Carbohydrate-binding modules from a thermostable *Rhodothermus marinus* xylanase: cloning, expression and binding studies

Maher ABOU HACHEM^{*1}, Eva NORDBERG KARLSSON^{*}, Eva BARTONEK-ROXÅ^{*}, Srinivasrao RAGHOTHAMA[†], Peter J. SIMPSON[‡], Harry J. GILBERT[‡], Michael P. WILLIAMSON[‡] and Olle HOLST^{*}

^{*}Department of Biotechnology, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden,

[†]Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, U.K., and

[‡]Department of Biological and Nutritional Sciences, University of Newcastle, Newcastle upon Tyne NE1 7RU, U.K.

The two *N*-terminally repeated carbohydrate-binding modules (CBM4-1 and CBM4-2) encoded by *xyn10A* from *Rhodothermus marinus* were produced in *Escherichia coli* and purified by affinity chromatography. Binding assays to insoluble polysaccharides showed binding to insoluble xylan and to phosphoric-acid-swollen cellulose but not to Avicel or crystalline cellulose. Binding to insoluble substrates was significantly enhanced by the presence of Na⁺ and Ca²⁺ ions. The binding affinities for soluble polysaccharides were tested by affinity electrophoresis; strong binding occurred with different xyans and β -glucan. CBM4-2 displayed a somewhat higher binding affinity than CBM4-1 for both

soluble and insoluble substrates but both had similar specificities. Binding to short oligosaccharides was measured by NMR; both modules bound with similar affinities. The binding of the modules was shown to be dominated by enthalpic forces. The binding modules did not contribute with any significant synergistic effects on xylan hydrolysis when incubated with a Xyn10A catalytic module. This is the first report of family 4 CBMs with affinity for both insoluble xylan and amorphous cellulose.

Key words: CBD, hemicellulose binding, modular proteins, Xyn10A.

INTRODUCTION

Xylanases and other glycosyl hydrolases have a modular structure [1–3]. Typically these enzymes are composed of catalytic modules joined to one or more auxiliary domains by recognizable linker sequences [2,4]. The most common auxiliary domains in hemicellulases (such as xylanases, mannanases, xylosidases and acetyl-esterases) and cellulases are the carbohydrate-binding modules (CBMs) [3,5], most of which bind preferentially to cellulose. These modules were previously referred to as cellulose-binding domains. A nomenclature that has been proposed recently (A. Boraston, personal communication) will be used here.

The CBMs found so far can be classified into different families (today more than ten), often based only on a comparison of primary structure with previously characterized sequences [3]. Many of the modules in this classification system are, however, not functionally characterized, and their precise roles in hemicellulase action are not yet fully understood [4,6–8].

It is now well documented that CBMs can enhance the hydrolysis of crystalline cellulose and complex cellulosic substrates [9–13]. Two proposals have been made to explain this enhancement [5,8]: the first suggests an increase in the effective enzyme concentration on the substrate surface that is caused by the attachment of the CBM [14]; the second is that the CBM has a role in disrupting non-covalent associations in the substrate, thereby increasing substrate accessibility [15]. Evidence so far suggests that, whereas all CBMs enhance hydrolytic rates in accordance with the first proposal, not all act in accordance with the second.

Hemicellulase catalytic modules are also connected to CBMs [16–18], and it has been shown that the hemicellulase activity on complex substrates (such as cellulose/hemicellulose complexes) is enhanced by the presence of a CBM [4,19]. The role of the binding modules with affinity for cellulose could be to participate

in the dissociation of cell-wall matrix polysaccharides such as xyans that are in close contact with microfibrils, or to mediate an intimate contact between the enzyme and the plant cell wall [4,19].

In most cases CBMs have been observed to bind to cellulose in both the crystalline and amorphous forms. A few, more specific, modules have been characterized, such as the two *Cellulomonas fimi* CBMs encoded by *xynD*, which bind preferentially to insoluble xylan [6,20], and *C. fimi* endoglucanase C (CenC) [21] and *Clostridium stercorarium* XynA [22], both with affinity only for amorphous cellulose.

The objective of the present study was to investigate the substrate-binding affinities of the isolated CBMs encoded by *xyn10A* (renamed in accordance with the nomenclature proposed by Henrissat et al. [23]) from *Rhodothermus marinus*, which is an aerobic thermophilic bacterium first isolated from marine hot springs on the west coast of Iceland [24]. The thermostable multidomain xylanase encoded by *xyn10A* possesses two putative CBMs suggested to belong to family 4 (previously identified with roman numerals) [25]. Using CBMs separately produced in *Escherichia coli*, we show that both modules have an affinity for insoluble xylan and amorphous cellulose, thus exhibiting a binding affinity different from that of the previously described modules of this family. Despite high similarities in sequence, the two modules exhibit differences in binding properties.

MATERIALS AND METHODS

Plasmids and strains

The gene fragments encoding the 165-residue CBM4-1 and the 166-residue CBM4-2 from Xyn10A [25] were cloned in the expression vector pET-25b(+) (Novagen, Madison, WI, U.S.A.).

The recombinant plasmid derived from pET-25b(+) was

Abbreviations used: AE, affinity electrophoresis; C₆, cellohexaose; CBM, carbohydrate-binding module; CenC, *Cellulomonas fimi* endoglucanase C; CMC, carboxymethylcellulose; PASC, phosphoric-acid-swollen cellulose; X₆, xylohexaose.

