper (*15*).

MATERIALS AND METHODS

Materials. Restriction endonucleases were from New England Biolabs (Mississauga, ON, Canada). T4 DNA ligase was from Gibco-BRL (Burlington, ON, Canada). PWO DNA polymerase and the Expand High Fidelity PCR system were from Boehringer Mannheim (Laval, PQ, Canada). Isopropyl β-D-thiogalactoside (IPTG), barley β-glucan, arabinogalactan, starch, locust bean gum, xylose, arabinose, *N*-acetylglucosamine, cellobiose, lactose, fructose, mannose, and shrimp chitin were from Sigma (St. Louis, MO). Glucose, maltose, galactose, and sucrose were from Fischer Scientific (Nepean, Ontario, Canada). Galactan, arabinan, xyloglucan (tamarind), arabinoxylan (wheat), and xylobiose were from Megazyme (Dublin, Ireland). Birch wood xylan was from Carl Roth AG (Karlsruhe, Germany). Avicel, PH101, was obtained from FMC International (Little Island, County Cork, Ireland). Bacterial microcrystalline cellulose (BMCC) and phosphoric acid swollen Avicel (PASA) were prepared as described previously (*6*, *16*). Regenerated Avicel (RA) was prepared by completely dissolving 1 g of Avicel in 200 mL of 85% phosphoric acid followed by precipitation of the cellulose by the addition of cold acetone. The precipitated cellulose was extensively washed with distilled water and finally resuspended in distilled water.

Bacterial Strains, Plasmids, Media, and Growth Conditions. T. maritima strain MSB8 was obtained from the ATCC. The *Escherichia coli* strains used were DH5α and JM101 (*17*). The plasmid used was pET28a (Novagen, Milwaukee, MI). *E. coli* cultures were grown routinely in TB or TYP medium (*17*) at 37 °C or at 30 °C for production of protein. The medium contained 50 μg of kanamycin mL^{-1} for strains carrying pET28 plasmids.

DNA Manipulations. Plasmid preparation, bacterial transformation, and agarose gel electrophoresis were done as described previously (*17*). Restriction endonucleases were used as recommended by the suppliers. DNA fragments were purified from the agarose gels with the Qiaquick Spin PCR purification kit or the Qiaex II kit (Qiagen, Valencia, CA). Oligodeoxynucleotide primers were synthesized by the Nucleic Acid and Protein Service (NAPS) Unit of the University of British Columbia, using an Applied Biosystems Model 380A DNA synthesizer, and purified by extraction with 1-butanol. DNA was sequenced by the NAPS Unit, using the AmpliTaq dye termination cycle sequencing protocol and an Applied Biosystems Model 377 sequencer.

Cloning of the DNA Fragment Encoding CBM9-2. The lyophilized cells of *T. maritima* received from the ATCC were suspended in 50 μL of sterile distilled water. The cells in 10 μL of the suspension were lysed as described previously (*18*). The lysate was used to amplify the DNA fragment encoding CBM9-2 by PCR. The primers used were 5′ ACGGTCCCATGGGCTAGCGGAATAATGGTAGCGAC-A 3′ (an *Nhe*I site is underlined) and 5′ CGACCGAAGC**T-TA**TCACTTGATGAGCCTGAGGT 3′ (a *Hin*dIII site is underlined; stop codon is in bold), allowing amplification of nucleotides 2849−3424 of *xyn10A*, encoding amino acids 869−1060 of Xyn10A (see Genbank Z46264). The PCR was run in a PTC-100 programmable thermal controller (MJ Research Inc., Watertown, MA). The 50 μL reaction mixtures contained 10 mM Tris-HCl, pH 8.8, 25 mM KCl, 5 mM $(NH_4)_2SO_4$, 2 mM $MgS0_4$, 5.0% DMSO, 400 μM dNTPs, 40 pmol of each primer, and 2 μL of 10 times diluted lysed *T. maritima.* The first five cycles were 98 °C for 1 min, 60 °C for 1 min, and 70 °C for 1 min. The last 25 cycles were 96 °C for 30 s, 45 °C for 1 min, and 70 °C for 1 min, followed by a final extension at 68 °C for 10 min. The purified fragment was digested with restriction endonucleases *Nhe*I and *Hin*dIII and ligated into pET28a digested with the same enzymes to give pETCBM9-2. The construct was verified by DNA sequencing after transformation into *E. coli* DH5α. It was transformed into *E. coli* BL21(DE3) for production of CBM9-2. The polypeptide encoded by pETCBM9-2 comprises a hexahistidine tag fused to the N-terminus of CBM9-2 by a thrombin cleavage site.

