Identification and glucan-binding properties of a new carbohydrate-binding module family

Alisdair B. BORASTON^{*†‡}, Mazyar GHAFARI[‡], R. Antony J. WARREN^{*‡} and Douglas G. KILBURN^{*†‡}

^{*}The Protein Engineering Network of Centres of Excellence, PENCE Inc., National Business Centre, 750 Heritage Medical Research Centre, Edmonton, Alberta, Canada T6G 2S2, [†]The Biotechnology Laboratory, 237-6174 University Boulevard, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3, and

[‡]Department of Microbiology and Immunology, 300-6174 University Boulevard, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3

The C-terminal 191-residue module of Cel5A from the alkalophilic *Bacillus* sp. 1139 comprises a carbohydrate-binding module (CBM) belonging to a previously unidentified family that we have classified as CBM family 28. This example, called CBM28, bound specifically to cello-oligosaccharides and mixed β -(1,3)(1,4)-glucans (barley β -glucan) with association constants of approximately $(1-4) \times 10^4 \text{ M}^{-1}$. Its binding to barley β -glucan was remarkably insensitive to pH between 7.0 and 10.9, in keeping with its alkalophilic source. CBM28 bound to cellulose having a significant non-crystalline content with an association constant similar to that for its binding to soluble glucans.

CBM17 (CBM family 17) and CBM28 modules naturally occur as tandems. The CBM17/CBM28 tandem from Cel5A bound with apparent co-operativity to barley β -glucan. The association of CBM28 with cello-oligosaccharides was driven enthalpically and marked by the different thermodynamic contribution of three putative binding subsites that accommodate a cellobiose molecule.

Key words: association, alkalophilic, cellulose, co-operativity, thermodynamics.

INTRODUCTION

Glycoside hydrolases are modular proteins comprising one or more catalytic modules and often multiple ancillary modules [1]. The most common ancillary modules are the carbohydrate-binding modules (CBMs), which mediate the interaction of these enzymes with polysaccharides. In many cases this function of CBMs has been shown to be important to the efficient degradative action of glycoside hydrolases [2–4]. Thus identifying new CBMs and studying the mechanisms of CBM–polysaccharide interaction is fundamental in the process of understanding polysaccharide degradation.

Many family 5 glycoside hydrolases from clostridia and alkalophilic bacilli share a common modular structure: CD5/CBM17/X11 (designated in accordance with [5]), where X11 represents a family of modules with undefined function. X11 modules are nearly always found in tandem with a CBM17 module immediately preceding it, a motif that suggests related functions. On the basis of this it was hypothesized that X11 modules are CBMs. Indeed, we show here that the family 11 X-module from the alkalophilic *Bacillus* sp. 1139 Cel5A is a CBM and we describe its glucan-binding properties. The family including this module has been classified as CBM family 28. Thus we shall refer to this CBM from *Bacillus* sp. 1139 Cel5A as CBM28.

MATERIALS AND METHODS

Carbohydrates and polysaccharides

Barley β -glucan (molecular mass standard 44 kDa; degree of polymerization 271) was purchased from MegaZyme Ltd (Bray, Co. Wicklow, Ireland). Cello-oligosaccharides were from

Seikagaku (Tokyo, Japan). Bacterial microcrystalline cellulose was prepared as described previously [6]. Regenerated cellulose was prepared from Avicel[®] as described previously [7]. Avicel[®], PH101, was obtained from FMC International (Little Island, County Cork, Ireland).