¹ To whom correspondence should be addressed (e-mail maher.abou-hachem@biotek.lu.se).

designed to produce each CBM construct fused with a hexahistidine (His) tag at the C-terminus. Appropriate restriction sites (underlined) were introduced by PCR at the 5' and 3' ends by using genomic *R. marinus* DNA as template and the Expand[®] high-fidelity PCR system (Boehringer Mannheim, Mannheim, Germany). Forward and reverse primers with the following sequences were used for the amplification of CBM4-1 and CBM4-2 respectively: forward CBM4-1, 5'-ACGTGATACATATGACGCCGGCAAATGTCAATGGAAG-3'; forward CBM4-2, 5'-ACGTCATACATATGCTTGTGCGCAACATCAACGGTGG-3'; reverse CBM4-1, 5'-TACGACCTGCTAGCCACGAGGGCATCGATATAAATGG-3'; reverse CBM4-2, 5'-TACGATAAGCTAGCAATGGCCAGGCCATCAATGTAGATGG-3'.

A hot-start PCR protocol of 35 cycles of denaturation at 95 °C for 1 min, annealing at 60 °C for 1 min and primer extension at 72 °C for 1 min was followed. Minor mutations in terms of the insertion or deletion of a maximum of two residues at the termini of the CBMs were introduced along with the PCR amplification. The His-tag is terminated by a stop codon and is preceded by an hsv-tag. The hsv-tag is comprised of 11 amino acid residues, derived from herpes simplex virus glycoprotein D. The resulting 511 and 517 bp PCR fragments, encoding CBM4-1 and CBM4-2 respectively, were separated on an agarose gel and subsequently purified by using the QIAEX protocol from Qiagen (Diagen GmbH, Hilden, Germany), digested with *Nde*I and *Nhe*I in two steps and cloned by using the *Nde*I-*Nhe*I restriction sites in the polylinker of pET-25b(+). The vector was transformed into *E. coli* strain XL1-Blue (Stratagene, San Diego, CA, U.S.A.) by electroporation with a Gene Pulser II electroporation system (Bio-Rad, Hercules, CA, U.S.A.) by using a standard protocol. Presence of the CBM sequences in correct reading frames was verified by restriction digestion with *Nhe*I and by partial sequencing by the dideoxy chain-termination method with the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer, Foster City, CA, U.S.A.). Plasmids were prepared from *E. coli* with Qiagen's mini plasmid kit.

Production

The plasmids encoding CBM4-1 and CBM4-2 were purified and transformed into *E. coli* production strain BL21(DE3). Shake flask cultures were grown on Luria-Bertani medium at 37 °C to an attenuation (D_{600}) of 0.75, expression was induced by the addition of isopropyl β -D-thiogalactoside to a final concentration of 0.5 mM, and growth continued for 3 h. The cells were harvested by centrifugation at 6000 *g* for 5 min. The cell pellets were washed with 0.9% NaCl, centrifuged as above and resuspended in 20 mM Tris/HCl (pH 7.4)/20 mM imidazole/0.75 M NaCl (binding buffer used in subsequent purification). The cells were disintegrated by ultrasonication at 60 W/cm² (three treatments of 120 s, cycle 0.5) with a UP400S sonicator (Dr. Hielscher GmbH, Stahnsdorf, Germany) and centrifuged at 27000 *g* for 10 min; the supernatants were frozen and stored for later purification.

Purification

Immobilized metal-ion-affinity chromatography was employed in the purification of the recombinant proteins. Iminodiacetic acid gel, prepared by covalently binding iminodiacetic acid to an epoxy-activated Sepharose CL-6B matrix, was mixed with 5 vol. of a CuSO₄ · 5H₂O solution (5 mg/ml) for 30 min and packed in

a column. The column was washed consecutively with at least 5 vol. of water, 50 mM imidazole and water, and subsequently equilibrated with binding buffer. Supernatants of the cell homogenates obtained above were loaded on the column; unbound protein was washed with several volumes of binding buffer until there was no further decrease in A_{280} . The bound protein was eluted with 20 mM Tris/HCl (pH 7.4)/250 mM imidazole/0.75 M NaCl and subsequently dialysed extensively against 20 mM Tris/HCl, pH 7.5. Crude cell homogenate and purified samples, along with Kaleidoscope prestained standard (Bio-Rad), were analysed by SDS/PAGE [12.5% (w/v) gel] by the method of Laemmli [26]. Recombinant protein production was estimated by densitometry.

Portions of the two purified polypeptides were treated with carboxypeptidase-A type II from bovine pancreas (Sigma, St. Louis, MO, U.S.A.) for 1 h at 37 °C. After the treatment with protease, the solutions were loaded on the affinity column and portions that failed to adsorb were assumed to lack the His-tag and possibly residues from the hsv-tag. This was verified by SDS/PAGE analysis.

Thermal denaturation

The thermal denaturation of each CBM was performed by using an HP 8453 UV diode-array spectrophotometer with a heating cell (Hewlett-Packard, Waldbronn, Germany). The spectrum was monitored between 190 and 1100 nm. The cuvettes were rinsed with deionized water before the addition of 2.5 ml of protein solutions containing CBM4-1 or CBM4-2 (0.1 mg/ml) in 20 mM Tris/HCl, pH 7.5. The temperature was then subsequently raised in 0.5 °C increments and held for 0.5 min to allow thermal equilibrium before collection of the spectra. The temperature interval was 50–100 °C; protein denaturation was illustrated by plotting the absorbance difference $A_{286} - A_{264}$ against temperature [27].