Production and Purification of CBM9-2. TYP medium (60 L), inoculated with *E. coli* BL21(DE3)/pETCBM9-2, was grown at 30 °C to an OD_{600nm} of ~0.6. IPTG was added to 0.3 mM, and incubation was continued overnight. The cells were harvested by centrifugation, resuspended in 1 L of 50 mM potassium phosphate buffer, pH 7.0 (KP), and ruptured by two passages through a French pressure cell (21000 lb/in^2). A 300 mL sample of the clarified cell extract was loaded at 1 mL min^{-1} onto a 200 mL column (packed bed volume) of beaded cellulose (95% porosity, 100−250 μM diameter; Sigma, St. Louis, MO). The column was washed with 5 column volumes of 50 mM potassium phosphate, pH 7.0, containing 500 mM NaCl, followed by a five column wash with 50 mM potassium phosphate, pH 7.0. Bound CBM9-2 was desorbed with 1 M glucose (in 50 mM potassium phosphate, pH 7.0). Fractions of the glucose eluate were screened by SDS−PAGE, and those containing pure polypeptide of the expected size (21 kDa) were pooled. Glucose was removed by ultrafiltration using a 1K Ultrafiltron membrane (Filtron Technology Corp., Northborough, MA) and the polypeptide concentrated into 50 mM potassium phosphate, pH 7.0. Complete removal of the glucose was checked by the phenol sulfuric acid assay (*19*). The concentrated solution was stored at 4 °C.

Preparation of Reduced Polysaccharides. One gram of Avicel (PH 101) or 1 g of oat spelt xylan (Sigma) was suspended in 15 mL of 100 mM NH_4OH (pH 10.8). Then 150 mg of $NaBH_4$ was added to this, and the mixture was covered and incubated for 6 h or overnight with stirring at room temperature. The polysaccharide was pelleted by centrifugation at 27000*g* and washed with 40 mL of distilled water. This process was repeated four times. The final polysaccharide pellet was resuspended in 10 mL of distilled water. A control sample to which $NaBH_4$ was not added was prepared in parallel to assess the effect of high-pH treatment

on the polysaccharide. The final concentration of these samples was determined by the dry weights of triplicate 1 mL samples. One hundred milligrams of barley β-glucan (Sigma, St. Louis, MO) was dissolved in 5 mL of 100 mM NH_4OH (pH 10.8). Then 50 mg of $NaBH_4$ was added to this, and the mixture was covered and incubated for 6 h with stirring at room temperature. The sample was dialyzed against distilled water for 4 days using a 1000 Da cutoff membrane (Spectra-Por 6; Spectrum Laboratories Inc., Rancho Dominiguez, CA). A control sample, to which $NaBH_4$ was not added, was prepared in parallel. The final concentration of these samples was determined by the phenol−sulfuric acid assay for total sugar (*19*) using glucose standards. The concentration of free reducing ends in the Avicel, xylan, and barley β-glucan preparations was determined by the hydroxybenzoic acid hydrazide (HBAH) assay (*20*) using glucose standards.

Protein Concentration Determination. The concentration of purified protein was determined by UV absorbance (280 nm) using a calculated molar extinction coefficient (*21*) of 43 430 M^{-1} cm^{-1}.

Polysaccharide Binding Assays. For insoluble polysaccharides binding constants at 22 °C were determined from depletion isotherms as described previously (*6*, *7*, *9*, *22*). The equilibrium association constants (K_a) were determined by nonlinear regression of bound versus free protein concentrations to a one binding site langmuir model as described previously (*8*, *22*). The reversibility of binding was determined as described previously using absorbance at 280 nm to quantify protein (*23*).

Binding to soluble polysaccharides was determined by affinity electrophoresis as described previously (*4*, *24*, *25*). The separating gels (7.5% polyacrylamide) contained 0.1% polysaccharide. Five micrograms of native gel electrophoresis standards (Sigma) was used as references. Electrophoresis was at 4 °C and 150 V for 2.5 h. Proteins were visualized by staining with Coomassie brilliant blue.

Competition Binding Assay. CBM9-2 (40 μM) was equilibrated with regenerated cellulose, 4 mg mL^{-1}, in 50 mM potassium phosphate, pH 7.0. Aliquots of this solution were diluted 4-fold by the addition of sugar solutions (5 mM in 50 mM potassium phosphate, pH 7.0) to give final concentrations of 10 μM, 1 mg mL^{-1}, and 3.75 mM for CBM9-2, regenerated cellulose, and sugar, respectively. Buffer, rather than sugar, was added to control tubes. All samples were prepared in triplicate. After overnight equilibration at 22 °C samples were centrifuged at 13 000 rpm in a benchtop microfuge to pellet the cellulose. Protein concentrations in the supernatant were determined from A_{280nm} as above. Control samples (no competitor added) represented the free protein concentration at 0% competition; the total concentration of added CBM9-2 (10 μM) represented 100% competition.

Fluorescence Spectroscopy. All fluorescence binding experiments were performed as described previously (*26*) except that 50 mM potassium phosphate buffer, pH 7.0, was used to buffer the system and 0.5 μM CBM9-2 was used. The fluorescence intensity was measured using an excitation wavelength of 280 nm and emission wavelengths of 330 and 350 nm. Data were analyzed as described previuosly (*26*).