Preparation of genomic DNA

Bacillus sp. 1139 (ATCC 43226) was grown for 48 h at 37 °C in 50 ml of ATCC culture medium 1513 PY-CMC (consisting of 10 g/l carboxymethylcellulose, 5 g/l peptone, 5 g/l yeast extract, 5 g/l NaCl, 1 g/l K_2HPO_4 , 2 g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ and 10 g/l Na_2CO_3). The cells were pelleted by centrifugation and washed with 2 ml of TE buffer [25 mM Tris/HCl (pH 8.0)/10 mM EDTA]. The cells were then taken up in 0.7 ml of TE buffer containing 50 mM glucose, 0.5% SDS, 10 mg/ml lysozyme and 1 mg/ml RNase A and incubated at 37 °C for 4 h. An equal volume of phenol/chloroform (1:1, v/v) was added and the mixture was vortex-mixed, then centrifuged for 30 min at 30000 g. This procedure was repeated three times. Genomic DNA was precipitated by the addition of NaCl to 0.5 M and propan-2-ol to 40% (v/v), then recovered by spooling the precipitated material on the tip of a glass Pasteur pipette. The genomic DNA was air-dried at 22 °C and dissolved in TE buffer.

DNA amplification and cloning of CBM28 and CBM17/CBM28

Gene fragments encoding amino acid residues 571–762 and 388–762 of *Bacillus* sp. 1139 cellulase 5A (GenBank[®] accession number D00066) were obtained and amplified by PCR from genomic DNA. An *Nhe*I site (underlined) was introduced at the

Abbreviations used: CBM, carbohydrate-binding module; CBM28, family 28 carbohydrate-binding module from *Bacillus* sp. 1139 Cel5A; CBM17/CBM28, family 17 and family 28 CBM tandem from *Bacillus* sp. 1139 Cel5A; ITC, isothermal titration calorimetry; K_a , association constant.

[†] To whom correspondence should be addressed. Present address: Department of Chemistry and the York Structural Biology Laboratory, University of York, Heslington, York YO10 5DD, U.K. (e-mail boraston@ysbl.york.ac.uk).

5' end of the *CBM28* gene fragment by using the oligonucleotide 5'-AGCCATATGGCTAGCGGAACAGAAGTTGAAATT-3' (*CBM28F*) as a primer. A *HindIII* (underlined) restriction site was introduced at the 3' end of the *CBM28* gene fragment sequence by using the oligonucleotide 5'-TGCGGCCGCAAGCTTTCAGCAGCCCCCTCAAAGCA-3' (*CBM28R*) as primer. An *NheI* site (underlined) was introduced at the 5' end of the *CBM17/CBM28* gene fragment by using the oligonucleotide 5'-AGCCATATGGCTAGCAAAGGTGTGAAGTATGAG-3' (*CBM17F*) as a primer and *CBM28R* as a flanking primer. The resulting 0.56 kb (*CBM28*) and 1.12 kb (*CBM17/CBM28*) PCR fragments were digested with *NheI* and *HindIII* and ligated into pET28a (Novagen, Madison, WI, U.S.A.) previously digested with the same restriction enzymes, to give pET-*CBM28* and pET-*CBM17/CBM28*. DNA was sequenced by the NAPS Unit (Biotechnology Laboratory, University of British Columbia) with the AmpliTaq dye-termination cycle-sequencing protocol and an Applied Biosystems Model 377 sequencer. The polypeptides encoded by pET-*CBM28* and pET-*CBM17/CBM28* comprise a His₆ tag fused to the N-terminus by a thrombin cleavage site.

Production and purification

Erlenmeyer flasks (2-litre) containing 0.5 litre of TYP medium [8], inoculated with *Escherichia coli* BL21(DE3) containing pET-*CBM28* or pET-*CBM17/CBM28* constructs, were incubated at 30 °C to a D_{600} of approx. 0.8. Isopropyl β -D-thiogalactoside was added to 0.5 mM; incubation continued for 15 h. The cells were harvested, resuspended to approx. 1/50 of the original culture volume in 20 mM Tris/HCl buffer, pH 7.9, containing 0.5 M NaCl, and ruptured by two passages through a French pressure cell (144.9 MPa; 21 000 lb/in²). The polypeptides were purified from the clarified cell extract by immobilized-metal-affinity chromatography in accordance with the manufacturer's protocols (Novagen). Purified polypeptides were concentrated and exchanged into distilled water as described previously [9]. Purity, assessed by SDS/PAGE, was greater than 95%. A pET-*CBM17* clone was also made; however, exceedingly poor protein yields precluded further characterization of this module in isolation.

