β-1,3-Glucan Binding by a Thermostable Carbohydrate-Binding Module from *Thermotoga maritima*†

Alisdair B. Boraston,*,‡,§,|| 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 September 10, 2001; Revised Manuscript Received October 3, 2001

ABSTRACT: The C-terminal 155 amino acids of the putative laminarinase, Lam16A, from *T. maritima* comprise a highly thermostable family 4 CBM that binds β-1,3- and β-(1,3)(1,4)-glucans. Laminarin, a β-1,3-glucan, presented two classes of binding sites for TmCBM4-2, one with a very high affinity (3.5×10^7 M^{-1}) and one with a 100-fold lower affinity (2.4×10^5 M^{-1}). The affinities for laminarioligosaccharides and β-(1,3)(1,4)-glucans ranged from $\sim 2 \times 10^5$ to $\sim 2.5 \times 10^6$ M^{-1}. Cellooligosaccharides and laminariobiose were bound only very weakly (K_as $\sim 5 \times 10^3$ M^{-1}). Spectroscopic and mutagenic studies implicated the involvement of three tryptophan residues (W28, W58, and W99) and one tyrosine residue (Y23) in ligand binding. Binding was enthalpically driven and associated with large negative changes in heat capacity. Temperature and osmotic conditions profoundly influenced binding. For the first time in solution, the direct uptake and release of water in CBM binding are demonstrated.

Carbohydrate-binding modules (CBMs)[1] are contiguous amino acid sequences within carbohydrate-active enzymes, most often glycoside hydrolases, which have discrete folds and carbohydrate-binding activity. Their biological role appears to be to mediate the association of glycoside hydrolases with polysaccharides, thereby enhancing the degradative process (*1*). There are currently 28 families of CBMs grouped by amino acid similarity (*2*). In total, these families show great variation in specificity including binding to cellulose, chitin, xylan, mannan, galactan, and starch. Thus, CBMs are excellent models for studies on the mechanisms of polysaccharide recognition by polypeptides.

CBM family 4 contains modules with varied polysaccharide binding specificity. CfCBM4-1 and CfCBM4-2 from *Cellulomonas fimi* Cel9B are specific for amorphous cellulose, cellooligosaccharides, and mixed β-(1,3)(1,4)-glucans (*3*); RmCBM4-1 and RmCBM4-2 from *Rhodothermus marinus* Xyn10A bind xylan and xylooligosaccharides in addition to amorphous cellulose, cellooligosaccharides, and mixed β-(1,3)(1,4)-glucans (*4*). Family 4 CBMs are also found in a β-1,3-glucanase (laminarinase) from the hyperthermophilic marine eubacteria *Thermotoga neopolitana*. These CBMs appear to be specific for β-1,3- and β-(1,3)(1,4)-glucans (*5*).

Thermotoga maritima produces an extracellular enzyme having laminarinase activity (*6*). ORF TM0024 identified by the *T. maritima* sequencing project (*7*) encodes a putative 642 amino acid polypeptide comprising N- and C-terminal modules with similarity to family 4 CBMs flanking a larger module with similarity to family 16 glycoside hydrolases. This putative enzyme, which has 87% amino acid sequence identity with a confirmed laminarinase from *T. neopolitana* (*8*), presumably accounts for the laminarinase activity of *T. maritima*. We show here that the C-terminal module, TmCBM4-2, from this *T. maritima* enzyme, here called Lam16A, is a carbohydrate-binding module. We present the first detailed binding studies of a CBM with specificity for β-1,3- and β-(1,3)(1,4)-glucans and extend these studies to include the effects of temperature and osmotic conditions. Furthermore, the role of aromatic amino acids in the binding activity of TmCBM4-2 is investigated by UV absorbance spectroscopy and site-directed mutagenesis.

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 was from Boehringer Mannheim (Laval, PQ, Canada). Isopropyl β-D-thiogalactoside (IPTG), arabinogalactan (larch), xylan (larch), laminarin [from *Laminaria digitata*; average degree of polymerization = 25 (*9*)], and cellobiose were from Sigma (St. Louis, MO). Glucose was from Fischer Scientific (Nepean, ON, Canada). Pectic-galactan (potato), oat β-glucan (DP 41 determined by reducing sugar analysis), barley β-glucan (MW 44 000; information provided by manufacturer), arabinan, xyloglucan (tamarind), and arabinoxylan (from wheat and rye) were from Megazyme (Dublin, U.K.). Birch wood xylan was from Carl Roth AG (Karlsruhe, Germany). Cellooligosaccharides were

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

* To whom correspondence should be addressed.

‡ Protein Engineering Network of Centres of Excellence.

§ The Biotechnology Laboratory.

|| Present address: Department of Chemistry, University of York, Heslington, York, United Kingdom YO10 5DD. Phone: 44 1904 434529; Fax: 44 1904 410519; E-mail: boraston@ysbl.york.ac.uk.

⊥ Department of Microbiology and Immunology.

[1] Abbreviations: AE, affinity electrophoresis; CBM, carbohydrate-binding module; CBD, cellulose-binding domain; CBM, carbohydrate-binding module; GuHCl, guanidine hydrochloride; IPTG, isopropyl-1-thio-β-D-galactopyranoside; K_a, association constant.

10.1021/bi015760g CCC: $20.00
Published on Web 11/07/2001

obtained from Seikagaku America Inc. (Falmouth, MA). Laminarioligosaccharides were from Dextra Laboratories Ltd. (Reading, Berkshire, U.K.).

