Structure, Vol. 12, 1177–1187, July, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.str.2004.04.022

Ab Initio Structure Determination and Functional Characterization Of CBM36: A New Family of Calcium-Dependent Carbohydrate Binding Modules

Sheelan Jamal-Talabani,[1,2] Alisdair B. Boraston,[1,3] Johan P. Turkenburg,[1] Nicolas Tarbouriech,[1,4] Valérie M.-A. Ducros,[1] and Gideon J. Davies[1,*]

[1]Structural Biology Laboratory
Department of Chemistry
The University of York
Heslington, York YO10 5YW
United Kingdom

Summary

The enzymatic degradation of polysaccharides harnesses multimodular enzymes whose carbohydrate binding modules (CBM) target the catalytic domain onto the recalcitrant substrate. Here we report the ab initio structure determination and subsequent refinement, at 0.8 Å resolution, of the CBM36 domain of the *Paenibacillus polymyxa* xylanase 43A. Affinity electrophoresis, isothermal titration calorimetry, and UV difference spectroscopy demonstrate that CBM36 is a novel Ca^{2+}-dependent xylan binding domain. The 3D structure of CBM36 in complex with xylotriose and Ca^{2+}, at 1.5 Å resolution, displays significant conformational changes compared to the native structure and reveals the molecular basis for its unique Ca^{2+}-dependent binding of xylooligosaccharides through coordination of the O2 and O3 hydroxyls. CBM36 is one of an emerging spectrum of carbohydrate binding modules that increasingly find applications in industry and display great potential for mapping the "glyco-architecture" of plant cells.

Introduction

Turnover of plant cell wall material, which makes up the vast bulk of biological material on the planet, is exceedingly important to the carbon cycle. Central to the breakdown of cell wall biomass are polysaccharolytic microbes. They achieve the breakdown of polysaccharides with complex systems of degradative enzymes comprising glycoside hydrolases, polysaccharide lyases, and carbohydrate esterases. Their modular architecture is a hallmark (Gilkes et al., 1991; Henrissat and Davies, 2000). Indeed, current evidence indicates that multimodularity is crucial to the efficient hydrolysis of plant cell wall polysaccharides, which are otherwise quite recalcitrant to degradation (Bolam et al., 1998; Boraston et al., 2003a; Charnock et al., 2000; Ali et al., 2001; Zverlov et al., 2001; Maglione et al., 1992; Hall et al., 1995; Tomme et al., 1988; Din et al., 1994). Single bacterial enzymes frequently contain up to five or six modules, sometimes more, and include at least one catalytic module and multiple ancillary modules. The most frequently occurring ancillary modules are the carbohydrate binding modules (CBMs) that mediate the tight interaction of the parent enzyme with substrate.

Polysaccharides are a chemically and structurally diverse class of macromolecule. This complexity reflects the large variety of monosaccharide building blocks, the numerous glycosidic linkages that may link these monomers, and the potential for branching and "decoration." Even a reducing hexasaccharide has over 10^{12} potential isomers (Laine, 1994)! Plant cell walls contain polysaccharides with backbones of glucose (most notably cellulose and xyloglucan), xylose, mannose, galactose, or arabinose. Frequently these backbones are decorated with sugar side chains, as exemplified by arabino-xylan, a xylose polymer with α-1,2 and α-1,3 linked arabinose side chains. Polysaccharide backbones may also contain a mixture of linkages, such as with the cereal β-1,3-1,4 glucans, or a mixture of monosaccharides, such as the β-1,4 linked glucose and mannose of glucomannan. Thus, polysaccharolytic enzymes and their constituent CBMs must contend with great heterogeneity in polysaccharide structures. Biochemical studies have made it increasingly evident that CBMs have evolved fine specificities for particular target polysaccharides (Boraston et al., 2003a; McLean et al., 2002; Carrard et al., 2000). Structural studies in tandem with functional characterization are just beginning to unravel the molecular determinants of CBM binding specificity.

A number of CBM structures have been reported, and these have revealed predominantly β sheet "jelly-roll" structures. CBM "topography" matches that of the polysaccharide target, and aromatic residues have been shown to contribute significantly to binding. Hence, type A CBMs have a flat aromatic-lined surface to match the hydrophobic surface of crystalline substrates, type B CBMs have open grooves to accommodate single glucan chains, and type C CBMs are more "lectin like" and accommodate oligosaccharides in surface pockets and indentations (Boraston et al., 1999) (Figure 1). The "X9" domain from the *Paenibacillus polymyxa* xylanase Xyn43A is of interest because of its unique characteristics. Here we show, by calorimetric and UV difference methods, that it is a Ca^{2+}-dependent xylan binding domain. The structure has been solved ab initio by direct methods, at 1.2 Å, using a SeMet derivatized form of the protein and subsequently refined in native form at 0.8 Å resolution. Analysis of a form cocrystallized in the presence of xylotriose reveals both conformational changes and the molecular basis for the Ca^{2+} dependence of ligand binding: one of the xylose residues is coordinated to this Ca^{2+} via its O2 and O3 hydroxyls. We propose that the family of X9 domains be renamed in accordance with standard practice in the field now that the carbohydrate binding properties of a family member have been demonstrated. Thus the "CBM36"

*Correspondence: davies@ysbl.york.ac.uk
[2]Present address: Department of Biological Sciences, Lancaster University, Lancaster, LA1 4YQ, United Kingdom.
[3]Present address: Department of Biochemistry & Microbiology, University of Victoria, P.O. Box 3055 STN CSC, Victoria, British Columbia, V8W 3P6, Canada.
[4]Present address: EMBL Grenoble Outstation BP. 181, 38042 Grenoble Cedex 9, France.