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) was performed using a MCS ITC (MicroCal, Inc., Northampton, MA). All samples were pH 7.0 in 50 mM phosphate buffer. Titrations were performed by injecting 5 or 10 μL aliquots of sugar solution into the ITC sample cell (volume = 1.3528 mL) containing CBM9-2. The concentrations of CBM9-2 in the sample cell were determined spectrophotometrically at 280 nm. C values [$C = K_a \times M \times n$, where K_a is the association constant, M is the macromolecule concentration (i.e., CBM9-2), and n is the number of binding sites on the macromolecule] were always greater than 4 and, in most cases, greater than 10. All ITC data were corrected for the heat of dilution of the titrant by subtracting mixing enthalpies for injections of sugar solution into protein-free buffer. Two or three independent titration experiments were performed at 25 °C to determine the binding constant of CBM9-2 to sugars. Binding stoichiometry, enthalpy, and the equilibrium association constant were determined by fitting the corrected data to a bimolecular interaction model. Additional sets of titrations were performed at 30, 35, and 40 °C to measure the heat capacity of binding for all sugars except cellobiose, for which binding to CBM9-2 was measured up to 70 °C in increments of 5 °C. Due to the low binding constant of glucose to CBM9-2 and therefore the high protein concentrations required, full titrations were only performed at 25 and 35 °C. To determine ΔC_p for this system, additional enthalpies of binding were measured between 25 and 40 °C by single 10 μL injection of CBM9-2 (107 μM) into a 35 mM glucose solution. ΔG and K_a values at 90 °C were calculated using the equation:

$$\Delta G_T = \Delta H_{T_0} - T\Delta S_{T_0} + \Delta C_p[T - T_0 - T\ln(T/T_0)] \quad (1)$$

where T_0 is the reference temperature, T is the temperature, and ΔG is the calculated free energy at temperature T. The predicted association constant is then calculated from the ΔG value using the fundamental relation $\Delta G = -RT\ln(K_a)$.

Table 1: Affinity Constants of CBM9-2 Binding to Insoluble Polysaccharides Determined by Depletion Isotherms

substrate[a]	K_a ($\times 10^{-5}$ M^{-1})	N_o (μmol/g)
xylan	1.5 ± 0.4	1.2 ± 0.1
PASA	1.8 ± 0.3	7.7 ± 0.4
RA	3.0 ± 0.2	12.3 ± 0.5
Avicel	1.5 ± 0.1	1.3 ± 0.3
BMCC	1.3 ± 0.1	1.8 ± 0.1

[a] Xylan from oat spelt, insoluble fraction; PASA, phosphoric acid swollen Avicel; RA, regenerated Avicel; BMCC, bacterial microcrystalline cellulose.

RESULTS

Binding to Insoluble Substrates. CBM9-2 binds to several cellulose allomorphs and the insoluble fraction of oat spelt xylan with affinities similar to those of family 1 and family 4 CBMs for their respective ligands (Table 1) (*4*, *27*, *28*). The affinites are slightly greater for amorphous celluloses (PASA and RA) than for crystalline celluloses (BMCC and Avicel); amorphous celluloses had significantly greater binding capacity than crystalline celluloses (K_a and N_o values, respectively, Table 1). Irreversible binding of some CBMs to cellulose has been documented (*29*). However, the binding of CBM9-2 to PASA was tested by performing a dilution

Table 2: Competition of Various Saccharides with Regenerated Cellulose for Binding to CBM9-2[a]

saccharide	% competition
1-*O*-Me-β-D-glucopyranoside, 1-*O*-Me-α-D-glucopyranoside, 6-deoxy-L-mannose, sucrose	0
D-mannose, D−*N*-acetylglucosamine	<5
D-ribose, 6-deoxy-L-galactose, raffinose, D-glucuronic acid	5−10
L-arabinose	31
D-xylose	37
gentibiose	42
D-glucose	57
D-galactose	60
cellobiose, lactose, maltose	>95

[a] Percent competition was determined as described in Materials and Methods.

Table 3: Affinity Constants of CBM9-2 Binding to Various Saccharides Quantified by Fluorescence Titrations[a]

saccharide	composition	K_a ($\times 10^{-5}$ M^{-1})	ΔG (kJ/mol)
D-glucose		0.134 ± 0.018	−23.1
D-galactose		0.038 ± 0.001	−19.7
D-xylose		0.051 ± 0.003	−21.0
xylobiose	XylDβ-1,4-XylD	1.90 ± 0.03	−30.1
xylopentaose	(XylDβ-1,4-XylD)$_5$	1.71 ± 0.03	−29.8
gentibiose	GluDβ-1,6-GluD	0.036 ± 0.004	−20.1
maltose	GluDα-1,4-GluD	3.70 ± 0.10	−31.6
lactose	GalDβ-1,4-GluD	13.0 ± 0.80	−34.6
cellobiose	GluDβ-1,4-GluD	22.6 ± 2.60	−35.6
cellotriose	(β-1,4-GluD)$_3$	16.0 ± 3.10	−34.8
cellotetraose	(β-1,4-GluD)$_4$	19.9 ± 1.30	−35.3
cellopentaose	(β-1,4-GluD)$_5$	27.9 ± 3.40	−36.2
cellohexaose	(β-1,4-GluD)$_6$	19.1 ± 1.50	−35.2

[a] Errors represent the standard deviation of triplicate or quadruplicate experiments.

isotherm, which showed that the binding is reversible (results not shown).

Binding to Soluble Polysaccharides. Affinity electrophoresis at 4 °C (*25*) shows that CBM9-2 binds to barley β-glucan and xyloglucan and weakly to soluble birchwood xylan and wheat arabinoxylan. There is no detectable binding to arabinan, arabinogalactan, CM-cellulose, hydroxyethylcellulose, dextran, starch, pectic galactan, laminarin, or locust bean gum (results not shown).

Binding to Small Soluble Saccharides. CBM9-2 is desorbed from cellulose by 0.1 M cellobiose (*13*) or 0.1 M glucose (data not shown). Several other mono- and disaccharides are able to compete effectively with the binding of regenerated cellulose to CBM9-2 (Table 2).