Determination of protein concentration

The concentration of purified protein was determined by measurement of A_{280} and using calculated molar absorption coefficients [10] of 32 290 M⁻¹·cm⁻¹ and 70 900 M⁻¹·cm⁻¹ for *CBM28* and *CBM17/CBM28* respectively.

UV difference scans

UV scans were collected with a Varian Cary 100e UV/visible 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. *CBM28* (25 μ M, 1 ml) or *CBM17/CBM28* (12.5 μ M, 1 ml) was scanned and corrected for a buffer-only baseline. A saturating amount of ligand (40-fold molar excess) was added; the sample was left to equilibrate and then rescanned. The latter scan was corrected for dilution and a difference scan was calculated by subtracting the protein-only scan from the protein-ligand scan.

Quantitative UV analysis of substrate binding

Absorbance spectra were collected on a S2000 charge-coupled-device-array fibre-optic spectrometer with a MINI-D2 deuterium UV light source and water-thermostatted cuvette holder (Ocean

Optics, Dunedin, FL, U.S.A.). The temperature was controlled by a recirculating heating/cooling bath (Forma Scientific, Marietta, OH, U.S.A.) and the temperature of the sample was monitored in the cuvette with a 0.8 mm thermistor temperature probe (Cole-Parmer, Vernon Hills, IL, U.S.A.) 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, U.S.A.). Data collection and pump control were automated and synchronized by software developed in-house.

Titration were performed in a 2 ml volume and buffered in 25 mM Tris/HCl, pH 7.0. Absorbance data from titration experiments with cello-oligosaccharides were processed as described previously [9]. The trough-to-peak heights at $\Delta A_{294.6} - \Delta A_{289.8}$ nm were plotted against total ligand concentration. The association constant (K_a) and the maximum molar change in the UV difference signal at saturation (ΔA_m) were determined by assuming a 1:1 stoichiometry and fitting the data to a bimolecular interaction that took ligand depletion into account. For binding experiments with barley β -glucan the data were converted to a molar change in UV difference/mol of glucan polymer injected, effectively the derivative data, and plotted against the molar ratio of glucan polymer to *CBM28* or *CBM17/CBM28*. These data were analysed with MicroCal MCS Origin (MicroCal, Northampton, MA, U.S.A.) by using a bimolecular interaction model that treated the barley β -glucan as the macromolecule and the *CBM* as the ligand.

Adsorption on insoluble cellulose was assayed by depletion isotherms as described previously [6,11,12], with UV absorbance to determine protein concentrations.

Isothermal titration calorimetry (ITC)

ITC was performed essentially as described previously [7] with an MCS ITC calorimeter (MicroCal). All samples were at pH 7.0 in 50 mM potassium phosphate buffer. Titrations were performed by injecting 10 μ l aliquots of cello-oligosaccharide solution into the ITC sample cell (volume 1.3528 ml) containing 202 μ M *CBM28*. All ITC data were corrected for the heat of dilution of the titrant by subtracting the excess heats at high molar ratios of cello-oligosaccharide to *CBM28*. The binding stoichiometry, enthalpy and equilibrium association constant were determined by fitting the corrected data to a 1:1 bimolecular interaction model. Two or three independent titration experiments were performed. Additional sets of titrations were performed at 15, 20 and 35 °C to measure the heat capacity of binding for cellopentaose and cellohexaose.

RESULTS

UV analysis of soluble ligand binding

Analysis of tyrosine exposure [13] in cellohexaose-saturated *CBM28* and *CBM17/CBM28* indicated that there was no change in tyrosine exposure on cellohexaose binding (results not shown). The UV difference signals of *CBM28* and *CBM17/CBM28* with saturating amounts of cellohexaose were virtually identical except for a larger signal from *CBM17/CBM28*, and composed entirely of features arising from tryptophan residues (Figure 1). Neither manno-oligosaccharides nor xylo-oligosaccharides generated a UV difference signal when added in excess to *CBM28*, indicating that these sugars are not ligands for *CBM28*. The 294.6–289.8 nm peak–trough pair was chosen as a quantitative monitor of binding to cello-oligosaccharides and barley β -glucan (Figure 2).