Adsorption and enzyme assays

Insoluble xylan was prepared by washing oat-spelt xylan (Sigma) twice, the first wash being overnight. Avicel was purchased from Merck (Darmstadt, Germany). Phosphoric-acid-swollen cellulose (PASC) was prepared by the treatment of Avicel with phosphoric acid, as described by Wood [28]. Cellulose of higher crystallinity was prepared from cotton [28], and bacterial microcrystalline cellulose was prepared as described previously [14].

Binding studies involving these insoluble substrates were conducted at room temperature in 10 ml glass tubes; 5 ml of approx. 13 μ M CBM4-2 (0.26 mg/ml) in 10 mM Tris/HCl, pH 7.5, was added to 50 mg of PASC, 50 mg of Avicel, 50 mg of xylan, 10 mg of dewaxed cotton or 0.5 mg of bacterial microcrystalline cellulose. Each solution was vortex-mixed and the samples were incubated on a head-to-head shaker for 3 h to allow the system to equilibrate. The samples were then centrifuged for 5 min at 1500 *g*; free protein in the supernatants was determined with a bicinchoninic acid protein assay kit (Sigma). BSA (0.26 mg/ml) was employed as a negative control; the measurements were done in at least triplicate. The ratio (as a percentage) between the initial and the free protein concentration after incubation with the different substrates was defined as the residual protein concentration. Total protein was determined with the bicinchoninic acid kit.

The same procedure as that described above was used to investigate the influence of Na⁺ and Ca²⁺ ions on the binding to insoluble xylan and PASC respectively. Protein controls containing Ca²⁺ or Na⁺ ions but lacking the insoluble substrates were made to account for eventual precipitation due to poly-

peptide aggregation. The adsorption of CBM4-2 lacking the His-tag was tested to assess whether the presence of the tag had any significant impact on the binding.

The adsorption of a concentration range of CBM4-1 added on a fixed amount of insoluble xylan and PASC in the presence of 5 mM CaCl_2 was described by binding isotherms determined by using the same procedure as above. Free protein concentration in the supernatant ($[P]$, μM) was determined by A_{280} . The absorption coefficient (1 cm, 1 mg/ml, 280 nm), determined by far-UV absorbance [29], was 1.37. The bound polypeptide concentration ($[PX]$, $\mu\text{mol/g}$ of xylan or PASC) was calculated from the concentration of added polypeptide and $[P]$. A coefficient (1/g of substrate), defined as the initial slope of the binding isotherm, was determined as a measure of the relative affinities for the tested substrates.

To investigate possible synergistic effects of the CBMs on insoluble xylan hydrolysis, the isolated catalytic module of Xyn10A, designated Xyn10A Δ NC [30] was used. Xylanase enzyme assays on washed oat-spelt xylan and with pure Xyn10A Δ NC were performed as described previously [30]. The activity of Xyn10A Δ NC, Xyn10A Δ NC in the presence of equimolar concentrations of one or both CBMs, and Xyn10A Δ NC in the presence of a 10-fold excess of one or both CBMs was measured.

Affinity electrophoresis (AE)

The AE method described by Tomme [21] was employed with a Bio-Rad mini-gel apparatus. The 12.5% (w/v) native polyacrylamide gels, with and without substrate (0.05–0.1%), were polymerized side by side, separated by an internal spacer within the same glass plates. CBM4-1 and CBM4-2 (5–10 μg) were separated on gels with and without substrate at room temperature and 90 V. As a negative non-interacting control, a Kaleidoscope prestained standard was used. Proteins were detected by staining with Coomassie Brilliant Blue. Screening with AE included the following substrates: oat-spelt xylan, larchwood xylan, carboxymethylcellulose (CMC), β -glucan (barley), glucomannan (locust-bean gum), arabinogalactan, lichenan and laminarin (all purchased from Sigma), birch xylan (Birch 7500; Roth, Karlsruhe, Germany) and wheat arabinoxylan (Megazyme, Bray, Ireland).

NMR binding studies

Protein solutions were prepared in 10 mM Tris/HCl, pH 7.5, in $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ (9:1, v/v). Protein concentrations were estimated by comparison with an internal standard. NMR titrations were performed on a Bruker DRX-500 spectrometer equipped with z -gradients on a 5 mm probe, with the use of the Watergate sequence for decreasing the water signal [31]. The temperature was calibrated with a standard sample of ethylene glycol. Titrations were performed by adding aliquots of cellohexaose (C_6 ; Seikagaku Corp., Tokyo, Japan) or xylohexaose (X_6 ; Dextra Laboratories, Reading, Berks., U.K.) in the same buffer to the protein and measuring chemical shift changes compared with 3-trimethylsilyl 2,2,3,3- d_4 -propionate as an internal standard. All spectra were processed by using Felix 97.0 (Molecular Simulations). Chemical shift changes were fitted to a standard binding equation as described [14].