The binding of CBM9-2 to the ligands that were positive in the competition assay, as well as to some related ligands, was quantified by fluorescence titrations (Table 3). CBM9-2 appears to prefer the β-1,4-linked sugars, though clearly other linkages are also appropriate. No increase in affinity is seen for cellooligosaccharides or xylooligosaccharides longer than two sugars, suggesting that a disaccharide completely occupies the binding site.

Binding to Reduced Polysaccharides. The competition studies show that C1-O-methylated sugars do not bind to CBM9-2. Thus, the reducing ends of small sugars appear to contain the recognition determinants required for their binding to CBM9-2. This hypothesis was confirmed by altering the reducing ends of microcrystalline cellulose (Avicel), insoluble oat spelt xylan, and barley β-glucan through reduction of the hemiacetal linkage at C-1 with sodium borohydride, thus opening the sugar ring.

CBM9-2 binds unmodified Avicel with an association constant and capacity of 2.6 (±0.1) × 10^5 M^{-1} and 1.6 (±0.0) μmol/g of cellulose, respectively. The concentration of reducing ends is 3.7 μmol/g of cellulose. Binding to the $NaBH_4$-treated Avicel, which had a reducing end concentration of 0.3 μmol/g of cellulose, could not be detected. $NaBH_4$ treatment of Avicel does not, however, affect the binding capacity or affinity of CBM2a, a CBM that binds crystalline cellulose and crystalline regions of amorphous cellulose (*9*). This shows that the gross structure of the cellulose is not significantly perturbed by the reducing agent. It is of interest that the maximum binding capacity of Avicel is approximately 50% of the measured reducing end concentration. The accessibility of CBM9-2 to only a proportion of the reducing ends presumably reflects steric problems, with some of the ends being closely associated with the crystalline cellulose surface or residing in pores too small to allow the entry of CBM9-2.

CBM9-2 binds unmodified insoluble oat spelt xylan with an association constant and capacity of 1.1 (±0.3) × 10^5 M^{-1} and 1.1 (±0.1) μmol/g of xylan, respectively. The concentration of reducing ends is 25.1 μmol/g of xylan. Extended treatment of oat spelt xylan with $NaBH_4$ decreased the reducing end concentration to 16.4 μmol/g of xylan. The binding affinity of CBM9-2 to the partially reduced xylan was unchanged, but the capacity was decreased by 33% to 0.7 (±0.1) μmol/g of xylan. This decrease was consistent with the 34% drop in the reducing end concentration. These results show that CBM9-2 also recognizes the reducing ends of insoluble xylan. It appears that very few of the reducing ends of oat spelt xylan act as binding sites (~4%). The reason for this is unclear, but because insoluble xylan does not have the ordered structure of crystalline cellulose, it would be expected to have a large proportion of its reducing ends accessible. It may indicate that many of the terminal xylose residues in oat spelt xylan are blocked by arabinosyl side chains or simply inaccessible to CBM9-2.

A similar specificity for the reducing ends of barley β-glucan was found (Figure 1). The reducing end concentrations of the control barley β-glucan and the $NaBH_4$-treated barley β-glucan were 6.41 μmol/g and 0.76 μmol/g, respectively. The small amount of intact reducing ends in the $NaBH_4$-treated sample is likely responsible for the weak binding to the reduced sample. CBM17, which can bind internally on the glucan polymer (*26*), binds as well to $NaBH_4$-treated barley β-glucan as it does to the control barley β-glucan (Figure 1).

Thermodynamics of Substrate Binding. The thermodynamics of CBM9-2 binding was investigated by ITC. The binding stoichiometry to mono-, di-, and oligosaccharides was 1:1, consistent with the single binding site found by X-ray crystallography (Table 4, Figure 2) (*15*). The binding of CBM9-2 to small soluble ligands and cellulose is highly exothermic, with the favorable changes in enthalpy partly compensated by unfavorable changes in entropy (Table 4). This relationship and the general magnitudes of ΔH and $T\Delta S$ are similar to those found in some lectin−carbohydrate systems (*30*). The association constants determined by ITC

Table 4: Thermodynamics of CBM9-2 Binding to Various Saccharides Quantified by Isothermal Titration Calorimetry at 25 °C[a]

saccharide	K_a ($\times 10^{-5}$ M^{-1})	ΔG (kJ mol^{-1})	ΔH (kJ mol^{-1})	$-T\Delta S$ (kJ mol^{-1})	ΔC_p (J mol^{-1} K^{-1})	n[b]
glucose	0.03	−20.19	−31.18 ± 0.53	11.37	−143 ± 65	0.75
cellobiose	7.45 ± 0.06	−33.52 ± 0.02	−46.32 ± 0.11	12.80 ± 0.11	−280 ± 9	0.90 ± 0.01
cellotetraose	5.70 ± 0.20	−32.85 ± 0.09	−45.26 ± 0.33	12.41 ± 0.34	−168 ± 14	0.88 ± 0.02
RA	2.43 ± 0.45	−30.71 ± 0.44	−48.24 ± 3.63	17.53 ± 3.66	−929 ± 170	ND
xylobiose	1.22 ± 0.13	−29.02 ± 0.28	−37.51 ± 0.91	8.49 ± 0.96	ND	0.79 ± 0.01
lactose	5.39 ± 0.19	−32.72 ± 0.09	−43.75 ± 0.12	11.03 ± 0.15	−285 ± 31	0.89 ± 0.01
maltose	2.10 ± 0.02	−30.39 ± 0.03	−39.75 ± 0.06	9.37 ± 0.07	−297 ± 14	0.91 ± 0.01

[a] Errors represent the standard deviations of triplicate or quadruplicate experiments or propagation of error. [b] n refers to the stoichiometry of the interaction.