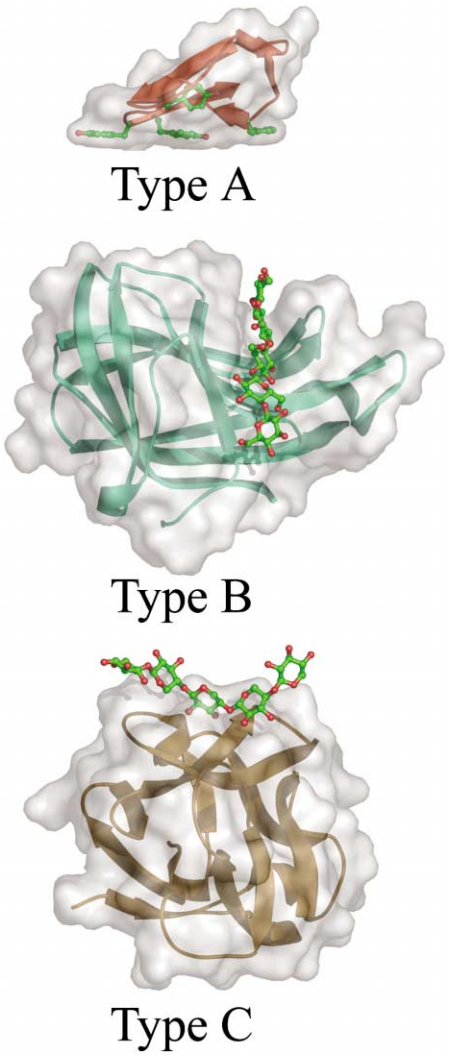

Figure 1. Three Classes of Carbohydrate Binding Modules

CBMs have been grouped intro three "types" on the basis of topography and binding properties of the domain (Boraston et al., 1999). Type A CBMs (for example, CBM1, 1cbh.pdb) bind via a flat planar surface which is complementary to the surface of crystalline cellulose; type B CBMs (for example, CBM4, 1gu3.pdb) display an open groove that can accommodate a single glucan chain, while type C CBMs (for example, CBM13, 1mc9.pdb) are more "lectin like." This figure was drawn with PyMOL (DeLano Scientific; http://pymol.sourceforge.net/).

structure, reported here at 0.8 Å resolution, is the first structural representative of an emerging group of metal ion dependent CBMs. Calcium-gated binding suggests that these CBMs may have potential applications in affinity labeling.

Results and Discussion

Xylanase 43 from *Paenibacillus polymyxa*

Xyn43 is a three-module protein comprising an N-terminal family GH43 catalytic module with xylanase/arabinofuranosidase activity (Gosalbes et al., 1991; Morales et al., 1995), an internal module having amino acid identity with family 6 CBMs, and a C-terminal module, here called CBM36. This latter module has sequence similarity with a family of modules of unknown function that occur exclusively in enzymes whose catalytic domains either have been shown to have xylanolytic activity or have been assigned as xylanases through sequence similarities (Figure 2). That noncatalytic modules appended to plant cell wall degrading enzymes are frequently carbohydrate binding, and the conservation of a number of aromatic residues (the signature motifs of polysaccharide recognition [Boraston et al., 1999]), led us to hypothesize that CBM36 modules are xylan binding CBMs.

CBM36 Is a Ca^{2+}-Dependent Xylan Binding CBM

Initial preliminary experiments using qualitative affinity electrophoresis (data not shown) showed that migration of CBM36 through native polyacrylamide gels, loaded with the appropriate polysaccharide, was slowed. CBM36 could thus be shown to bind tightly (i.e., large change in apparent mobility) to the polysaccharides glucuronoxylan and arabino-xylan, weakly to barley β-glucan [β-(1,3)(1,4)-glucan] and glucomannan (konjac) (i.e., small change in mobility), and not at all to derivatized cellulose (hydroxyethyl cellulose) or galactomannan (carob). This suggested that CBM36 was primarily a xylan binding CBM.