Bacterial Strains, Plasmids, Media, and Growth Conditions. Thermotoga maritima strain MSB8 was from the ATCC. The *Escherichia coli* strains used were TOPP10 and BL21(DE3). The plasmid used was pET28a (Novagen, Madison, WI). *E. coli* cultures were grown routinely in TYP medium (*10*) 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. Bacterial transformations and agarose gel electrophoresis were done as described previously (*10*). Restriction endonucleases were used as recommended by the suppliers. Small-scale plasmid preparation was done using the Qiaprep Spin Mini-prep kits (Qiagen, Valencia, CA). DNA fragments were purified from the agarose gels with the Qiaex II DNA purification kit. Oligodeoxynucleotide primers were synthesized by the Nucleic Acid and Protein Service (NAPS) Unit (Biotechnology Laboratory, 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 TmCBM4-2. The DNA fragment encoding TmCBM4-2 was amplified by PCR from lysed *Thermotoga maritima* (MS8B) cells using a procedure described previously (*11*). The oligonucleotide primers used were 5′ACGGTCCCATGGGCTAGCATAAATAACGGCACTTTC 3′ (an *Nhe*I site is underlined) and 5′ CGACCGAAGCTT**TCA**TTGAGGGCTCACCGT 3′ (a *Hin*dIII site is underlined; the stop codon is in boldface), allowing amplification of nucleotides 1464−1929 of *lam16A*, encoding amino acids 488−643 of Lam16A (see GenBank accession no. AE001690; ORF TM0024). The purified fragment was digested with restriction endonucleases *Nhe*I and *Hin*dIII. It was ligated into pET28a that had been digested with the same enzymes to give pET-TmCBM4-2. The construct was verified by DNA sequencing after transformation into *E. coli* TOPP10. It was transformed into *E. coli* BL21(DE3) for production of TmCBM4-2. The polypeptide encoded by PET-TmCBM4-2 comprises a His_6 tag fused to the N-terminus of TmCBM4-2 by a thrombin cleavage site. Site-directed mutations were introduced using a "mega-primer" PCR method as described previously (*12*).

Production and Purification of TmCBM4-2 and Mutants. Six 2 L Erlenmeyer flasks containing 0.5 L of TYP medium, inoculated with *E. coli* BL21(DE3)/pET-TmCBM4-2, were grown at 30 °C to an $OD_{600\ nm}$ of ∼0.8. IPTG was added to 0.3 mM, and incubation was continued for 15 h. The cells were harvested by centrifugation (8500*g*) for 10 min at 4 °C and resuspended to about 1/50 of the original culture volume in 20 mM TrisHCl buffer, pH 7.9, containing 0.5 M NaCl. Cells were ruptured by two passages through a French pressure cell (144.9 MPa; 21 000 lb/in.^{2}), and cell debris was removed by centrifugation for 45 min at 27000*g* and 4 °C. TmCBM4-2 was purified from the clarified cell extract by immobilized metal affinity chromatography (IMAC) using a column containing 70 mL (packed volume) of His-Bind resin (Novagen, Madison, WI). Standard column preparation and purification procedures recommended by the resin supplier (Novagen, Madison, WI) were used. Protein fractions were analyzed for purity by SDS−PAGE (16% gels). Pure TmCBM4-2 fractions were pooled and exchanged into the water buffer and concentrated in a stirred ultra-filtration unit (Amicon, Beverly, MA) with 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 (*13*) of 45 660 M^{-1} cm^{-1}.

Affinity Electrophoresis. Affinity gel electrophoresis (AE) (*14*) was performed using a Mini-ProteanII system (Bio-Rad, Richmond, CA) with 0.75 mm spacers. The separating gels (7.5% polyacrylamide) contained 0.5% polysaccharide. TmCBM4-2 was applied (5−10 μg each lane) in native loading buffer (200 g of glycerol L^{-1}, 1 g of bromophenol blue L^{-1}, in 125 mM TrisHCl, pH 8.8). Then 5 μg of Native gel electrophoresis standards (Sigma) was used as a reference. Electrophoresis was at 4 °C and 150 V for 1 h. Proteins were visualized by staining with Coomassie brilliant blue. Data were collected and analyzed as described previously (*14*).

UV Difference Scans and Derivative Spectroscopy. UV scans were collected using a Varian Cary 100e UV−Vis spectrophotometer at wavelengths from 260 to 310 nm with a spectral bandwidth of 0.5 nm, an average integration time of 0.2 s, and a data interval of 0.2 nm. One milliliter of 20.6 μM TmCBM4-2 was scanned and corrected for a buffer-only baseline. A saturating amount of ligand (3−100-fold molar excess) was added, and the sample was allowed to equilibrate and rescanned. The latter scan was then corrected for dilution and a difference scan calculated by subtracting the protein-only scan from the protein−ligand scan.

Second-derivative spectra of the absorbance spectra were calculated using Origin v.6.0 (MicroCal, Northhampton, MA). Scans of denatured TmCBM4-2 and CfCBM4-1 were taken in 4 M guanidine hydrochloride (GuHCl) at 70 °C. Tyrosine exposure in native and ligand-bound forms was calculated using the method established by Ragone et al. (*15*).

Quantitative UV Analysis of Substrate Binding. Absorbance spectra were collected on a S2000 CCD array fiber optic spectrometer with a MINI-D2 deuterium UV light source and water-thermostated cuvette holder (Ocean Optics, Dunedin, FL). The temperature was controlled by a recirculating heating/cooling bath (Forma Scientific, Marietta, OH), 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. Ligand was delivered from Hamilton syringes by a KDS200 computer-controlled micrometer driven syringe pump (KD Scientific, New Hope, PA). Data collection and pump control were automated and synchronized by software developed in-house.