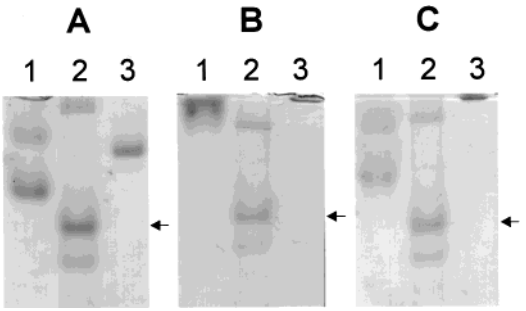

FIGURE 1: Affinity electrophoresis of CBM9-2. Panels: A, 10% native polyacrylamide gel; B, 10% native polyacrylamide gel containing 0.13% barley β-glucan; C, 10% native polyacrylamide gel containing 0.13% $NaBH_4$ treated barley β-glucan. Lanes: 1, 10 μg of CBM9-2; 2, 10 μg of native gel standards; 3, 10 μg of CBM17. The relative mobilities of CBM9-2, calculated as the mobility of CBM9-2 divided by the mobility of a reference (indicated by arrows), were as follows: native gel, 0.72; barley β-glucan, 0.17; reduced barley β-glucan, 0.69. CBM17 appears at the top of the gels in panels B and C.

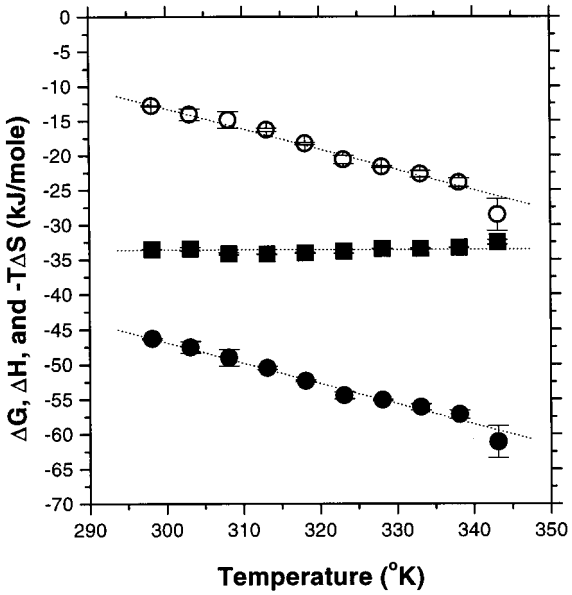

FIGURE 2: Temperature dependence of ΔG (■), ΔH (●), and $T\Delta S$ (○) for binding to cellobiose. Error bars represent the standard deviation of triplicate experiments.

are 2−3-fold lower than those determined by fluorescence titrations. The fitting of fluorescence data assumes a perfect 1:1 stoichiometry whereas the stoichiometry is one of the parameters regressed in the fitting of ITC data. This difference may contribute to the discrepancy in the association constants. Moreover, binding constants determined by different techniques often differ by a factor of about 2−3 (*4*, *31*).

CBM9-2 is remarkably thermostable, resisting denaturation at temperatures above 100 °C (*32*). Using ITC, we investigated the thermodynamics of binding to cellulose between 25 and 70 °C (Figure 2). With increasing temperature, the binding enthalpy becomes increasingly favorable in a linear manner. The changes in ΔH are compensated by equal and unfavorable changes in $T\Delta S$ such that ΔG does not vary significantly. The ΔC_p values of −0.292 kJ mol^{-1} K^{-1} (correlation coefficient 0.99) and −0.293 kJ mol^{-1} K^{-1} (correlation coefficient 0.99), determined by linear regression of the ΔH and $T\Delta S$ values, respectively, were in excellent agreement, and ΔC_p can therefore be considered a constant up to 70 °C. The temperature dependence of ΔH on binding is often used to calculate the ΔC_p of sugar-binding systems; however, this is usually done over a relatively narrow temperature range, and the assumption is usually made that ΔC_p is temperature independent. Until now, 37 °C was the highest temperature for which ITC using a carbohydrate-binding protein was reported (*33*). Seventy degrees Celsius is the highest temperature for which titration calorimetry has been reported for a carbohydrate-binding protein. Furthermore, this is the first time that the dependence of ΔH on temperature has been determined up to such high temperatures and that ΔC_p, in this case, is temperature independent.

DISCUSSION

CBM9-2 Binding Specificity. CBM9-2 binds both insoluble cellulose and insoluble xylan with high affinity. Considering that this module comes from a xylanase, it would seem most likely that the biological ligand of CBM9-2 is xylan rather than cellulose. However, CBMs with cellulose-binding activity are often found on xylanases (e.g., Xyn10A from *C. fimi* and Xyn10A from *Pseudomonas fluorescens*). It is thought in these cases that because xylan is associated with cellulose in plant cell walls that binding to cellulose will still target the enzyme to its substrate. A similar scenario may be envisioned in *T. maritima* Xyn10A, though a source of xylan that is associated with cellulose in a hyperthermophilic marine environment is not obvious.