In order to quantify the binding of CBM36 to ligands and investigate the role of calcium in binding, isothermal titration calorimetry and UV difference spectroscopy were employed. Initial preparations of CBM36 that were extensively dialyzed against 50 mM potassium phosphate buffer yielded very small heats when xylooligosaccharides were titrated into this protein, and the data could not be confidently analyzed. This was anomalous for CBMs, which typically have large enthalpies of binding and, thus, yield large, significant heats in the calorimeter. Based on suspicions raised by the identification of a metal ion in the putative binding site of the CBM36 structure (see below), the ITC was subsequently performed in the presence of 2 mM $CaCl_2$, which then yielded larger heats (Figure 3). The thermodynamic values and stoichiometries (with $C > 1$, where $C = K_a \times [CBM] \times n$) could be obtained for xylopentaose and xylohexaose with reasonable confidence (using the method of Wiseman et al., 1989) (Table 1). In the case of xylotetraose, the low C values demanded that the stoichiometry be fixed at 1 and then analyzed by the same method.

The binding constants for CBM36 increase with oligosaccharide length up to xylohexaose, the longest oligosaccharide tested. This dependence on ligand length has been observed with many other CBMs (for examples, see Charnock et al., 2000; Tomme et al., 1996; Boraston et al., 2000). Under the experimental conditions, binding was dominated by a favorable change in enthalpy (ΔH) that was partially offset by an unfavorable change in entropy (ΔS), another feature common to CBM interactions with soluble glycans (for examples, see Charnock et al., 2000; Boraston et al., 2002). C values for binding to polymeric xylans were <1, and thus while generating affinity constants of $\sim 4 \times 10^3\ M^{-1}$ and $\sim 7 \times$

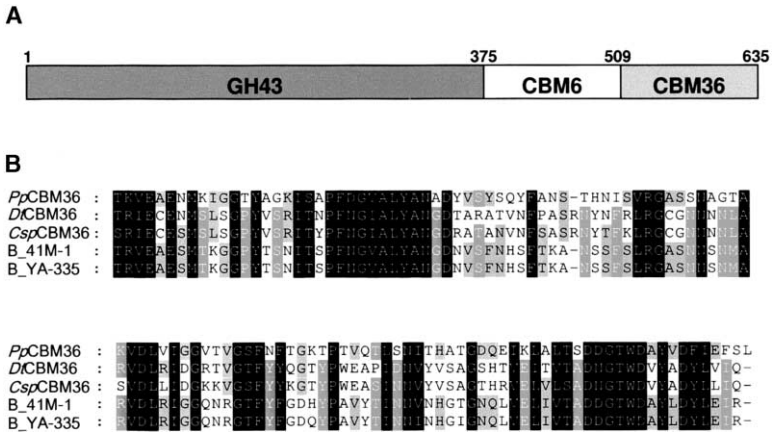

Figure 2. CBM36 and Its Occurrence in Enzymes

(A) The modular structure of *Paenibacillus polymyxa* Xyn43. Domain boundaries are given by amino acid sequence number above the module schematic.

(B) Alignment of modules showing significant identity with CBM36. Entries are as follows: *Pp*CBM36 from *Paenibacillus polymyxa* Xyn43 (GenBank accession no. X57094), *Dt*CBM36 from *Dictyoglomus thermophilum* xylanase 11 (U76545), *Csp*CBM36 from *Caldicellulosiruptor* sp. Rt69B.1 xylanase 11 (AF036925), B_YA-335 from *Bacillus sp*. (strain YA-335) xylanase 11 (X59059), and B_41M-1 from *Bacillus sp*. (strain 41M-1) xylanase 11 (AB029319).

10^3 M^{-1} for wheat arabinoxylan and birchwood glucuronoxylan, respectively, we could not deconvolute enthalpy and stoichiometry with confidence.

The sensitivity of UV-absorbing amino acid side chains to ligand binding was investigated by UV difference studies. Three UV difference spectra were collected: (1) CBM36 (with 10 mM calcium) perturbed by the addition of excess xylotriose, (2) CBM36 (with 10 mM calcium) perturbed by the addition of EDTA, and (3) a CBM36 (with 10 mM calcium)-xylotriose complex perturbed by the addition of EDTA. CBM36 (with 10 mM calcium) perturbed by the addition of excess xylotriose (spectrum 1) showed peaks and troughs that were diagnostic of tyrosine side chains moving into a more apolar environment (Figure 4). This is consistent with the xylotriose-complexed structure of CBM36 (discussed below) that shows two tyrosyl groups being shielded from solvent by the bound ligand. Interestingly, the UV difference spectrum of CBM36 (with 10 mM calcium) perturbed by the addition of EDTA to 25 mM (spectrum 2) showed a difference spectrum similar to the DMSO-perturbed *N*-acetyl-tryptophan difference spectrum (Figure 4) and, thus, indicative of a structural rearrangement involving tryptophan residues caused by the removal of the bound metal ions. In order to assess the influence of stripping bound metal ions from the ligand-bound form of CBM36, the last difference spectrum, a CBM36 (with 10 mM calcium)-xylotriose complex perturbed by the addition of EDTA (spectrum 3), was collected (Figure 4). This spectrum was very similar to spectrum 2 (i.e., dominated by tryptophan signal); however, the 285 nm peak of this difference spectrum had an additional small shoulder at ~282 nm, which was likely due to a small difference in tyrosine signal caused by the dissociation of xylotriose concomitant with stripping the bound metal ions. This was investigated by subtracting spectrum 3 from spectrum 2, which should isolate any UV difference signal unique to the dissociation of xylotriose and, therefore, should approximate the inverse of spectrum 1. Indeed, the result of this was the inverse of the xylotriose-induced UV difference spectrum (Figure 4) and suggestive of the movement of tyrosyl side chains into a more polar environment upon addition of EDTA, which was entirely consistent with the dissociation of xylotriose caused by the stripping of calcium from CBM36 by EDTA.