Absorbance data from titration experiments were processed as described previously (*16*). The trough to peak heights at $\Delta A_{297.8} - \Delta A_{292.1}$ nm (for laminaroligosaccharides, laminarin, and oat β-glucan) and $\Delta A_{301} - \Delta A_{293}$ nm (for glucose, laminariobiose, and cellooligosaccharides) were plotted versus total ligand concentration. Binding parameters were determined as described below.

The general binding equilibrium including a stoichiometric parameter is derived as follows:

$$P + L \leftrightarrow PLn \quad (1)$$

$$K_a = \frac{[PLn]}{[P][L]} \quad (2)$$

$$P_T = [P] + [PLn] \quad (3)$$

$$L_T = [L] + n[PLn] \quad (4)$$

where P and L represent free protein and free ligand, respectively. P_T and L_T are the total protein and ligand, respectively. PL*n* represents the protein−ligand complex where *n* mol of ligand is bound to 1 mol of protein. The UV difference signal upon binding can be derived as

$$\Delta A = \frac{\Delta A_{PLn} \times K_a \times L}{1 + K_a \times L} \quad (5)$$

where ΔA is the UV difference signal, K_a is the association constant, ΔA_{PLn} is the maximal UV difference signal of the complexed protein, and L is the free ligand concentration. L is unknown but can be solved for by substituting eqs 3 and 4 into eq 2 and solving the quadratic equation for L to give

$$L = 0.5 \times [(L_T - 1/K_a - n \times P_T) + \sqrt{(n \times P_T - L_T + 1/K_a) + 4 \times L_T/K_a}] \quad (6)$$

The binding parameters n, K_a, and ΔA_{PLn} can then be determined by simultaneous nonlinear regression of the data to eqs 5 and 6. For ligands where experimental stoichiometries are reported, C values [$C = K_a \times A \times n$, where K_a is the association constant, A is the acceptor concentration (i.e., TmCBM4-2), and n is the number of binding sites on the acceptor] were always greater than 4 and usually greater than 10, in which case binding approximated a pseudo-first-order reaction. For binding experiments using laminarin, the data were converted to a molar change in UV difference per mole of laminarin injected, effectively the derivative data, and plotted versus the molar ratio of laminarin to TmCBM4-2. These data were analyzed with MicroCal MCS Origin (MicroCal, Northhampton, MA) using a two binding site model and treating the laminarin as the macromolecule and the CBM as the ligand.

Calculation of Thermodynamic Parameters. Gibbs free energy (ΔG) was calculated from the K_a values using the fundamental equation:

$$\Delta G = -RT \ln(K_a) \quad (7)$$

Enthalpy (ΔH), entropy (ΔS), and heat capacity (ΔC_p) changes were calculated by nonlinear regression of $\ln(K_a)$ vs temperature data with the following equation:

$$\ln(K_a) = \frac{\Delta H - T\Delta S + \Delta C_p\left(T - T^* - T \ln\frac{T}{T^*}\right)}{-RT} \quad (8)$$

where T is the experimental temperature and T^* is an arbitrary reference temperature kept as a constant (e.g., 298.15 K) in the analysis.

Osmotic Stress. Osmotic stress experiments were performed and analyzed as described previously (*17*, *18*) using association constants determined by UV difference binding experiments. Glycerol solutions of the desired osmolality were prepared by weight in 25 mM TrisHCl, pH 7.0.

Thermal Melts. Samples of 20 μM TmCBM4-2 were prepared in 25 mM TrisHCl, pH 7.0, in concentrations of guanidine hydrochloride (GuHCl) ranging from 0 to 4.5 M. Additional sets of samples contained, in addition to buffer and GuHCl, either 10 mM $CaCl_2$ or 10 mM EDTA. Thermal melts with these samples were performed and analyzed as described previously (*12*) using a wavelength of 285 nm to monitor unfolding. T_m values were plotted vs GuHCl concentration, and linear regression was used to extrapolate the T_m at zero GuHCl concentration (*19*).

RESULTS AND DISCUSSION

Polysaccharide Binding Assessed by Affinity Electrophoresis. TmCBM4-2 bound tightly to β-1,3-glucan (laminarin) and β-(1,3)(1,4)-glucan (oat β-glucan), but weakly to xylans (birchwood, wheat, rye, and oat spelt) and xyloglucan (from tamarind)(results not shown). It did not bind to dextran, arabinan, arabino-galactan (larch), pectic-galactan (potato), or galactan (gum arabic). *T. maritima* is a marine bacterium; thus, it is most likely that β-1,3-glucans, such as those commonly found as structural polysaccharides in marine algae, are the biological ligands for TmCBM4-2 and the substrates for Lam16A. However, information available about the extreme environments this bacterium is found in is sparse, making it difficult to confidently discuss the true biological sources of polysaccharides that come in contact with *T. maritima*.