RESULTS

Production, purification and thermal denaturation

High production levels of the cloned CBMs were obtained; the recombinant proteins comprised between 40 and 50% of total cell protein as estimated by densitometry. The affinity-

Figure 1 SDS/PAGE of crude and purified Xyn10A CBMs

The gel was stained with Coomassie Brilliant Blue and the samples were loaded in the following order (from left to right): 1, protein molecular mass standard (molecular masses shown at the left); 2, CBM4-1 ultrasonicated crude cell extract; 3, CBM4-1 purified sample; 4, CBM4-2 ultrasonicated crude cell extract; 5, CBM4-2 purified sample.

Table 1 CBM4-2 binding to insoluble substrates at room temperature

CBM4-2 was incubated with Avicel, PASC or the insoluble fraction of oat-spelt xylan. A corresponding concentration of BSA was also incubated with each polymer as a negative non-interacting control. The results given (means $\pm$ S.D.) are the residual protein concentrations as percentages of the initial protein concentration.

Substrate	CBM4-2 binding (% of initial concentration)	
	BSA	CBM4-2
Avicel	96.4 $\pm$ 2.6	94.9 $\pm$ 3.9
PASC	95.2 $\pm$ 0.4	84.6 $\pm$ 1.8
Oat-spelt xylan	97.5 $\pm$ 3.6	78.9 $\pm$ 3.5

chromatography purification procedure yielded protein solutions containing as much as 85–90% of the recombinant polypeptides after one purification step (Figure 1).

The molecular masses were in agreement with the theoretical values but a band of higher molecular mass also appeared on denaturing PAGE of CBM4-1 but not CBM4-2. This extra band was not present in the crude untreated cell extracts (results not shown) and appeared only after treatment of the sample with ultrasonication or heat (70 $^{\circ}\text{C}$, 30 min; results not shown).

The estimated denaturation temperatures of CBM4-1 and CBM4-2 were 73.8 $^{\circ}\text{C}$ and 87.4 $^{\circ}\text{C}$ respectively.

Adsorption and enzyme assays

The qualitative binding experiments conducted show that CBM4-2 binds to PASC and insoluble xylan from oat spelt but not to crystalline cellulose (results not shown) or Avicel (Table 1). Avicel consists of approx. 50% crystalline cellulose and 50% amorphous cellulose. These results imply binding to both xylan and PASC. The same pattern of binding specificity was observed with CBM4-1 (results not shown); in accordance with the affinity results obtained from tests with soluble substrates (AE, see below), CBM4-2 displayed a higher relative binding level than CBM4-1.

The relative adsorption of CBM4-2 to both insoluble xylan and PASC increased significantly in the presence of Ca^{2+} ions

Figure 2 Effect of Ca^{2+} ions on adsorption

The relative adsorption of CBM4-2 on insoluble xylan in the presence of different CaCl_2 concentrations is shown.

Figure 3 Binding isotherms of CBM4-1 to insoluble xylan and PASC

The binding isotherms to insoluble xylan (A) and amorphous cellulose (B) were measured in the presence of 5 mM CaCl_2 . [P] is given in μM and [PX] in $\mu\text{mol/g}$ of substrate. An exponential trendline has been added to both panels.

(Figure 2). The presence of Na^+ ions also enhanced adsorption to xylan and a maximum of 60% of added polypeptide was adsorbed in this case, although the relative adsorption increase was not as pronounced as when Ca^{2+} ions were added. At 5 mM CaCl_2 , adsorption increased sharply; a plateau was reached at 20 mM. Some aggregation was observed at high Ca^{2+} concentrations. However, reversible binding was confirmed by repeated washing of the xylan followed by a single addition of EDTA, after which more than 50% of the adsorbed polypeptide was eluted from the polysaccharides (results not shown).

The CBM4-2 samples lacking the His-tag did not display any significant deviation in binding behaviour to PASC and insoluble

Table 2 Binding of *xyn10A* CBMs to various soluble substrates

Semi-quantitative screening of the binding affinity of *xyn10A* CBMs to various high-molecular-mass carbohydrate substrates detected by affinity electrophoresis. Semi-quantitative measurement: —, no detectable affinity; +, clearly observable affinity; ++, strong specific interaction; + + +, very strong specific interaction.

Substrate	Binding affinity of CBMs	
	CBM4-1	CBM4-2
Birch xylan	++	+++
Wheat arabinoxylan	++	+++
Oat-spelt xylan	++	+++
Larchwood xylan	++	+++
Barley β -glucan	++	+++
CMC	—	—
Glucomannan	—	—
Laminarin	+	++
Lichenan	+	++
Arabinogalactan	—	—

xylan (results not shown), implying that the His-tag did not interfere with the binding.

The affinity coefficients for the adsorption of CBM4-1 on insoluble xylan and PASC were estimated to be 0.74 and 0.68 l/g respectively. The experimentally determined maximum amounts of bound CBM4-1, [PX], to xylan and PASC were 16.0 and 7.0 $\mu\text{mol/g}$ of substrate respectively (Figure 3). Saturation was approached in both cases.

No significant synergistic effects were measured from the xylanase activity assays performed on insoluble xylan, with the use of Xyn10A catalytic module, in the presence or absence of the CBMs.