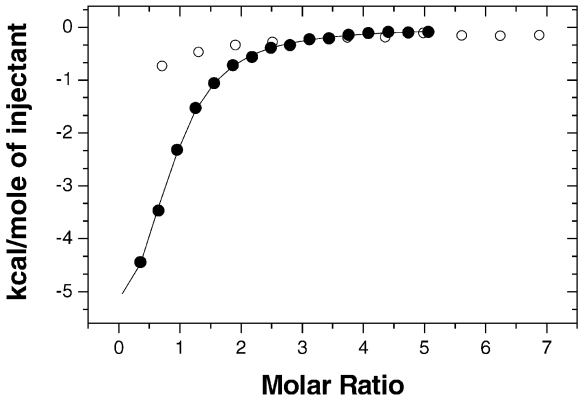

Figure 3. Isothermal Titration Calorimetry Isotherms of CBM36 Binding to Ligands

The experimental conditions were 200 μM of CBM36 binding to xylohexaose in 50 mM potassium phosphate buffer (open circles) and in 50 mM HEPES buffer with 2 mM $CaCl_2$ (closed circles). The solid lines show the fit of a one binding site model to the data obtained in HEPES buffer.

Quantitative UV difference titrations in the absence and presence of EDTA also allowed us to quantify the binding of short ligands to CBM36, which was not possible by ITC due to the low C values and very small heats. The affinity of CBM36 (with 10 mM calcium) for xylotriose thus measured is 1.1 (±0.1) × 10^3 M^{-1}. When EDTA was included at 25 mM, the affinity was reduced to the point where it was not quantifiable; however, the measurement of a UV difference signal did suggest a small degree of interaction (Figure 4C). Overall, these results demonstrated the dependence of ligand binding by CBM36 on the presence of calcium.

Ab Initio Structure Solution of CBM36

Initial crystals of native CBM36 were obtained from 1.6 M magnesium sulfate and 0.1 M MES (pH 6.5) and diffracted beyond 1.5 Å using CuKα radiation. In order to facilitate structure solution, the SeMet form of CBM36 was prepared but did not crystallize under the same conditions nor in standard screening using commercial crystal screens. Crystals of SeMet CBM36 were there-

Table 1. Xylooligosaccharide Binding Properties of CBM36

Ligand	K_a ($\times 10^{-3}$ M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol/K)	*n*
Xylohexaose	13.0 (± 1.1)	−5.6 (± 0.1)	−8.1 (± 0.5)	−8.4 (± 1.6)	0.8 (± 0.0)
Xylopentaose	7.3 (± 1.1)	−5.3 (± 0.1)	−8.4 (± 0.1)	−10.5 (± 0.4)	0.9 (±0.0)
Xylotetraose	3.2 (± 0.1)	−4.8 (± 0.3)	−8.1 (± 0.1)	−11.0 (± 0.5)	1[a]

Experiments were done in 50 mM HEPES (pH 7.5) supplemented with 2 mM $CaCl_2$ at 25°C. Errors represent the standard errors obtained from nonlinear fitting.
[a]This *n* value was fixed at unity during the analysis.

fore obtained by "streak-seeding" of crushed native crystals into preequilibrated hanging droplets under identical conditions as used for the native enzyme (see Experimental Procedures and Figure 5).

Single crystals of SeMet CBM36 diffracted to 1.0 Å at the Daresbury Synchrotron Radiation Source (SRS) on beamline PX 9.6. Data were collected, at a wavelength below that of the Se edge, to (partially) optimize the *f″* signal from the single Se site. This Se position was easily derived from manual inspection of the anomalous difference Patterson and was used as a starting atom for direct methods "ab initio" phase determination using ACORN (Foadi et al., 2000) from the CCP4 suite (CCP4, 1994) (Figure 5). Outstanding phases, with a main chain map correlation of 0.9 with the final F_{obs}, α_{calc} map, led to semiautomatic construction of the model (see Experimental Procedures for more details). Subsequently, the model of CBM36 was refined against 0.8 Å native data from the European Synchrotron Radiation Source.

The general fold of CBM36 is similar to other short polysaccharide CBM binders. It consists of eight β strands of two antiparallel β sheets, each of four strands, forming the classical β-jelly-roll fold observed in other CBMs (Figure 6). One face of the β sheet is concave, forming a long open cleft. At a superficial structural level, the fold most resembles a classical type B CBM. In the vast majority of type B CBM structures described thus far, the concave cleft formed by the curvature of the barrel forms the binding site (see Figure 1, for example) and is appropriately equipped with aromatic residues. This is not the case for CBM36, where instead this cleft is devoid of the aromatic residues that are the signature motif of CBMs.