Figure 1 UV difference spectra of CBM28 and CBM17/CBM28 complexed with cellohexaose

(A) Difference spectrum of *N*-acetyltryptophan (solid line) and *N*-acetyltyrosine (broken line) perturbed in 20% (v/v) DMSO. Samples containing 50 μ M *N*-acetyltryptophan or 100 μ M *N*-acetyltyrosine in 25 mM Tris/HCl, pH 7.5, were used to collect baseline spectra. Samples of the same concentrations but containing 20% (v/v) DMSO were used to collect difference spectra. Peak and trough wavelengths are labelled in bold and italics for *N*-acetyltryptophan and *N*-acetyltyrosine respectively. (B) Difference spectrum of CBM28 (solid line) and CBM17/CBM28 (broken line) induced by cellohexaose binding. Peak and trough wavelengths are labelled.

Figure 2 Quantitative UV monitoring of ligand binding to CBM28

The left axis shows unmodified cumulative UV difference data of cellohexaose titrated into CBM28 (○). The right axis shows UV difference data as derivative data of barley β -glucan titrated into CBM28 (●). Titrations were performed as described in the Materials and methods section.

Table 1 Binding parameters of CBM28 determined by UV difference titrations at 25 °C

Samples were buffered in 25 mM Tris/HCl, pH 7.0. Results are means $\pm$ S.D. for triplicate experiments.

Glucan	$10^4 K_a$ (M^{-1})	ΔG (kJ/mol)	ΔA_m (arbitrary units)
Cellohexaose	4.96 ± 0.55	-26.79 ± 0.28	7.1 ± 0.4
Cellopentaose	1.84 ± 0.43	-24.29 ± 0.61	7.3 ± 0.4
Cellotetraose	0.51 ± 0.14	-21.08 ± 0.71	8.4 ± 0.5

Figure 3 Representative calorimetric titration

Cellohexaose (6.32 mM) was titrated into CBM28 (202 μ M) at 25 °C in 50 mM potassium phosphate, pH 7.0.

Measurement of cello-oligosaccharide binding by UV difference titrations

The affinity of CBM28 for soluble glucans increased with increasing chain length (Table 1), which is typical of CBMs that bind to soluble glycans [9,14–17]. No binding to cellotriase could be detected. Binding to the individual CBM modules in CBM17/CBM28 was quantified by fitting a two-site model to the data and setting the molar absorbance signal of CBM28, determined from an independent titration of CBM28 with cellohexaose, and the stoichiometries as constants. The association constants for the individual modules in the CBM17/CBM28 tandem for cellohexaose were $(4.01 \pm 0.60) \times 10^4 M^{-1}$ and $(3.03 \pm 1.2) \times 10^4 M^{-1}$ for the CBM17 and CBM28 modules respectively (means $\pm$ S.D.).

Polymeric β -glucan binding

The affinity of CBM28 for barley β -glucan, a mixed β -(1,3)(1,4)-glucan, was $(4.0 \pm 1.0) \times 10^4 M^{-1}$ with a stoichiometry of 12 ± 1.5 CBM28 modules per glucan molecule. This approximates to one CBM28 module occupying a span of 23 glucose residues in the glucan polymer. The association of CBM28 with barley β -glucan was remarkably insensitive to pH between 7.0 and 10.9 with the association constants not varying outside the experimental error (results not shown). The retention of affinity at elevated pH was consistent with the presence of this module in an enzyme from an

Table 2 Binding parameters of CBM28 determined by isothermal titration calorimetry at 25 °C

Samples were buffered in 50 mM potassium phosphate, pH 7.0. Results are means $\pm$ S.D. for duplicate experiments. n is the stoichiometry of the interaction. Abbreviation: n.d., not determined.