UV Difference Spectroscopy. The UV absorbance spectrum of TmCBM4-2 was significantly perturbed in the presence of ligand (Figure 1). The difference spectra with glucose, cellooligosaccharides, and laminaribiose as ligands were very similar to the DMSO-perturbed difference spectrum of *N*-acetyltryptophan (Figure 1), suggesting the involvement of only tryptophan in the interaction of these ligands with TmCBM4-2. The difference spectra with laminarin, laminarioligosaccharides (three glucose units or greater), and β-(1,3)(1,4)-glucans as ligands resembled the difference spectrum of perturbed *N*-acetyltryptophan; however, the 293 nm peak and 289 nm trough of *N*-acetyltryptophan are shifted to slightly shorter wavelengths, probably due to contribution from a 286.5 nm tyrosine peak. The 279 nm−281 nm peak−trough pair found in the DMSO-perturbed difference spectrum of *N*-acetyltyrosine appears in the difference spectra of TmCBM4-2 with these ligands. This suggests that these interactions involve both tryptophans and tyrosines.

Second-derivative UV absorption analysis of tyrosine exposure in native TmCBM4-2 indicated that five of the nine tyrosine residues are solvent-exposed (not shown). In the presence of laminarin, laminarioligosaccharides (three glucose units or greater), and β-(1,3)(1,4)-glucan, the number of exposed tyrosines decreases to approximately 3.3. This is consistent with the involvement of tyrosines in binding these ligands and suggests that one or two tyrosine residues are involved in laminarin, laminarioligosaccharide, and β-(1,3)(1,4)-glucan binding by TmCBM4-2. In contrast, laminaribiose and cellooligosaccharides, regardless of length,

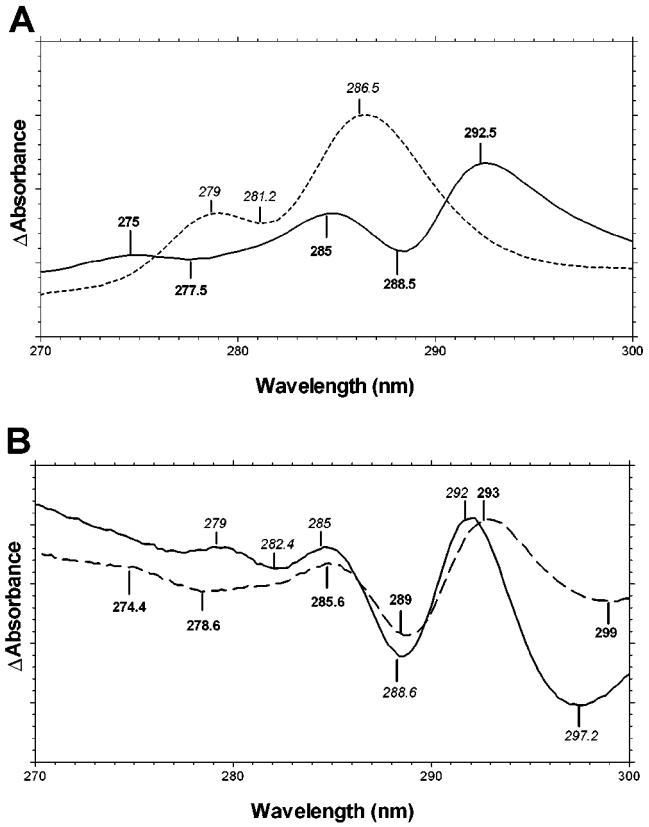

FIGURE 1: UV difference spectra of TmCBM4-2 complexed with soluble ligands. Panel A: Difference spectrum of *N*-acetyltryptophan (solid line) and *N*-acetyltyrosine (dashed line) perturbed in 20% DMSO. Samples containing 50 μM *N*-acetyltryptophan or 100 μM *N*-acetyltyrosine in 25 mM TrisHCl, pH 7.5, were used to collect baseline spectra. Samples of the same concentrations but containing 20% DMSO were used to collect difference spectra. Peak and trough wavelengths are labeled in boldface and italics for *N*-acetyltryptophan and *N*-acetyltyrosine, respectively. Panel B: Difference spectrum of TmCBM4-2 with laminarin (solid line) and laminariobiose (dashed line). Peak and trough wavelengths are labeled in boldface and italics for laminariobiose and laminarin, respectively.

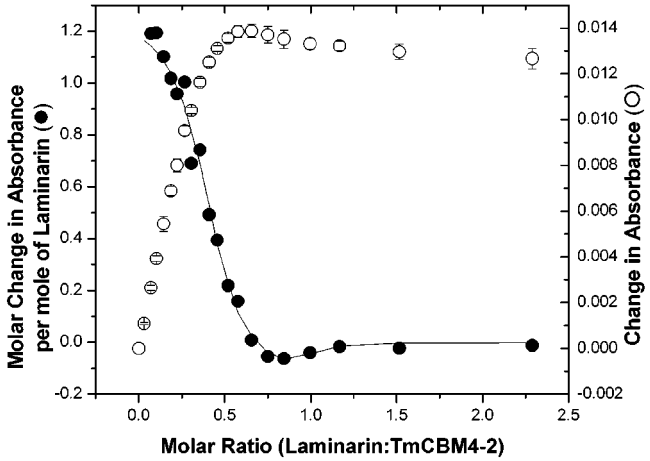

FIGURE 2: Titration of laminarin (2 mg/mL; 491.6 μM) into TmCBM4-2 (16.7 μM). The right axis shows unmodified cumulative UV difference data using wavelengths at 297.8 and 292.1 nm (○). Error bars on the cumulative data represent the standard deviations of triplicate experiments. The left axis shows the UV difference data as derivative data (●) and the best-fit line determined by nonlinear least-squares analysis to a two binding-site model. Samples were buffered in 25 mM Tris, pH 7.0.