Both edges of the CBM36 structure consist of loops. On one edge, three loops make a shallow cleft (~4.7 Å deep). This site is furnished by the single tryptophan in the structure, Trp121, and the aromatic side chains of Tyr40, Tyr26, and Tyr123. Close to Trp121, the loop region (Ser115-Asp116-Asp117-Gly118) is highly mobile, with poor side chain density. Given the demonstrated importance of these residues in binding, as indicated by UV difference spectra above, it was likely that this loop region, instead of the open groove, contributed to the binding site for ligands described below.

Two metal ions are found in the CBM36 native structure. The first, close to the short α-helical segment, is a calcium ion, located exactly as also seen in the *Cs*CBM6-3 structure, described more fully below. It is coordinated by the main chain carbonyls of Asp35 and Asp125, the OD1 and OE1 oxygens of Asp125 and Glu16, respectively, and a single water molecule. The classical heptacoordination is completed through a bidentate interaction with Glu18. All of these interacting partners are conserved in *Cs*CBM6-3. The second metal is coordinated by the main chain carbonyl of Tyr40, OD1 from Asp121, and four solvent molecules. Geometry and B value singly suggest that this is a Mg^{2+} ion, presumably from the 1.6 M $MgSO_4$ of the crystallization conditions. Metal ion coordination distances are shown in Figure 7. This metal site is that occupied by the second Ca^{2+} ion in the complex structure (discussed below), although in that case both ligand binding and conformational rearrangements contribute to markedly different coordination geometry.

The general fold of CBM36 shows closest similarity to members of CBM family 6, to which it also shows the most significant sequence similarity: structural similarity searches with DALI (Holm and Sander, 1993) reveal an overlap of 126 equivalent C-α atoms with rmsd 2.2 Å and 20% sequence identity with representative CBM6 family members. CBM6 modules are the "outliers" for CBM-ligand recognition. CBM6 modules consist of a classical lectin-like β-jelly-roll that contains two potential ligand binding clefts designated clefts A and B (Czjzek et al., 2001; Boraston et al., 2003b). Cleft B is the standard concave groove, but many CBM6 domains interact instead via a second surface area, cleft A, which corresponds approximately to the loop region of CBM36 furnished with aromatic residues described above. In order to explore ligand binding of CBM36, screening of complex crystals in the presence of Ca^{2+} was undertaken.

Structure of CBM36 in Complex with Xylotriose

A new crystal form of CBM36 was obtained in the presence of xylotriose and added Ca^{2+}. Electron density for two ordered xylose moieties and a third disordered group is evident in the 1.5 Å electron density map (Figure 8A). These have been assigned the numbers Xyl3, Xyl2, and Xyl1 (the latter disordered) in accordance with IUPAC nomenclature. Given the internal symmetry of xylooligosaccharides (the only "chemical" difference between the two possible orientations is the swapping of C5 with O5), it is difficult to assign the chain direction. The orientation has been assigned solely on the behavior of the C5 and O5 temperature factors of the central (most well-ordered) xyloside during refinement. In the chosen orientation, the C5 and O5 B values for xylosides 2 and 3 are both 24 $Å^2$, whereas in the "reverse" orientation the corresponding B values for C5/O5 refine to 21 and 31 $Å^2$, respectively. The corresponding values for the second xyloside are 33/31 $Å^2$ in the "correct" orientation and 29/36 $Å^2$ in the "reverse" orientation. We cannot,