Glucan	$10^4 K_a$ (M $^{-1}$)	ΔG (kJ/mol)	ΔH (kJ/mol)	$T\Delta S$ (kJ/mol)	ΔC_p (J $\cdot$ mole $^{-1}$ $\cdot$ K $^{-1}$)	n
Cellohexaose	4.04 ± 0.05	-26.29 ± 0.03	-61.24 ± 0.50	-34.95 ± 0.60	-563 ± 30	1.08 ± 0.01
Cellopentaose	2.23 ± 0.01	-24.82 ± 0.02	-68.35 ± 0.18	-43.53 ± 0.19	-135 ± 15	1.05 ± 0.00
Cellotetraose	1.49 ± 0.04	-21.43 ± 0.06	-62.32 ± 1.83	-40.90 ± 1.89	n.d.	1.04 ± 0.03

alkalophilic organism. CBM17/CBM28 had an affinity of $(2.6 \pm 0.3) \times 10^5$ M $^{-1}$ for barley β -glucan and a stoichiometry of 6.7 ± 0.2 CBM17/CBM28 polypeptides per glucan molecule. This approximates to one CBM17/CBM28 polypeptide occupying a span of 40 glucose residues or 20 glucose residues per module in the glucan polymer.

CBM28 did not bind to bacterial microcrystalline cellulose, a highly crystalline preparation of cellulose. However, it did bind to the regenerated cellulose with a K_a of $(0.26 \pm 0.03) \times 10^6$ M $^{-1}$ and a binding capacity of 5.45 ± 0.28 μ mol/g of cellulose.

Thermodynamic analysis of cello-oligosaccharide binding

Calorimetric data could be fitted with a bimolecular binding model for all cello-oligosaccharides tested (Figure 3). Dominating favourable changes in enthalpy (ΔH) were partly compensated for by unfavourable changes in the entropy term ($T\Delta S$) (Table 2). The K_a values determined by ITC for cello-oligosaccharides agreed well with those determined by UV difference titrations. An enthalpy–entropy compensation plot (ΔH against $T\Delta S$) including the data for cellopentaose and cellohexaose at four different temperatures gave a straight line ($R^2 = 0.99$) with a slope of 0.83 ± 0.03 (results not shown). The results for cellotetraose deviated from this relationship and were not included in the regression analysis. As with UV difference titrations, no binding to cellotriose could be detected by ITC.

DISCUSSION

New family of CBMs structurally and functionally related to CBM17 modules

We have shown that CBM28, the X11 module from *Bacillus* sp. 1139 Cel5A, is indeed a glucan-binding CBM. It has significant amino acid sequence identity with six other modules that are primarily from alkalophilic bacilli (Figure 4). Members of CBM family 17 are related to these modules at the amino acid sequence level with approximately 15–20% sequence identity and 30–35% sequence similarity. The amino acid similarity of CBM17 modules and CBM28 modules (before this classification) has been noted previously [18]. The secondary structures of CBM17 and CBM28 modules are predicted to be well conserved; however, functional residues are only partly conserved (Figure 4). Although this difference seems to be subtle, by definition the CBM family classifications should be predictive of both structural and functional elements. Despite the probability of similar structures, the low level of identity and the lack of conserved functional residues are sufficient to classify them as separate families. Thus these modules have been classified into CBM family 28.

Many CBMs occurring in tandem show a relatively high degree of identity and probably arose through the duplication of gene fragments. In general, these CBMs show very similar binding properties [19,20]. The co-occurrence of CBM17 and CBM28 modules and their distant similarity also suggests that CBM17/CBM28 tandems occurred by a duplication event.

The specificity and affinity of CBM28 for non-crystalline cellulose and soluble β -1,4-glucans make it seem almost identical in functional properties to CBM17 from *Clostridium cellulovorans* Cel5A [9,21]. On the assumption that all CBM17 modules share similar binding properties, an assumption supported by the cellohexaose-binding affinity of the CBM17 module in CBM17/CBM28, from an evolutionary perspective the sequence divergence between CBM17 and CBM28 modules seems unnecessary given that identical repeats would perform the same function. This suggests subtle functional differences between the CBM17 and CBM28 modules that are as yet unidentified.