Table 1: Binding Parameters at 25 °C for the Association of TmCBM4-2 with Soluble Ligands[a]

	K_a ($\times 10^{-6}$ M^{-1})	ΔG (kJ/mol)	n[b]
barly β-glucan (DP 271)	2.98 ± 0.55	−36.90 ± 0.44	0.043 ± 0.005
lamina riheptaose	0.77 ± 0.03	−33.59 ± 0.09	0.97 ± 0.15
laminarihexaose	1.11 ± 0.08	−34.50 ± 0.16	1.14 ± 0.07
laminaripentaose	2.46 ± 0.36	−36.46 ± 0.37	1.03 ± 0.04
laminaritetraose	0.89 ± 0.09	−33.94 ± 0.25	1.36 ± 0.07
laminaritriose	0.18 ± 0.03	−30.03 ± 0.36	1.78 ± 0.01
laminaribiose	0.0049 ± 0.0007	−21.02 ± 0.35	1
glucose	0.0074 ± 0.0003	−22.09 ± 0.10	1
cellohexaose	0.0053 ± 0.001	−21.21 ± 0.48	1
cellopentaose	0.0043 ± 0.0003	−20.72 ± 0.18	1
cellotetraose	0.0050 ± 0.0003	−21.11 ± 0.16	1
cellotriose	0.0034 ± 0.0001	−20.17 ± 0.09	1
cellobiose	0.0034 ± 0.0007	−20.15 ± 0.50	1

[a] Errors represent the standard deviations of 4−6 repeated titrations. [b] n, number of sugar molecules per TmCBM4-2 molecule. Values for laminariobiose, glucose, and cellooligosaccharides could not be experimentally determined and were assumed to be 1.

do not affect the tyrosine exposure in TmCBM4-2, supporting the lack of tyrosine involvement in their binding.

Polysaccharide Binding. Laminarin is a soluble β-1,3-glucan with an average degree of polymerization of 25, approximately 1 β-1,6-linked glucose branch for every 20 backbone residues, and terminal mannitol residues (*9*). The UV difference titration of this polysaccharide into TmCBM4-2 gave binding curves that suggested multiple classes of binding sites (Figure 2). When converted to derivative data, the binding curves fit a binding model with two classes of sites. The high-affinity site had an association constant of 3.5 (±1.6) × 10^7 M^{-1} and occurred on average 1.3 (±0.1) times per laminarin molecule (i.e., 1 binding site per 20 backbone sugar residues). The low-affinity site had an association constant of 2.4 (±0.8) × 10^5 M^{-1} and also occurred on average 1.3 (±0.2) times per laminarin molecule. These values were determined from the analysis of titrations performed in triplicate using three different concentrations of TmCBM4-2 (i.e., nine titrations). Laminarin forms tertiary and quaternary structures of single and triple helices, respectively (*20*). The results of binding experiments to NaOH-denatured laminarin (*20*) were qualitatively similar to native laminarin, suggesting that the conformation adopted by laminarin is not responsible for the multiple classes of binding sites. It is unknown if the β-1,6-linked glucose branches or terminal mannitol residues (*9*) play a role in laminarin recognition, but at this point it appears likely.

TmCBM4-2 also bound very tightly to barley β-glucan (Table 1), a mixed-linkage glucan consisting, on average, of cellotriose and cellotetraose units separated by β1−3 glycosidic bonds. Binding sites for TmCBM4-2 on barley β-glucan were much more frequent (~1 site per 12 glucoses) than for CfCBM4-1 (~1 site per 25) (*3*) or CBM17 (~1 site per 30) (*16*), which are β1,4-glucan-binding modules. TmCBM4-2 did not bind well to oligosaccharides consisting solely of β-1,4-linked glucose, suggesting that the β1−3 linkages in barley β-glucan are highly important in its recognition. Based on the frequency of binding sites and the probable number of sugar residues that occupy the binding site (5−6; see below), TmCBM4-2 could span one or two β1−3 linkages. Clearly, TmCBM4-2 must also be able to accommodate some β-1,4-linked glucose residues while

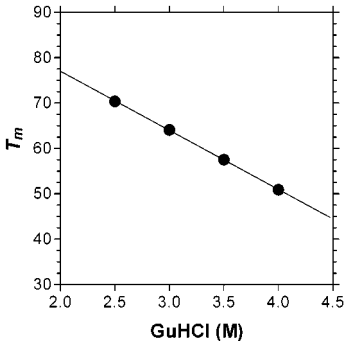

FIGURE 3: Melting temperature of TmCBM4-2 versus GuHCl concentration. Only data in the absence of EDTA and $CaCl_2$ are shown for clarity. Error bars are well within the sizes of the data points.

maintaining a high-affinity interaction. In contrast, the low frequency of binding sites for CfCBM4-2 and CBM17 suggests a more specific requirement for binding, probably extended stretches of β-1,4-linked glucose.

Oligosaccharide Binding. At 25 °C, TmCBM4-2 bound with association constants of $\sim 10^5-10^6$ M^{-1} to laminarioligosaccharides (>2 glucose units in length) (Table 1). Laminaribiose and cellooligosaccharides were bound very weakly. With most CBMs, increases in oligosaccharide length correlate with increases in binding affinity until the saccharide is of sufficient length to occupy the entire binding site, at which point the affinity ceases to improve (*3*, *16*, *21*, *22*). The affinity of TmCBM4-2 clearly peaks at an oligosaccharide length of 5, beyond which the affinity unexpectedly decreases. The reason for this is unknown. The affinity of TmCBM4-2 for laminarioligosaccharides (>2 glucose units in length) and mixed β-(1,3)(1,4)-glucans was on the order of 10−100-fold greater than the affinity of CfCBM4-1 for cellooligosaccharides and mixed β-(1,3)(1,4)-glucans (*3*, *23*).