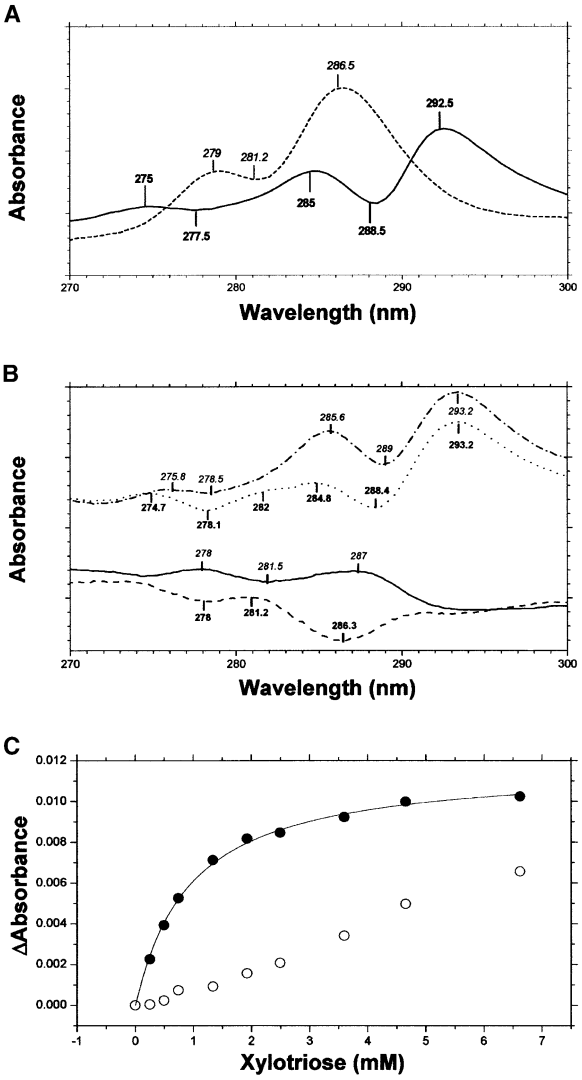

Figure 4. UV Difference Analysis of CBM36

(A) Difference spectrum of *N*-acetyltryptophan (solid line) and *N*-acetyltyrosine (dotted line) perturbed in 20% DMSO.

(B) Difference spectra are as follows: difference induced by excess xylotriose (solid line), difference induced by EDTA (dash-dotted line), difference induced by EDTA in the presence of excess xylotriose (dotted line), and net difference of the difference induced by EDTA and difference induced by EDTA in the presence of excess xylotriose (dashed line). Peak and trough wavelengths are indicated in both panels.

(C) UV difference binding isotherms of CBM36 binding to xylotriose. Experiments were performed in 50 mM Tris (pH 8.0) plus 10 mM $CaCl_2$ (closed circles) and in 50 mM Tris (pH 8.0) plus 10 mM $CaCl_2$ and 25 mM EDTA (open circles). The solid line shows the fit of a one binding site model.

however, rule out the possibility that the chain binds partially in the reverse orientation.

Xylotriose does not bind to the concave cleft B as observed for most CBM families. Consistent with its structural and sequence similarity to CBM6, xylotriose binds on one edge of the β sheets through a series of interactions with loop regions (Figures 6–8). Significantly, the complex reveals the basis for the unique calcium-dependent binding of xylo-oligosaccharides and xylan. The central xylose moiety, Xyl2, is coordinated, through both O2 and O3 hydroxyls, to a new heptacoordinate calcium ion (Ca-2) (Figures 7C and 8A). In the native structure, the position, but not the coordination, of this metal was taken by Mg^{2+}. In addition to the common coordination of this metal by the OD1 oxygen of Asp121 and the main chain carbonyl of Tyr40, the Ca-2 completes its equatorial plane coordination through interactions with the OD1 oxygen of Asp116 and the main chain carbonyl of Trp120, while a solvent water molecule completes the axial coordination (in addition to the main chain carbonyl of Tyr 40). The ordered xylose moiety at the nonreducing end (Xyl3) makes no direct hydrogen bonds to CBM36 in this structure; instead, it is wedged between the hydrophobic side chains of Tyr26 and Tyr40 (the latter also shielding Xyl2), consistent with the partial burial of these residues indicated by UV difference spectra (Figure 4).

Significant conformational changes, primarily in the loop from residue 113 to 121, accompany the Ca^{2+}-mediated binding of xylotriose (Figure 8B). These may result from the 4 Å movement of the side chain of Asp116 in order for it to interact with Ca-2. The most significant side chain movement is that of Trp120, which flips some 5–10 Å in order to fill the hydrophobic void caused by the main chain loop migration. Thus, Trp120 becomes internalized and more buried upon complex formation. These results are partially at odds with the UV difference results. Such a large change in the environment of a tryptophan upon binding, as was observed in the liganded versus unliganded crystal structures, would surely result in a noticeable tryptophan signal in the UV difference upon addition of xylotriose to calcium loaded CBM36. However, xylotriose binding to calcium-loaded CBM36 only perturbs tyrosine residues. A possible contributor to the observed structural change in the crystal structures may be the presence of the magnesium, rather than calcium, in the binding site of the native structure. Because of the different sphere of interactions between Mg^{2+} and the protein relative to Ca^{2+}, we are unable to state with certainty whether the movement of the tryptophan residue is a ligand-induced conformational change or whether it derives from the change in metal ion coordination on going from Mg^{2+} to Ca^{2+}. However, given that the UV difference results suggest a strong dependence of tryptophan absorbance on the presence of metal ions but a lack of tryptophan signal from xylotriose binding, it would appear that the conformational change is most likely a function of the binding of Ca^{2+} rather than oligosaccharide.

CBM36 interacts only weakly with xylotriose (K_d ~1 mM), and binding increases with chain-length up to oligosaccharides with a degree of polymerization of at least six (K_d for xylohexaose ~70 μM). There are additional hydrogen bonding residues at the reducing end where Xyl1 is currently observed but disordered. The protein surface is more extensive, but still limited, beyond the nonreducing end as currently observed. The structural similarity of CBM36 to *Cs*CBM6-3 (Boraston et al., 2003b) does not reveal any clues as to the binding of longer substrates (Figure 8C). Xylotriose binds to *Cs*CBM6-3 (in a manner not dependent on Ca^{2+}) in a