Perhaps the most novel property of CBM9-2 is that it interacts directly with reducing end of polysaccharides. More specifically, on the basis of the results of the small sugar-binding assays, CBM9-2 requires access to the C1-hydroxyl and an equatorial C2-hydroxyl. The C6-hydroxymethyl group and the C4-hydroxyl appear to be unimportant in binding. These results are consistent with the H-bonding pattern observed in the crystal structure of CBM9-2 in complex with several ligands (*15*), where the C1-hydroxyl, C2-hydroxyl, and pyranose oxygen of several sugars were involved in H-bonds with CBM9-2.

CBM9-2 binds relatively weakly to the soluble glycans xylan, xyloglucan, and barley β-glucan in affinity electrophoresis assays. This is in apparent variance with its high affinity for insoluble cellulose, insoluble xylan, and oligosaccharides. The reducing end concentrations in these soluble glycan preparations are relatively low because of their substantial degree of polymerization, thus, giving the impression of low-affinity binding in assays, such as affinity electrophoresis, where the stoichiometry of the interaction is not considered. Furthermore, quantitative binding studies using these polysaccharides is limited by their solubility; they are not soluble enough to prepare samples with sufficiently high reducing end concentrations to perform detailed binding studies. Aside from the reducing end of the backbone polymer, the arabinose and xylose branches on xylan and xyloglucan, respectively, are obvious potential targets for binding. However, our results indicate a strict requirement for reducing sugars; the branches of xylans and xyloglucans are nonreducing as the C1 moieties are involved in the glycosidic bonds.

CBM9-2's specificity for the reducing ends of polysaccharides is unique among currently characterized CBMs. The N-terminal family 4 CBM, CBM4-1 (formerly called CBDN1), from *C. fimi* Cel9B and CBM17 from *Cellulomonas cellulovorans* Cel5A (*26*) can bind internally on soluble glucan chains (*4*). The proposed model of xylan binding by the internal family 2b CBM, CBM2b-1 (or XBD1), from *C. fimi* Xyn10A also suggests that it binds internally on the xylan chain (*34*). Moreover, the architecture of the binding faces of family 1, 2a, 3, and 5 CBMs does not indicate that they can discriminate between the ends of cellulose chains (*11*, *35*−*37*). Indeed, we have shown here that the family 2a CBM, CBM2a, does not require the reducing ends of cellulose for binding. The reducing end specificity of CBM9-2 therefore suggests that it has a role in targeting the enzyme, Xyn10A. This in turn may impact the mode of attack of the enzyme, such as introducing processivity from the reducing end of the polysaccharide or targeting to "broken" portions of the polysaccharide chains.

The affinity of CBM9-2 for mono- and disaccharides has important consequences for its practical application as an affinity tag. Glucose and cellobiose can be used to desorb CBM9-2 from cellulose. Presumably the other sugars that were found to compete with cellulose for binding to CBM9-2 could be used as well. The low cost and availability of cellulose and these competitive eluents make CBM9-2 an attractive candidate for use as a purification tag.

Thermodynamic Mechanism of Binding. The binding of small soluble sugars to CBM9-2 is enthalpically driven, with ΔH values ranging from −31.56 kJ/mol for glucose to −46.32 kJ/mol for cellobiose, the strongest binding sugar at 25 °C. Exothermic binding can result from net favorable intermolecular hydrogen bond formation between the sugar and the protein, formation of new water−water hydrogen bonds in the bulk solution, and/or formation of close van der Waals contacts between the sugar and the protein. The strongly exothermic binding of soluble sugars to CBM9-2 is consistent with the X-ray crystallographic structures for the protein−carbohydrate complexes. The binding site of CBM9-2 is a shallow cleft containing a number of basic and acidic side-chain residues (*15*). Five to eight intermolecular hydrogen bonds are formed between the reducing end glycopyranoside unit and five to six neighboring residues in the binding cleft. Two tryptophans, the faces of which are approximately parallel, form the entrance to the cleft and appear to sandwich portions of the first and second units of the sugar chain by being centered over the glycosidic bond. Except in the unique case of cellobiose, where the C6-hydroxyl of the second sugar unit makes a hydrogen bond with CBM9-2, the interactions of the second glycopyranoside unit are limited to stacking interactions with the tryptophans. The structural elements within the binding site therefore provide a number of hydrogen bond donors and acceptors, as well as strong van der Waals contacts through ring stacking of the sugar and the two tryptophans.

Binding of sugars to CBM9-2 is characterized by a relatively small negative change in heat capacity, as observed for many sugar-binding reactions (*4*, *38*). Among the directly measurable thermodynamic parameters, ΔC_p provides the most sensitive gauge of changes in solvent environment (*39*, *40*). The observed small change in ΔC_p reflects the amphipolar nature of sugars, the CBM9-2 binding site, and the protein−sugar complex.