Mechanism of soluble glucan recognition by CBM28

UV studies of CBM28 with bound cello-oligosaccharides showed that the UV difference spectra resulted entirely from tryptophan residues. This, combined with the lack of change in tyrosine exposure as measured by second-derivative UV spectroscopy, suggested a binding site with tryptophan residues, but not tyrosine residues, that interact with ligand. The maximum molar UV difference signal did not change with ligand length, indicating that the full complement of tryptophan residues in the binding site was involved in interacting with cellotetraose (Table 1). The 1:1 stoichiometry of CBM28–cellotetraose binding (Table 2) shows that this observation could not have been due to multiple cellotetraose molecules occupying an extended binding site.

The thermodynamic data for cellohexaose and cellopentaose showed partial enthalpy–entropy compensation and suggested that these oligosaccharides bind by the same general mechanism. Cellotetraose did not follow this relationship. The slope of the compensation plot, which was less than unity, indicated a binding process that was most affected by changes in entropy. This, combined with negative ΔC_p values, suggested that binding was highly influenced by the burial of non-polar surface area and the concomitant reorganization of solvent [22,23].

Examination of the thermodynamic data yields some insights into the possible subsite architecture of the CBM28 binding site. Because the binding of CBM28 to cellotriose could not be detected by UV or ITC, cellotetraose seems to be the minimum ligand for CBM28 and provides the ‘base’ energy for the interaction in what will be considered the first subsite (S_1). The additional sugar of cellopentaose results in $\Delta\Delta G$, $\Delta\Delta H$ and $T\Delta\Delta S$ values of -3.4 , -6.0 and -2.6 kJ/mole respectively. Thus the binding of the fifth glucose moiety in the second subsite (S_2) contributes enthalpically, suggesting that this subsite gives mainly polar interactions. Cellohexaose, relative to cellopentaose, gives $\Delta\Delta G$, $\Delta\Delta H$, $T\Delta\Delta S$ and $\Delta\Delta C_p$ values of -1.5 kJ/mol, 7.1 kJ/mol, 8.6 kJ/mol and -430 J $\cdot$ mole $^{-1}$ $\cdot$ K $^{-1}$ respectively. The favourable entropic contribution, unfavourable enthalpic contribution and negative contribution to the ΔC_p of binding at the third subsite (S_3) suggest that S_3 is a non-polar subsite. The difference spectra of cellopentaose and cellohexaose were the same, which suggests that the non-polar amino acid side chains in the S_3 site are not sensitive to UV (for example leucine and

Figure 4 Amino acid sequence alignment of CBM17 modules and X11 (or CBM28) modules

Abbreviations are as follows: X11_1139, CBM28 from *Bacillus* sp. 1139 Cel5A (GenBank D00066); X11_C.j., from *Clostridium josui* Cel5A (GenBank D85526); X11_KSM-64, CBM28 from *Bacillus* sp., strain KSM-64, Cel5A (GenBank M84963); X11_22-28, from *Bacillus* sp. 22-28 Cel5A (GenBank D85236); X11_KSM-635, from *Bacillus* sp., strain KSM-635, Cel5A (GenBank M27420); X11_N-4, from *Bacillus* sp., strain N-4, Cel5A (GenBank M25500); X11_A.t., from *Anaerocellum thermophilum*, Cel5B (GenBank Z77855). The sequences of the CBM17 modules correspond to those described previously [9]. The predicted secondary (2°) structures [26] of CBM28 and the family 17 CBM from *Clostridium cellulovorans* (CBM17_C.c.) are shown beneath the X11 and CBM17 sequences respectively. The actual secondary structure of CBM17_C.c. [21] (PDB ID 1J84) determined from its three-dimensional structure is also shown. Aromatic amino acids involved in the binding of CBM17_C.c. to cello-tetraose are indicated above the sequences by ♦. Polar amino acids involved in binding are indicated by ●. This alignment was prepared with ClustalW [27]. Asterisks mark every 10 amino acid residues.

isoleucine) or to changes in environment on binding (for example phenylalanine).