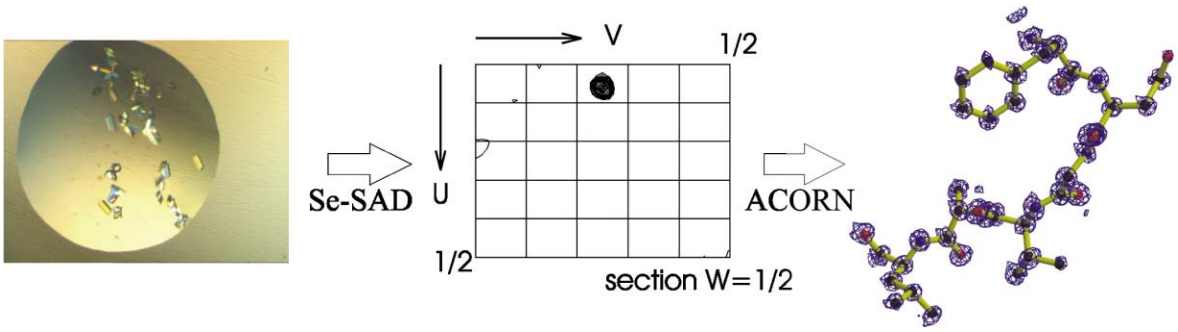

Figure 5. Ab Initio Structure Solution of CBM36

The structure solution of CBM36 is notable both for the requirement to streak-seed using native crystals in order to generate SeMet crystals and for the extremely rapid (<1 min) determination of almost perfect phases using ACORN and the anomalous Patterson-derived Se atom position as a starting "seed." The map shown is the 1.0 Å ACORN "E-map" with normalized structure factors as amplitudes and weighted ab initio phases.

position displaced approximately 7 Å from that observed in CBM36 (Figure 8C). Thus, the CBM36 ligand is not exactly in cleft A but is displaced to one edge where it instead interacts with the loop-coordinated Ca-2. While Tyr40 of CBM36 is equivalent to Tyr56 of the *Cs*CBM6 (where it interacts with the central xylose moiety), it is displaced about 4 Å.

Conclusions

The module of previously unknown function termed "X9" is shown to be a calcium-dependent xylan binding module. It has therefore been renamed the CBM36 module in recognition of this fact. The CBM36 module from *Paenibacillus polymyxa* xylanase 43A binds xylooligosaccharides and xylan chains with a low affinity that increases with the length of the ligand and is driven by favorable enthalpy. Such low-affinity constants are typical for CBMs derived from mesophilic organisms and assayed at the appropriate temperature. In contrast, high affinities tend to be reported for those derived from thermophilic organisms but assayed at low temperatures that (artifactually) enhance their affinity. The crystal structures of the unbound and liganded protein revealed a β-jelly-roll fold and a binding site at one edge of the fold, as observed for CBM6 and as UV difference spectroscopy suggested, rather than across the concave face as for most CBMs. The binding of ligands on the nonstandard face of the CBM strengthens the similarity between both *Cs*CBM6-3 and CBM36 and to lectins such as galactose and fucose specific lectins. We have commented previously (Boraston et al., 2003b) that this makes the distinction between "lectins" and enzyme-bound CBMs less clear. Indeed, the difference between type B and type C CBMs, described in the introduction, is also rather blurred by these observations.

As with many carbohydrate-active enzymes, *P. polymyxa* Xyn43, from which CBM36 was derived, is multimodular with a GH43 catalytic domain and CBM6 domain "upstream" of the C-terminal CBM36 module. The presence of multiple CBMs in bacterial and plant carbohydrate-active enzymes is very common, and the significance is unclear. Often, additional binding domains generate increased affinity through avidity effects, perhaps to overcome the intrinsically low binding constants

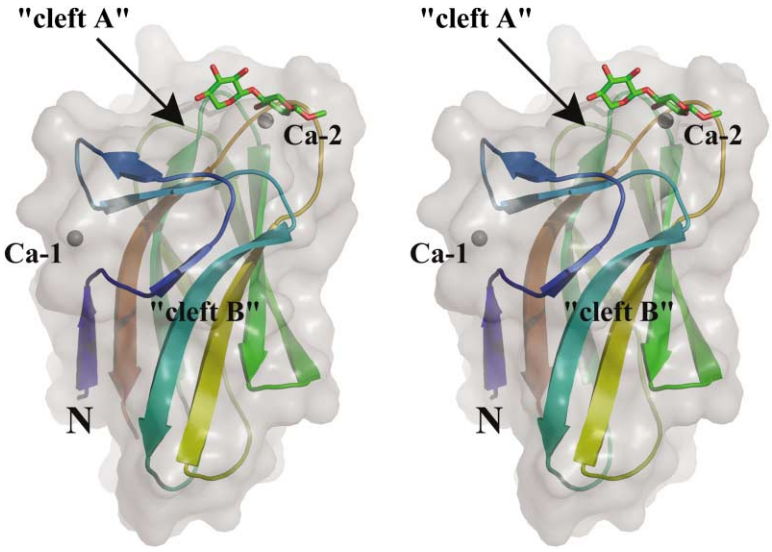

Figure 6. 3D Structure of CBM36

Protein cartoon of the xylotriose complex of CBM36, color-ramped from N terminus (blue) to C terminus (red). Ca^{2+} ions are shown as shaded spheres, and the xylotriose is shown in ball-and-stick representation. The clefts, "A" and "B," observed as ligand binding surfaces in CBMs from different families are indicated. This figure was drawn with PyMOL (DeLano Scientific; http://pymol.sourceforge.net/).