Within experimental error, the binding stoichiometries of laminaripentaose, laminarihexaose, and laminariheptaose to TmCBM4-2 were 1:1 (Table 1). The binding stoichiometries to laminaritriose and laminaritetraose deviated significantly from the expected 1:1 ratio. The values greater than unity suggest that on average more than one ligand molecule is bound to a TmCBM4-2 molecule. Interestingly, when these values are converted to moles of glucose per mole of TmCBM4-2, the stoichiometries were 5.3 (±0.3) and 5.4 (±0.3) for laminaritriose and laminaritetraose, respectively. These results suggest that on average 5−6 glucose residues occupy the TmCBM4-2-binding site.

Thermal Stability of TmCBM4-2 and Binding at Elevated Temperatures. TmCBM4-2 is extremely thermostable with a T_m of 103.0 (±0.4) °C (Figure 3). The T_m was not affected by the presence of 10 mM $CaCl_2$ [T_m of 102.1 (±2.0) °C] but was reduced to 96.4 (±0.9) °C in the presence of 10 mM EDTA (not shown). The ~7 °C reduction in T_m in the presence of EDTA suggests that TmCBM4-2 contains one or more metal-binding sites that contribute to the thermostability of the polypeptide. The lack of effect of $CaCl_2$ on the T_m suggests that the untreated preparation of TmCBM4-2 was fully saturated with calcium or that the bound metal is not calcium. These results also show that TmCBM4-2 remains fully folded at temperatures exceeding those at which *T. maritima* is viable.

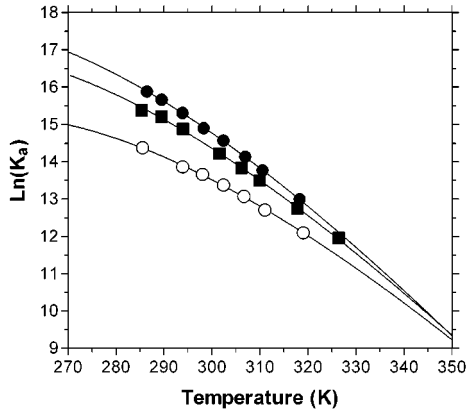

FIGURE 4: ln(K_a) versus temperature plots of TmCBM4-2 binding to laminaritetroase (○), laminariopentaose (●), and laminarihexaose (■). Errors calculated as the standard deviation of 4−6 titrations were well within the size of the symbols. Solid lines represent best-fit lines determined by nonlinear regression of the data to eq 8.

Table 2: van't Hoff Thermodynamic Parameters Calculated at 25 °C (298.15 K) for the Association of TmCBM4-2 with Soluble Ligands[a]

	ΔH (kJ/mol)	ΔS (J/mol/K)	$-\Delta C_p$ (J $mol^{-1} K^{-1}$)
laminarihexaose	−59.4 ± 2.6	−78.5 ± 8.8	1100 ± 350
laminaripentaose	−64.9 ± 2.1	−93.5 ± 7.1	1170 ± 430
laminaritetraose	−47.7 ± 1.7	−46.6 ± 5.8	1060 ± 280

[a] Errors represent the standard errors of regressed data.

The van't Hoff thermodynamic values for TmCBM4-2 binding to laminarioligosaccharides were determined by fitting eq 8 to the natural logarithm of K_as determined at several different temperatures (Figure 4) (*24*). At 25 °C, all of the interactions had favorable changes in enthalpy (ΔH) with partially compensating unfavorable changes in entropy ($T\Delta S$) (Table 2). This observation is highly common for carbohydrate-binding proteins, including CBMs (*3*, *11*, *22*, *25*).

Within experimental error, the ΔC_p values for laminaritetraose, laminaripentaose, and laminarihexaose were the same (Table 2). ΔC_p is postulated to depend on the change in polar and apolar solvent-accessible surface area upon ligand binding (*26*). This suggests the burial of similar and predominantly apolar surface areas with the accompanying solvent rearrangement upon binding of these ligands (*27*, *28*). Thus, the binding site of TmCBM4-2 is predicted to contain a relatively high proportion of amino acids with apolar side chains.

Most protein−carbohydrate interactions give relatively small changes in ΔG over a temperature range. However, the fundamental relationship $\Delta G = -RT \ln(K_a)$ dictates that these small changes give dramatic decreases in K_a with increasing temperature, especially if ΔC_p is significantly negative. Assuming that one of the basic biological functions of CBMs is to enhance the association of the catalytic module with the substrate, the above thermodynamic condition is detrimental for CBMs found in *T. maritima*. The affinity of TmCBM4-2 for laminarioligosaccharides decreased approximately 15-fold when going from 25 to 45 °C (Figure 4). As this CBM is stable to temperatures above boiling, the K_as can be extrapolated to 90 °C using the determined thermodynamic values. At this temperature, the K_as are

Table 3: Binding of TmCBM4-2 Mutants to Laminariohexaose at 25 °C[a]

	K_a ($\times 10^{-6}$ M^{-1})	$\Delta\Delta G$ (kJ/mol)[b]
wild-type	1.11 ± 0.08	–
W24A	0.007 ± 0.001	12.66 ± 0.70
W58A	0.019 ± 0.002	10.08 ± 0.30
W99A	0.037 ± 0.001	8.45 ± 0.19
Y29A	0.037 ± 0.008	8.72 ± 0.60
Y45A	1.24 ± 0.1	−0.27 ± 0.15
Y47A	1.29 ± 0.03	−0.37 ± 0.06

[a] Errors represent the standard deviations of 2−3 repeated titrations. [b] Calculated relative to wild-type.

predicted to decrease approximately 1000-fold. This may in part explain why this module has evolved to have an affinity for soluble ligands that is approximately 100-fold higher or, in the case of binding to the high-affinity site of laminarin, 1000-fold higher than the affinity of mesophilic CBMs for their respective soluble ligands (*3*, *16*, *22*, *29*).