The sensitivity to entropic effects of the binding of cello-pentaose compared with that of cellohexaose is apparently at odds with the enthalpically dominated free energies of binding. Favourable changes in enthalpy are very common in CBM-carbohydrate interactions [7,14,15,21] and result primarily from polar interactions, such as hydrogen bonding, and van der Waals interactions. Unfavourable changes in entropy are also characteristic of carbohydrate-binding events and are thought to result mainly from the freezing of sugar and amino acid side chain conformations [14,24]. This probably describes the overall properties of a CBM28-cello-oligosaccharide interaction. An under-

standing of the sensitivity of a CBM28-cello-oligosaccharide interaction to entropic effects is provided by the subsite architecture of CBM28. The sum of the contributions of S_2 and S_3 to the binding energetics gives a favourable net entropic contribution (+6 kJ/mol) but unfavourable net enthalpic contribution (+1.1 kJ/mol), showing that, overall, the swings in entropy dominate when considering binding to cello-tetraose with cello-pentaose and with cellohexaose. On the basis of the thermodynamic values, the inferred polarities of the interacting surfaces in subsites S_2 and S_3 are quite different. The sensitivity of the binding to entropy probably comes from the differential entropic contributions of solvent reorganization owing to the varied polarities of the subsites and from the manner in which this compensates for the entropic penalty of fixing the conformations of both ligand (carbohydrate) and macromolecule (CBM28).

Tandem CBM17 and CBM28 modules

Glycoside hydrolases frequently contain multiple CBMs. Occasionally these modules co-operate to give an enhanced affinity for polysaccharides [19]. However, equally often, multiple CBMs seem to provide no advantage in terms of binding affinity [14,25]. The related specificity of CBM28 and CBM17 modules suggested that a function of the CBM17/CBM28 tandem might be to enhance the affinity for β -glucans. Indeed, the tandem did have an approx. 10-fold higher affinity for the mixed β -(1,3)(1,4)-glucan, barley β -glucan, than for cellohexaose. On the assumption that the CBM17 module shows the same affinity for barley β -glucan as it did for cellohexaose (in other words, the same pattern as CBM28), the enhanced affinity probably resulted from the co-operative binding of the modules within the tandem to proximal binding sites on the same glucan molecule. Binding of the modules within the tandem to separate glucan molecules would not result in an increase in affinity, as is evident from the results of cellohexaose binding to CBM17/CBM28. However, the possibility cannot be excluded that the CBM17 module had a ten-fold higher affinity for barley β -glucan than for cellohexaose.

We thank Patrick Chiu for technical assistance, and Dr Charles Haynes and Dr Louise Creagh for providing assistance with the calorimetry. This work was supported by the Natural Sciences and Engineering Research Council and the Protein Engineering Network Centres of Excellence.

REFERENCES

- Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. 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.
- 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) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochem. J.* **331**, 775–781.
- Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A. and Miller, Jr, R. C. (1993) The nature of the cellulose-binding domain affects the activities of a bacterial endoglucanase on different forms of cellulose. *FEMS Microbiol. Lett.* **113**, 211–218.
- Meinke, A., Gilkes, N. R., Kilburn, D. G., Miller, R. C. J. and Warren, R. A. (1991) Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: functions and relatedness to domains in other polypeptides. *J. Bacteriol.* **173**, 7126–7135.
- Henrissat, B., Teeri, T. T. and Warren, R. A. (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Lett.* **425**, 352–354.
- Gilkes, N. R., Jervis, E., Henrissat, B., Tekant, B., Miller, R. C. J., Warren, R. A. and Kilburn, D. G. (1992) The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. *J. Biol. Chem.* **267**, 6743–6749.