AE

AE was employed in this work for testing the affinity of the cloned CBM4-1 and CBM4-2 for different soluble high-molecular-mass carbohydrate polymers (Table 2). Both CBMs showed highest affinity for the different xylns and β -glucan. Qualitatively, CBM4-1 and CBM4-2 displayed similar specificities for the different substrates tested. However, the general pattern observed implied that CBM4-2 had a higher binding affinity than CBM4-1 as seen by a greater retardation in AE gels (Figure 4).

NMR titrations

NMR titrations of CBM4-1 with X_6 and C_6 were fitted to a standard binding isotherm (Figure 5) and yielded the dissociation constants listed in Table 3. The binding affinity for X_6 is more than 15-fold that for C_6 . The titration curves with C_6 were in many cases poor fits to a simple single-site model, suggesting that C_6 might bind in several possible sites and/or orientations. Similar results were seen for CBM4-2. In titrations of CBM4-1 with X_6 performed at 65 °C, the NMR signals moved from the free to the bound chemical shifts with little or no broadening during the titration. This corresponds to fast exchange, suggesting an off-rate faster than 500 s^{-1} [32]. However, in the titrations performed with X_6 at 30 °C, many resonances broadened during

Figure 4 AE of CBM4-1 (N1) and CBM4-2 (N2)

Native PAGE was conducted in the absence (—) and the presence (+) of larchwood xylan (0.25%, w/v) (1), barley β -glucan (0.05%, w/v) (2), laminarin (0.1%, w/v) (3) or CMC (0.1%, w/v) (4). Kaleidoscope prestained standard (C) was employed as a negative non-interacting control.

the titration and then sharpened up again at high ligand concentration, indicating intermediate exchange on the chemical-shift timescale. Thus, at this temperature, the off-rates were approx. $10\text{--}100\text{ s}^{-1}$. The results therefore suggest that the off-rate is slower by an approximate factor of 10 or more at the lower temperature, which would be consistent with the typical temperature dependence of off-rates. In comparison with the dissociation constant, which is in the low-micromolar range, this suggests that the on-rate is only one or two orders of magnitude slower than the diffusion-controlled rate. The binding affinities to the oligohexaoses were not greatly affected by the presence of Ca^{2+} ions.

DISCUSSION

Binding specificity

In a previous study [30] we presented evidence for the specific binding of the full-length xylanase encoded by *xyn10A* to xylan, β -glucan and amorphous (but not crystalline) cellulose and suggested that the binding was mediated by the two *N*-terminal putative CBMs. In the present paper we prove that binding is indeed mediated by each of these modules independently. The results obtained show that the individual CBMs have an affinity for insoluble xylan and for amorphous cellulose but do not bind to the highly ordered flat surface of crystalline cellulose.

Of the soluble polysaccharides tested, the highest binding affinity was obtained with the four xylans and β -glucan (Table 2). Weak interaction was observed with lichenan and laminarin, whereas no binding was observed to CMC, arabinogalactan or glucomannan (Table 2 and Figure 4). The same behaviour was

observed with hexasaccharides, with only weak binding to C_6 (Table 3). These binding specificities can be rationalized on the basis of the structures of the different polysaccharides. Xylan is a polymer of β -1,4-linked xylose residues with various degrees of substitution and branching depending on its origin, hardwood xylans being less substituted than those of softwoods [33]. β -Glucan [(1,3),(1,4)- β -D-glucan] can be considered as mainly a non-substituted unbranched chain of β -1,3-linked cellotriosyl and cellotetraosyl units arranged randomly, containing approx. 70% β -1,4 and 30% β -1,3-links [34]. Lichenan has a structure similar to that of β -glucan, with a non-substituted unbranched chain of β -1,3 linked cellotriosyl units (according to Sigma) but here the β -1,3-linkages occur with a higher frequency. Laminarin has a backbone of β -1,3-linked glucose units with only a small quantity of β -1,6 linkages (according to Sigma). CMC is evenly substituted with carboxymethyl groups at three sites (C-2, C-3 and C-6) [35]. Arabinogalactans are composed of β -1,3-linked D-galactopyranose residues, each of which carries a side chain on its 6-position [33]. Glucomannan is an unbranched, non-substituted molecule with a backbone of β -1,4-linked mannose and glucose units, the former being in the majority [33].

From the above results it is therefore expected that the Xyn10A CBMs prefer xylan as a ligand but can accept unbranched linear β -1,3 linked polymers. β -1,4-linked hexoses (i.e. glucose and mannose) are bound less well. One possible explanation for the discrimination is that β -1,4-linked hexoses present a planar three-dimensional structure, with a 180° rotation between monomers, whereas xylan [36] and β -1,3-linked glucose units [37] both have a marked helical character, with an approx. 120° rotation between every second monomer. It is unclear whether this explanation is valid for insoluble complex substrates.

Figure 5 NMR titrations of CBM4-1 and CBM4-2

Oligosaccharide titrations were performed at 65 °C with C_6 with CBM4-1 (a), X_6 with CBM4-1 (b) and X_6 with CBM4-2 (c). The curves show the best fitted isotherms.

Table 3 Dissociation constants obtained with soluble oligosaccharides

Dissociation constants for X_6 and C_6 were determined by NMR.