Role of Aromatics in Ligand Binding. TmCBM4-2 has approximately 20% identity and 35% similarity with CfCBM4-1, the well-characterized N-terminal family 4 CBM from *Cellulomonas fimi* Cel9B (*3*, *23*, *30*, *31*). The three binding site aromatic amino acids of CfCBM4-1, Y19, Y43, and Y85, are conserved with W24, W58, and W99 of TmCBM4-2, suggesting these tryptophans play a role in ligand binding by TmCBM4-2. The comparison of conserved residues does not give any clue as to the identity of the tyrosine residue(s) involved in binding. However, based on the structure of CfCBM4-1 (*30*), Y29, Y45, and Y47 of TmCBM4-2 are predicted to be near the putative binding site.

The involvement of W24, W58, W99, Y29, Y45, and Y47 in ligand binding was probed by site-directed mutagenesis. The W24A, W58A, W99A, and Y29A mutants had a markedly lower affinity for laminariohexaose than wild-type TmCBM4-2 (Table 3), suggesting their involvement in binding. The Y39A, Y45A, and Y47A mutants had unchanged affinity. The T_ms of the mutants in 4 M GuHCl were unchanged relative to wild-type, suggesting that the mutations did not grossly perturb the polypeptide fold (not shown). Unlike all of the other mutants, UV difference and second-derivative analysis of the Y29A mutant in the presence of laminariohexaose did not indicate any participation of tyrosines in ligand binding and suggest that Y29 is the sole tyrosine participating in ligand binding. These four aromatic amino acid side chains would provide a considerable apolar surface area in the binding site and contribute to the large negative ΔC_ps of binding. The large effect on binding of removing of these aromatic amino acid side chains from the interaction is consistent with the results found for many CBMs (*31*−*34*).

Role of Water in Ligand Binding. The participation of water in ligand binding by TmCBM4-2 was investigated by osmotic stress (*17*, *35*). The K_as decreased with increasing osmolyte concentration (Figure 5), indicating the net uptake of ∼5 water molecules for laminaritriose to laminaripentaose (Table 4). In contrast, there was a net of ∼0 water molecules involved in laminarihexaose and laminariheptaose binding. This indicates the release of ∼5 water molecules resulting from binding the sixth sugar in laminarihexaose. In apparent contradiction to the net requirement of no water molecules in the binding of longer laminariooligosaccharides, an

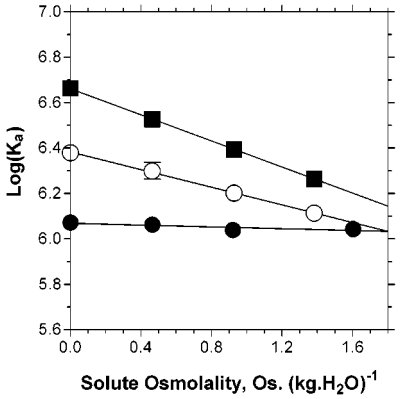

FIGURE 5: Osmotic sensitivity of the association constant for the binding of TmCBM4-2. Only data for laminarihexaose (●), laminariopentaose (○), and laminarin (■) as a function of neutral solute osmolality at 25 °C are shown for clarity. Titration data for laminarin were fit to a single binding-site model so K_a [and log-(K_a)] values approximate the average affinity of TmCBM4-2 for binding sites on laminarin. Straight lines were obtained by linear regression. Error bars represent the standard deviations of four titrations.

Table 4: Changes in the Number of Solute-Excluding Water Molecules (Δn_w) Coupled to the Binding of Sugars to TmCBM4-2 as a Function of Osmotic Stress[a]

	Δn_w
laminarin (DP 25)[b]	7.01 ± 0.04
laminariheptaose	0.08 ± 0.45
laminarihexaose	0.48 ± 0.69
laminaripentaose	4.72 ± 0.13
laminaritetraose	5.01 ± 0.41
laminaritriose	4.81 ± 0.74

[a] Errors represent the standard errors of regressed data. [b] Represents the approximate average number of water molecules coupled to binding at both the high- and low-affinity sites of laminarin.

average of ∼7 water molecules are taken up upon binding of laminarin. The strong need for water during binding of this polysaccharide may relate to its triple helical structure or possibly to the potential role of the β-1,6-glucose substitutions, though it is not clear how. Previously, the requirement of water in ligand binding would only have become evident in high-resolution structural studies. This is the first demonstration of the direct participation of water in ligand binding by a CBM and one that shows requirements for the uptake and release of water by different portions of the ligand.

Conclusions. TmCBM4-2 has amino acid similarity with family 4 CBMs and is a β-1,3-glucan-binding module that can also accommodate β-(1,3)(1,4)-glucans. This is in keeping with its presence in a putative laminarinase (β-1,3-glucanase) and the presence of closely related modules in other laminarinases or lichenases (*5*, *36*). Thus, as with other CBMs, it can be hypothesized that this module functions to target the enzyme to the substrate on which it is most active.