Binding of sugars to CBM9-2 is opposed by a net decrease in entropy. The similar binding entropies for cellobiose and cellotetraose suggest that changes in the conformational entropy of the sugar chain beyond the second glucose unit do not significantly contribute to the observed ΔS. This leads to a binding model in which the sugar chain is anchored at its reducing end to CBM9-2, with the second glucose unit interacting specifically with W71 and W175, and any additional glucopyranoside units extending away from the binding site and free to adopt their full range of chain conformations. Although the second glycopyranoside unit in a disaccharide significantly contributes to the binding affinity, its net contribution to the binding entropy is small. The $T\Delta\Delta S$ values for binding of disaccharides relative to glucose at 25 °C range from 2.0 to −1.43 kJ/mol. In comparison, the net changes in enthalpy ($\Delta\Delta H$) range from −8.19 to −14.76 kJ/mol. This indicates a highly favorable enthalpic interaction between the protein and the second sugar unit which, unlike the first glucopyranoside unit, is not strongly opposed by a concomitant decrease in entropy. This suggests that CBM9-2 specifically binds the individual glycopyranoside units of a disaccharide in distinctly different ways.

Figure 3 shows a linear relationship between ΔH and $T\Delta S$ for CBM9-2 binding to cellobiose, cellotetraose, lactose, maltose, and xylobiose, indicating a common binding mechanism. Perfect enthalpy−entropy compensation (a slope of ΔH vs $T\Delta S$ equal to 1) results when, on average, each unique contact formed results in an equivalent change in enthalpy and entropy. Sugar binding to CBM9-2 is characterized by partial enthalpy−entropy compensation, with a slope of 1.8, indicating that the unique contacts formed between the protein and each sugar influence ΔH more than ΔS. On the basis of the structural information reported by Notenboom et al. (*15*) for CBM9-2 complexed with glucose, cellobiose, cellotetraose, and xylopentaose, binding always involves formation of a hydrogen bond between N172 and the C1-OH proton, as well as a hydrogen bond between R161 and the C1-OH oxygen. These two specific interactions, which

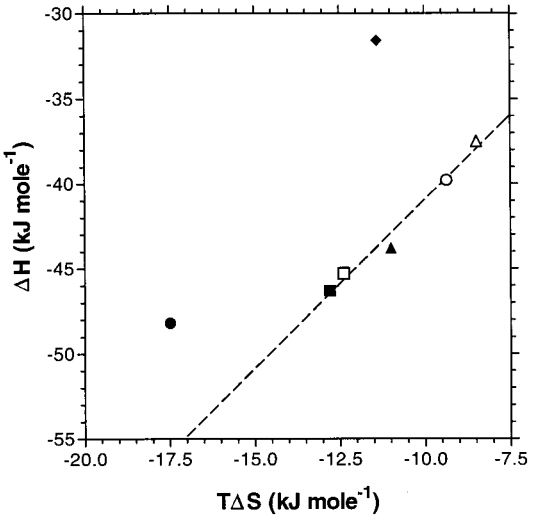

FIGURE 3: Enthalpy−entropy compensation plot of CBM9-2−sugar interactions. The symbols are as follows: ●, regenerated Avicel; ■, cellobiose; □, cellotetraose; ◆, glucose; ▲, lactose; ○, maltose; △, xylobiose. The dashed line is the best-fit line determined by linear regression. The slope is 1.8 with a correlation coefficient of 0.97.

are common to all sugar complexes studied and therefore do not contribute to the slope of the enthalpy−entropy compensation plot, serve to fix each sugar within the binding cleft. Formation of these two common contacts drastically reduces the entropy of the sugar (as well as the conformational entropy of N172 and R161). Unique interactions formed between the sugar and the protein therefore have a relatively small influence on the entropy of the sugar. They can, however, strongly influence ΔH, particularly if the interaction involves formation of a hydrogen bond or a large and close van der Waals contact area. The regressed partial compensation curve in Figure 3 does not include binding−thermodynamics data for glucose or regenerated Avicel. The departure of glucose from the regressed compensation line reflects the importance of the second sugar unit, including its interactions with W71 and W175, to the overall binding mechanism and energetics. Regenerated Avicel has a complex tertiary structure. The different mechanism of association with this polysaccharide may relate to the environment and organization of the sugar chains such as differences in the solvation state of free ends due to the proximity of cellulose in the solid phase. This is supported by the large negative ΔC_p for binding to regenerated Avicel, which also suggests considerable rearrangement of water. Alternatively, entropy effects resulting from the immobilization of portions of the cellulose chain to which the CBM binds may decrease enthalpy−entropy compensation.

Implications for Binding at Elevated Temperatures. Because Xyn10A functions at extremely high temperatures, it is important to determine how CBM9-2 performs at these temperatures. Although it has been established that CBM9-2 remains folded at the growth temperature of *T. maritima* (80−90 °C) (*32*), for practical reasons it is difficult to do standard binding assays at temperatures this high. However, using ITC we obtained a complete thermodynamic description of ligand binding over a temperature range extending to 70 °C, allowing the calculation of ΔH, ΔS, ΔC_p, ΔG, and, thus, the equilibrium constant at any temperature. Assuming that the ΔC_p for cellobiose remains constant above 70 °C and that this relationship is true for other ligands, the binding constants at 90 °C were calculated.