- 7 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) Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. *Biochemistry* **40**, 6240–6247
- 8 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 9 Boraston, A. B., Chiu, P., Warren, R. A. J. and Kilburn, D. G. (2000) Specificity and affinity of substrate binding by a family 17 carbohydrate-binding module from *Clostridium cellulovorans* cellulase 5A. *Biochemistry* **39**, 11129–11136
- 10 Mach, H., Middaugh, C. R. and Lewis, R. V. (1992) Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. *Anal. Biochem.* **200**, 74–80
- 11 McLean, B. W., Bray, M. R., Boraston, A. B., Gilkes, N. R., Haynes, C. A. and Kilburn, D. G. (2000) Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. *Protein Eng.* **13**, 801–809
- 12 Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G. and Haynes, C. A. (1996) Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12229–12234
- 13 Ragone, R., Colonna, G., Balestrieri, C., Servillo, L. and Irace, G. (1984) Determination of tyrosine exposure in proteins by second-derivative spectroscopy. *Biochemistry* **23**, 1871–1875
- 14 Tomme, P., Creagh, A. L., Kilburn, D. G. and Haynes, C. A. (1996) Interaction of polysaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 1. Binding specificity and calorimetric analysis. *Biochemistry* **35**, 13885–13894
- 15 Charnock, S. J., Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M., Davies, G. J. and Fontes, C. M. (2000) The X6 'thermostabilizing' domains of xylanases are carbohydrate-binding modules: structure and biochemistry of the *Clostridium thermocellum* X6b domain. *Biochemistry* **39**, 5013–5021
- 16 Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J. and Williamson, M. P. (1999) A family 11b xylan-binding domain has a similar secondary structure to a homologous family 11a cellulose-binding domain but different ligand specificity. *Structure Fold. Des.* **7**, 853–864
- 17 Meissner, K., Wassenberg, D. and Liebl, W. (2000) The thermostabilizing domain of the modular xylanase XynA of *Thermotoga maritima* represents a novel type of binding domain with affinity for soluble xylan and mixed-linkage β -1,3/ β -1,4-glucan. *Mol. Microbiol.* **36**, 898–912
- 18 Sunna, A., Gibbs, M. D. and Bergquist, P. L. (2001) Identification of novel β -mannan- and β -glucan-binding modules: evidence for a superfamily of carbohydrate-binding modules. *Biochem. J.* **356**, 791–798
- 19 Bolam, D. N., Xie, H., White, P., Simpson, P. J., Hancock, S. M., Williamson, M. P. and Gilbert, H. J. (2001) Evidence for synergy between family 2b carbohydrate binding modules in *Cellulomonas fimi* xylanase 11A. *Biochemistry* **40**, 2468–2477
- 20 Brun, E., Johnson, P. E., Creagh, A. L., Tomme, P., Webster, P., Haynes, C. A. and McIntosh, L. P. (2000) Structure and binding specificity of the second N-terminal cellulose-binding domain from *Cellulomonas fimi* endoglucanase C. *Biochemistry* **39**, 2445–2458
- 21 Notenboom, V., Boraston, A. B., Chiu, P., Frelove, A., Kilburn, D. G. and Rose, D. R. (2001) Recognition of cello-oligosaccharides by a family 17 carbohydrate-binding module: an X-ray crystallographic, thermodynamic and mutagenic study. *J. Mol. Biol.*, in the press
- 22 Sturtevant, J. M. (1977) Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2236–2240
- 23 Murphy, K. P. and Gill, S. J. (1990) Group additivity thermodynamics for dissolution of solid cyclic dipeptides in water. *Thermochim. Acta* **172**, 11–20
- 24 Chervenak, M. C. and Toone, E. J. (1995) Calorimetric analysis of the binding of lectins with overlapping carbohydrate-binding ligand specificities. *Biochemistry* **34**, 5685–5695
- 25 Gill, J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P. and Gilbert, H. J. (1999) The type II and X cellulose-binding domains of *Pseudomonas xylanase* A potentiate catalytic activity against complex substrates by a common mechanism. *Biochem. J.* **342**, 473–480
- 26 Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. and Barton, G. J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892–893
- 27 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix. *Nucleic Acids Res.* **22**, 4673–4680

Received 17 August 2001/3 October 2001; accepted 23 October 2001