Ligand	Temperature (°C)	Dissociation constant (μ M)	
		CBM4-1	CBM4-2
X_6	65	12 ± 1	10 ± 2
C_6	65	200 ± 50	360 ± 50
X_6	30	—	5 ± 2

Binding of Xyn10A CBMs to both insoluble xylan and PASC was markedly enhanced in the presence of Ca^{2+} and Na^+ ions, which suggests a strong dependence on the ionic strength of the binding solution. With insoluble substrates the binding results obtained in the presence of 5 mM $CaCl_2$ imply a better binding

capacity to xylan. There was, however, no major difference in affinities, estimated as affinity coefficients from the initial slope of the isotherms. Scatchard plots (adsorption data plotted in the form $[PX]/[P]$ against $[PX]$) were both non-linear (results not shown), suggesting a complex interaction between the polypeptide and the substrates tested.

At present we can only speculate on the reasons for the different affinities with soluble and insoluble substrates in the absence of Ca^{2+} ions, and insoluble substrates in the presence of these ions. A plausible explanation for this behaviour might be the decrease in ionic repulsion between the negatively charged polypeptides (theoretical pI values 4.39 and 4.31 for CBM4-1 and CBM4-2 respectively) and negatively charged xylan substituents such as 4-*O*-methylglucuronic acid. However, because the structure and substitution of xyans in general are not well defined, it is hard to assess the validity of this hypothesis. Moreover, the effect of ions on binding to PASC, which is non-ionic, is not resolved by the above. Another possible explanation might be that positive ions in general and multivalent ions in particular can bind to the polypeptide, altering the overall pI and neutralizing the negative charge or inducing conformational changes that result ultimately in more favourable binding. An ion-dependent adsorption of CBMs has been proposed previously [38,39].

The structure of the *C. fimi* xylanase D *N*-terminal xylan-binding CBM2b module is similar to that of the *C. fimi* exocellulase cellulose-binding CBM2a module [20], although the former binds xylan and the latter binds cellulose. The different ligand specificities of these two modules were shown to arise from the orientations of two surface-exposed tryptophan residues that form key interactions with the polysaccharides. In the xylan-binding module in xylanase D they are roughly perpendicular, forming a twisted binding site, whereas in the cellulose-binding module in *C. fimi* exocellulase they are coplanar. We are currently undertaking a structural study on CBM4-1 and CBM4-2 to investigate whether a similar explanation holds here.

We have shown previously [30] that the predicted secondary structures of the Xyn10A CBMs are in good agreement with the secondary structure of CBM4-1 of CenC, whose binding residues are all located in a binding cleft [40]. The results discussed above are indicative of a binding cleft able to accommodate single polysaccharide chains. The fast on-rates measured by NMR suggest that access to the binding-site cleft is not hindered. Nevertheless, bulky or charged substituents do not bind, implying that the cleft has a restrictive geometry that does not allow larger ligands to bind.

Relation to family 4 CBMs

The only representative from family 4 CBMs that has been subject to structural characterization [40], and whose binding properties have been experimentally demonstrated [21], is CenC CBM4-1 from *C. fimi* (CfCBM4-1, previously CBDN-1). In addition, CfCBM4-2 (previously CBDN-2) and two other CBMs have been classified to this family on the basis of a comparison of primary structures [3]. Structural similarities within the members of this family are expected because the residues forming strands of the jelly-roll β -sandwich motif of CfCBM4-1 are generally well conserved [40]. Consequently, a comparison between the binding modules from *R. marinus* (RmCBM4-1 and RmCBM4-2) and CfCBM4-1 is relevant in the context of defining the relatedness between the CBMs studied and those belonging to family 4. We have earlier proposed that the Xyn10A CBMs belong to family 4 on the basis of a 25% sequence identity with CfCBM4-1 [25]. The aromatic amino acids, believed to be

involved in substrate binding, are well conserved between the modules, and predictions of the secondary structure of RmCBM4-2 and CfCBM4-1 [30] are in good agreement. The apparent lack of affinity for crystalline cellulose is another major similarity between Xyn10A CBMs and CfCBM4-1. Despite these functional and plausibly structural similarities, there are some clear differences between the Xyn10A binding modules and CfCBM4-1. The two cysteine residues that form a disulphide bridge in the CBM4-1 of CenC are lacking from the CBMs from *R. marinus*. Another functional difference is the xylan affinity that is displayed in contrast with that of CfCBM4-1. These differences, along with the relatively low sequence identity between the Xyn10A CBMs and other members of family 4, impose some lack of clarity on the classification of the CBMs. To emphasize these differences, we suggest that family 4 CBMs should be divided into two subfamilies, with CenC CfCBM4-1 representing one, and RmCBM4-1 and RmCBM4-2 representing the other. Functional characterization of several family 4 CBMs will be required to test the validity of such a subdivision.

Binding interaction

The results from the use of high-molecular-mass ligands all indicate that CBM4-2 binds more tightly than CBM4-1. In contrast, the NMR results from the use of hexasaccharides indicate very similar binding constants. The sequences of the two modules and their binding specificities are very similar, implying that the difference in binding is probably due to steric restraints imposed on the binding site of CBM4-2 by a few amino acid changes at the periphery of the binding site, rather than by any gross change in structure of the binding site. Further discussion will require a structure for either or both modules.