The environmental conditions that a glycoside hydrolase functions in are important not only to the catalytic module but also to the CBM. The effect of temperature on ligand binding is well-known for carbohydrate-binding proteins that bind small sugars but, except in one case (*11*), this is considered over a relatively narrow temperature range. This is the first demonstration of the temperature dependence of the association constant over a broad temperature range for

a true polysaccharide-binding CBM. It shows that for thermostable CBMs, especially ones such as TmCBM4-2 where the ΔC_ps of binding are large and negative, the temperature at which it functions profoundly influences the equilibrium. While the impact of this phenomenon on the hydrolytic activity of polysaccharidases is currently unknown, it is clearly a factor that cannot be overlooked. The requirement for the uptake and/or release of water upon CBM binding is also an important consideration. The osmotic conditions of an environment are influenced by the presence of osmolytes, such as minerals, salts, or even carbohydrates, or simply by the lack of water. This osmotic state can influence association constants positively or negatively by severalfold and is a previously overlooked aspect in the biological function of CBMs and glycoside hydrolases.

NOTE ADDED AFTER ASAP POSTING

This article was inadvertently released ASAP on 11/07/01 before final corrections were made. In Table 2 the fourth column heading should read $-\Delta C_p$ (J mol^{-1} K^{-1}). The negative sign was originally omitted. The correct version was posted 11/12/01.

ACKNOWLEDGMENT

We thank Patrick Chiu and Alphonsus Au for technical assistance.

REFERENCES

1. 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.
2. 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.
3. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
4. Abou-Hachem, M., Nordberg-Karlsson, E., Bartonek-Roxa, E., Raghothama, S., Simpson, P. J., Gilbert, H. J., Williamson, M. P., and Holst, O. (2000) *Biochem. J. 345*, 53−60.
5. Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, G. A., and Schwarz, W. H. (2001) *Microbiology 147*, 621−629.
6. Bronnenmeier, K., Kern, A., Liebl, W., and Staudenbauer, W. L. (1995) *Appl. Environ. Microbiol. 61*, 1399−1407.
7. Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., and Fraser, C. M. (1999) *Nature 399*, 323−329.
8. Zverlov, V. V., Volkov, I. Y., Velikodvorskaya, T. V., and Schwarz, W. H. (1997) *Microbiology 143*, 1701−1708.
9. Read, S. M., Currie, G., and Bacic, A. (1996) *Carbohydr. Res. 281*, 187−201.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY.
11. Boraston, A. B., Creagh, A. L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. A., Warren, R. A., and Kilburn, D. G. (2001) *Biochemistry 40*, 6240−6247.
12. Boraston, A. B., Tomme, P., Amandoron, E. A., and Kilburn, D. G. (2000) *Biochem. J. 350*, 933−941.
13. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) *Anal. Biochem. 200*, 74−80.
14. Tomme, P., Boraston, A., Kormos, J. M., Warren, R. A., and Kilburn, D. G. (2000) *Enzyme Microb. Technol. 27*, 453−458.
15. Ragone, R., Colonna, G., Balestrieri, C., Servillo, L., and Irace, G. (1984) *Biochemistry 23*, 1871−1875.
16. Boraston, A. B., Chiu, P., Warren, R. A. J., and Kilburn, D. G. (2000) *Biochemistry 39*, 11129−11136.
17. Swaminathan, C. P., Surolia, N., and Surolia, A. (1998) *J. Am. Chem. Soc. 120*, 5153−5159.
18. Swaminathan, C. P., Nandi, A., Visweswariah, S. S., and Surolia, A. (1999) *J. Biol. Chem. 274*, 31272−31278.
19. Guo, Y., Kammerer, R. A., and Engel, J. (2000) *Biophys. Chem. 85*, 179−186.
20. Young, S. H., Dong, W. J., and Jacobs, R. R. (2000) *J. Biol. Chem. 275*, 11874−11879.
21. 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.
22. 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.
23. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
24. van den Elsen, J. M., van Unen, L. M., van Bloois, L., Busquets, M. A., Jiskoot, W., Hoogerhout, P., Wilting, J., Herron, J. N., and Crommelin, D. J. (1997) *Biochemistry 36*, 12583−12591.
25. Charnock, S. J., and Davies, G. J. (1999) *Biochemistry 38*, 6380−6385.
26. Edgcomb, S. P., and Murphy, K. P. (2000) *Curr. Opin. Biotechnol. 11*, 62−66.
27. Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2236−2240.
28. Murphy, K. P., and Gill, S. J. (1990) *Thermochim. Acta 172*, 11−20.
29. 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.
30. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry 35*, 14381−14394.
31. Kormos, J., Johnson, P. E., Brun, E., Tomme, P., McIntosh, L. P., Haynes, C. A., and Kilburn, D. G. (2000) *Biochemistry 39*, 8844−8852.
32. 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.
33. Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghothama, S., Fontes, C. M., Dias, F. M., Ferreira, L. M., and Bolam, D. N. (2001) *Biochemistry 40*, 9167−9176.
34. Ponyi, T., Szabo, L., Nagy, T., Orosz, L., Simpson, P. J., Williamson, M. P., and Gilbert, H. J. (2000) *Biochemistry 39*, 985−991.
35. Parsegian, V. A., Rand, R. P., and Rau, D. C. (2000) *Proc. Natl Acad. Sci. U.S.A 97*, 3987−3992.
36. Sunna, A., Gibbs, M. D., and Bergquist, P. L. (2001) *Biochem. J. 356*, 791−798.

BI015760G