The association constants decrease quite profoundly when mesophilic and hyperthermophilic temperatures are compared (e.g., 70-fold and 250-fold decreases for cellobiose and RA, respectively). This translates into a shift toward higher concentrations of free CBM at high temperatures. For example, at 90 °C a concentration of CBM9-2 250 times that at 25 °C is required to reach 50% saturation of a sample of regenerated Avicel. The biological function of CBMs, at least in part, is postulated to be in the concentration of the enzyme onto the substrate. It would appear that at extreme temperatures much of this advantage is lost. However, the effectiveness of a CBM depends on the relative concentrations of substrate and enzyme. If CBM9-2 is indeed effective in this targeting role, this implies that the biological substrate of xylanase 10A must be in great excess such that the bound species is favored and a high-affinity interaction is not required. Alternatively, the reduction in affinity of this module binding at high temperatures is compensated by the affinity of other ancillary modules present in xylanase 10A, such as the tandem N-terminal "A" modules (*13*). These A modules have been shown to be xylan-binding CBMs (*41*) belonging to the recently classified CBM family 22 (*42*). Such a multipoint attachment, as seen in lectin systems, can result in large increases in affinity.

ACKNOWLEDGMENT

We thank Emily Amandoron-Akow, Emily Kwan, and Vinit Srivastava for technical assistance.

REFERENCES

1. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. J., and Warren, R. A. (1991) *Microbiol. Rev. 55*, 303−315.
2. Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A. L., Sturch, K., Gilkes, N. R., Haynes, C. A., Warren, R. A. J., and Kilburn, D. G. (1998) *J. Chromatogr. B 715*, 283−296.
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. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
5. Coutinho, J. B., Gilkes, N. R., Warren, R. A., Kilburn, D. G., and Miller, R. C. J. (1992) *Mol. Microbiol. 6*, 1243−1252.
6. 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.
7. Ong, E., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A., and Kilburn, D. G. (1993) *Biotechnol. Bioeng. 42*, 401−409.
8. Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G., and Haynes, C. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12229−12234.
9. McLean, B. W., Bray, M. R., Boraston, A. B., Gilkes, N. R., Haynes, C. A., and Kilburn, D. G. (2000) *Protein Eng. 13*, 801−809.
10. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry 35*, 14381−14394.
11. 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.
12. Goldstein, M. A., Takagi, M., Hashida, S., Shoseyov, O., Doi, R. H., and Segel, I. H. (1993) *J. Bacteriol. 175*, 5762−5768.
13. Winterhalter, C., Heinrich, P., Candussio, A., Wich, G., and Liebl, W. (1995) *Mol. Microbiol. 15*, 431−444.

14. 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.
15. Notenboom, V., Boraston, A. B., Kilburn, D. G., and Rose, D. R. (2001) *Biochemistry 40*, 6248−6256.
16. Wood, T. M., and Garcia-Campayo, V. (1990) *Biodegradation 1*, 147−161.
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Liebl, W., Gabelsberger, J., and Schleifer, K. H. (1994) *Mol. Gen. Genet. 242*, 111−115.
19. Chaplin, M. F. (1986) in *Carbohydrate Analysis: a Practical Approach* (Chaplin, M. F., and Kennedy, J. F., Eds.) pp 1−36, IRL Press, Oxford.
20. Lever, M. (1973) *Biochem. Med. 7*, 274−281.
21. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) *Anal. Biochem. 200*, 74−80.
22. 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.
23. Linder, M., and Teeri, T. T. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12251−12255.
24. Takeo, K. (1985) *Electrophoresis 5*, 187−195.
25. Tomme, P., Boraston, A., Kormos, J. M., Warren, R. A., and Kilburn, D. G. (2000) *Enzyme Microb. Technol. 27*, 453−458.
26. Boraston, A. B., Chiu, P., Warren, R. A. J., and Kilburn, D. G. (2000) *Biochemistry 39*, 11129−11136.
27. Linder, M., Mattinen, M. L., Kontteli, M., Lindberg, G., St*o*hlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G., and Annila, A. (1995) *Protein Sci. 4*, 1056−1064.
28. Brun, E., Johnson, P. E., Creagh, A. L., Tomme, P., Webster, P., Haynes, C. A., and McIntosh, L. P. (2000) *Biochemistry 39*, 2445−2458.
29. Jervis, E. J., Haynes, C. A., and Kilburn, D. G. (1997) *J. Biol. Chem. 272*, 24016−24023.
30. Lis, H., and Sharon, N. (1998) *Chem. Rev. 98*, 637−674.
31. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
32. Wassenberg, D., Schurig, H., Liebl, W., and Jaenicke, R. (1997) *Protein Sci. 6*, 1718−1726.
33. Sigurskjold, B. W., and Bundle, D. R. (1992) *J. Biol. Chem. 267*, 8371−8376.
34. Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J., and Williamson, M. P. (1999) *Structure Fold. Des. 7*, 853−864.
35. Mattinen, M. L., Linder, M., Drakenberg, T., and Annila, A. (1998) *Eur. J. Biochem. 256*, 279−286.
36. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
37. Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) *Biochemistry 36*, 16074−16086.
38. Surolia, A., Sharon, N., and Schwarz, F. P. (1996) *J. Biol. Chem. 271*, 17697−17703.
39. Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci U.S.A. 74*, 2236−2240.
40. Murphy, K. P., and Gill, S. J. (1990) *Thermochim. Acta 172*, 11−20.
41. Meissner, K., Wassenberg, D., and Liebl, W. (2000) *Mol. Microbiol. 36*, 898−912.
42. 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.

BI0101695