From measurements of the binding affinity of X_6 to CBM4-2 at different temperatures, it is possible to estimate the contributions of enthalpy and entropy to the overall free energy of binding. However, it is first necessary to estimate the dependence of enthalpy on temperature (i.e. the heat capacity), which we have done with measurements of the binding of X_6 to the xylan-binding module from *C. fimi*. We have assumed that the heat capacities for the binding of X_6 to the two xylan-binding modules are similar, because heat capacity depends largely on hydrophobic area buried in the binding, which will be very similar in the two cases. For the interaction with *C. fimi* CBM, we measured dissociation constants at 10 and 40 °C as 1300 and 280 μ M respectively. This implies a free energy change of 3.8 kJ/mol over 30 °C. It is already known [20] that $T\Delta S$ at 25 °C (i.e. approximately at the midpoint of these two temperatures) is -26.4 kJ/mol. Assuming that the entropy does not change with temperature, this results in a calculated change in enthalpy for the interaction of *C. fimi* CBM with X_6 of only 1.17 kJ/mol over 30 °C. This is a very small change: it corresponds to a heat capacity of only $-38.5 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$, equivalent to a buried hydrophobic surface area of only 50 Å² [41]. When this figure is used to estimate the change in enthalpy for the interaction of X_6 with CBM4-2, we calculate that the enthalpic contribution to binding is approx. -25.9 kJ/mol and that the entropic contribution ($T\Delta S$) is approx. $+4.2$ kJ/mol. In other words, the binding is driven almost entirely enthalpically.

Similar results were obtained for the binding of a variety of ligands to the CenC CBM, which also binds to an amorphous polysaccharide [21]; Tomme et al. [21] observed an enthalpically driven binding associated with a small change in heat capacity. In contrast, the CBM from *C. fimi* exocellulase, which binds crystalline cellulose, binds with a large favourable entropy change [42]. The difference between these observations can be

rationalized by noting that crystalline cellulose is in a fixed conformation and does not lose conformational entropy on binding, whereas amorphous cellulose and xylan are both disordered and therefore lose significant conformational entropy on binding. Nevertheless, the enthalpy of the interaction with the amorphous ligands is larger, probably (as suggested by Tomme et al. [21]) because (1) the interaction is stabilized by van der Waals forces and a number of hydrogen-bonding interactions and (2) the more restrictive binding site found on the modules that bind amorphous polysaccharides serves both to minimize the loss of conformational entropy of the protein on binding and to permit interactions with a larger area of the ligand.

Why should there be two CBMs?

A preliminary comparison of the relative adsorption to PASC of the full-length xylanase [30] and the individual CBMs gives no indications of any co-operative mode of action. Rather, they seem to exhibit an additive binding mode. The same observation was noted for the tandemly repeated CBMs of CenC from *C. fimi* [21]. Furthermore, the two modules have the same substrate specificities. The question therefore arises: What is the advantage of having two binding modules?

The two modules have different affinities: CBM4-2 binds more strongly than CBM4-1 to polysaccharides. The denaturation temperatures determined spectroscopically and verified by differential scanning calorimetry studies (M. Abou Hachem, E. Nordberg Karlsson and M. Andersson, unpublished work) suggest that the modules have different thermal stabilities. Moreover, ultrasonication had different effects on the two recombinant products: whereas the molecular mass of CBM4-2 on denaturing PAGE remained unaffected by sonication (Figure 1), an additional band of a higher molecular mass (not seen in the untreated controls) appeared after the ultrasonication of CBM4-1. The same pattern was observed when cells were heat-treated. These results all suggest that the CBM4-1 structure might be less rigid than that of CBM4-2. This might confer different properties on the two modules, which could for example enable the two modules to bind preferentially to substrates differing in accessibility, covalent modification or degree of branching. We are currently determining the tertiary structure of the modules to throw more light on this question.

Conclusions

Here we have confirmed our earlier theory that both *xyn10A* encoded N-terminal modules from *R. marinus* mediate the binding of carbohydrate by the full-length xylanase. The ligand specificity is for xylan, β -glucan and amorphous cellulose, which is unique among members of CBM4, and consistent with the modules' having a binding cleft rather than a flat surface as has been observed for most CBMs.

The two *xyn10A*-encoded CBMs show 88 % sequence identity but differ in their substrate affinities and stabilities. So far we can only speculate about the exact function and binding mechanism of the tandemly repeated CBMs. Tertiary structures of the CBMs as well as detailed binding studies and investigation of their interaction with Ca^{2+} would provide us with valuable information.

We gratefully acknowledge financial support from EU-Biotech program (BIO4-CT96-0488), CF-Environmental Fund, National Board for Industrial and Technical Development (NUTEK), Crafoord Foundation and BBSRC (U.K.).

REFERENCES

- 1 Gilbert, H. J. and Hazlewood, G. P. (1993) *J. Gen. Microbiol.* **139**, 187–194
- 2 Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, Jr., R. C. and Warren, R. A. J. (1991) *Microbiol. Rev.* **55**, 303–315
- 3 Tomme, P., Warren, R. A. J., Miller, Jr., R. C., Kilburn, D. G. and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Carbohydrates*, vol. 618 (Saddler, J. N. and Penner, M. H., eds.), pp. 142–161, American Chemical Society, Washington, DC
- 4 Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Theodorou, M. K., Morris, P. and Gilbert, H. J. (1996) *Biochem. J.* **319**, 515–520
- 5 Tomme, P., Warren, R. A. J. and Gilkes, N. R. (1995) in *Advances in Microbial Physiology*, vol. 37 (Poole, R. K., ed.), pp. 2–81, Academic Press, London
- 6 Black, G. W., Hazlewood, G. P., Millward-Sadler, S. J., Laurie, J. I. and Gilbert, H. J. (1995) *Biochem. J.* **307**, 191–195
- 7 Hazlewood, G. P. and Gilbert, H. J. (1992) in *Xylans and Xylanases* (Visser, J., Beldman, G., Kusters-van Someren, M. A. and Voragen, A. G. J., eds.), pp. 259–273, Elsevier Science Publishers, Amsterdam
- 8 Teeri, T., Reinikainen, T., Ruohonen, L., Jones, T. A. and Knowles, J. K. C. (1992) *J. Biotechnol.* **24**, 169–176
- 9 Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J. and Miller, Jr., R. C. (1993) *FEMS Microbiol. Lett.* **113**, 211–218
- 10 Gilkes, N. R., Warren, R. A. J., Miller, Jr., R. C. and Kilburn, D. G. (1988) *J. Biol. Chem.* **263**, 10401–10407
- 11 Linder, M. and Teeri, T. T. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12251–12255
- 12 Knowles, J., Lehtovaara, P. and Teeri, T. (1987) *Trends Biotechnol.* **5**, 255–260
- 13 Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeysens, M. (1988) *Eur. J. Biochem.* **170**, 575–581
- 14 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
- 15 Din, N., Damude, H. G., Gilkes, N. R., Miller, Jr., R. C., Warren, R. A. J. and Kilburn, D. G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11383–11387
- 16 Millward-Sadler, S. J., Poole, D. M., Henrissat, B., Hazlewood, G. P., Clarke, J. H. and Gilbert, H. J. (1994) *Mol. Microbiol.* **11**, 375–382
- 17 Millward-Sadler, S. J., Davidson, K., Hazlewood, G. P., Black, G. W., Gilbert, H. J. and Clarke, J. H. (1995) *Biochem. J.* **312**, 39–48
- 18 Winterhalter, C., Heinrich, P., Candussio, A., Wich, G. and Liebl, W. (1995) *Mol. Microbiol.* **15**, 431–444
- 19 Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Ferreira, L. M. A., Bolam, D. N. and Gilbert, H. J. (1997) *J. Biotechnol.* **57**, 59–69
- 20 Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J. and Williamson, M. P. (1999) *Structure* **7**, 853–864
- 21 Tomme, P., Creagh, A. L., Kilburn, D. G. and Haynes, C. A. (1996) *Biochemistry* **35**, 13885–13894
- 22 Takada, G., Karita, S., Takeuchi, A., Ahsan, M. M., Kimura, T., Sakka, K. and Ohmiya, K. (1996) *Biosci. Biotechnol. Biochem.* **60**, 1183–1185
- 23 Henrissat, B., Teeri, T. T. and Warren, R. A. J. (1998) *FEBS Lett.* **425**, 352–354
- 24 Alfredsson, G. A., Kristjánsson, J. K., Hjörleifsdóttir, S. and Stetter, K. O. (1988) *J. Gen. Microbiol.* **134**, 299–306
- 25 Nordberg Karlsson, E., Bartonek-Roxä, E. and Holst, O. (1997) *Biochim. Biophys. Acta* **1353**, 118–124
- 26 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- 27 Pace, C. N., Shirley, B. A. and Thomson, J. A. (1989) in *Protein Structure: a Practical Approach* (Creighton, T. E., ed.), pp. 311–313, IRL Press, Oxford
- 28 Wood, T. M. (1988) *Methods Enzymol.* **160**, 19–25
- 29 Scopes, R. K. (1974) *Anal. Biochem.* **59**, 277–282
- 30 Nordberg Karlsson, E., Bartonek-Roxä, E. and Holst, O. (1998) *FEMS Microbiol. Lett.* **168**, 1–7
- 31 Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR* **2**, 661–665
- 32 Sandström, J. (1982) *Dynamic NMR Spectroscopy*, Academic Press, London
- 33 Puls, J. and Schuseil, J. (1993) in *Hemicellulose and Hemicellulases* (Coughlan, M. P. and Hazlewood, G. P., eds.), pp. 1–27, Portland Press, London
- 34 Gómez, C., Navarro, A., Manzanares, P., Horta, A. and Carbonell, J. V. (1997) *Carbohydr. Polym.* **32**, 17–22
- 35 Kulicke, W.-M., Kull, A. H., Kull, W., Thielking, H., Engelhardt, J. and Pannek, J.-B. (1996) *Polymer* **37**, 2723–2731
- 36 Nieduszynski, I. A. and Marchessault, R. H. (1972) *Biopolymers* **11**, 1335–1344
- 37 Ogawa, K. and Yui, T. (1998) in *Polysaccharides* (Dumitriu, S., ed.), pp. 101–129, Marcel Dekker, New York
- 38 Sun, J. L., Sakka, K., Karita, S., Kimura, T. and Ohmiya, K. (1998) *J. Ferment. Bioeng.* **85**, 63–68
- 39 Sakka, K., Takada, G., Karita, S. and Ohmiya, K. (1996) *Ann. N. Y. Acad. Sci.* **782**, 241–251
- 40 Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G. and McIntosh, L. P. (1996) *Biochemistry* **35**, 14381–14394
- 41 Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci.* **4**, 2138–2148
- 42 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

Received 5 May 1999/2 September 1999; accepted 14 October 1999