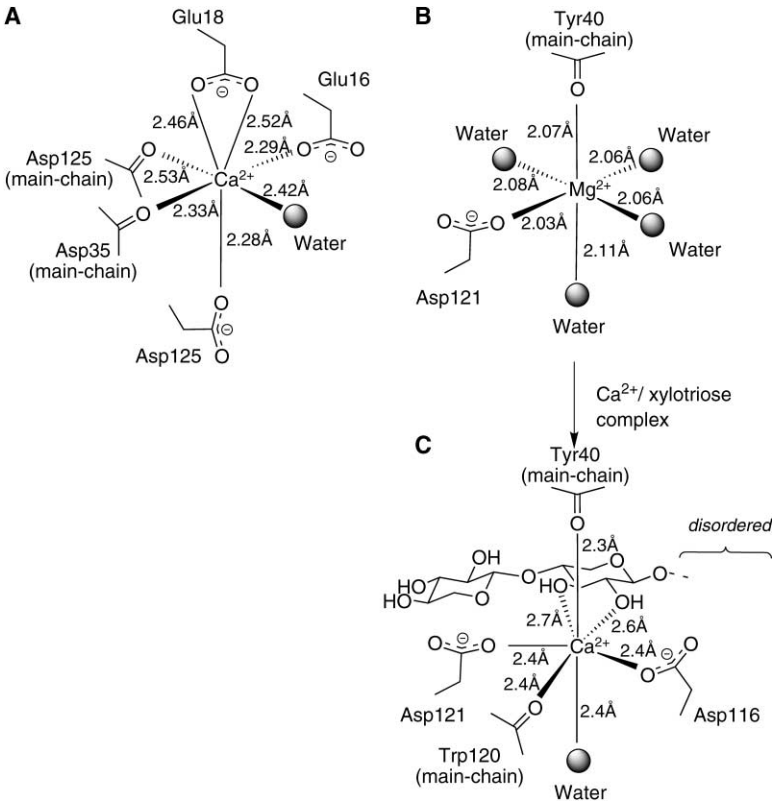

Figure 7. Schematic Diagram of the Refined Metal Ion Coordination Geometries of CBM36

(A) Ca^{2+} and (B) Mg^{2+} in the native CBM36 at 0.8 Å resolution, and (C) the interactions of the xylotriose-coordinating Ca^{2+} (which takes the place of the observed Mg^{2+} of the native structure) in the oligosaccharide complex.

which may be needed to provide high off-rates and the potential for diffusion over the complex plant cell wall substrates.

We believe that CBM36 is the first structure of a CBM for which Ca^{2+}-dependent binding has been reported, although recently X4/CBM35 modules have also been shown to be Ca^{2+} dependent (Bolam et al., 2004). Given the burgeoning roles of CBMs, not only in enzyme targeting but in analysis of cellular glyco-architecture (McCartney et al., 2004) and more tangentially in proteomic approaches such as tandem affinity purification (TAP)-tagging (for example, Gavin et al., 2002), the Ca^{2+} "gating" of ligand binding and subsequent elution with EDTA hints at new applications for this expanding and unusual class of carbohydrate binding module.

Experimental Procedures

Carbohydrates and Polysaccharides

Xylo-oligosaccharides, barley β-glucan, konjac glucomannan, and carob galactomannan were purchased from MegaZyme Ltd. (Bray, Co. Wicklow, Ireland). Birchwood xylan (Roth 7500; MW ~25000) was obtained from Carl Roth RG (Karlsruhe, Germany). Water-soluble xylan was prepared according to previously described procedures (Blake and Richards, 1971). Hydroxyethyl cellulose was from Sigma-Aldrich (Gillingham, U.K).

Preparation of Genomic DNA

Paenibacillus polymyxa (ATCC 842) was grown for 48 hr at 30°C in 50 ml yeast extract-phosphate medium (TYP) (Sambrook et al., 1989). The cells were pelleted by centrifugation and washed with 2 ml of TE (25 mM Tris [pH 8.0], 10 mM EDTA). The cells were then taken up in 0.7 ml TE containing 50 mM glucose, 0.5% sodium dodecyl sulfate, 10 mg/ml lysozyme, and 1 mg/ml RNaseA and incubated at 37°C for 4 hr. An equal volume of a solution containing phenol and chloroform (1:1) was added, and the mixture was vortexed and then centrifuged for 30 min at 30,000 × g. This procedure was repeated three times. Genomic DNA was precipitated by the addition of NaCl to 0.5 M and isopropanol to 40% and then recovered by spooling the precipitated material onto the tip of a glass Pasteur pipette. The genomic DNA was air-dried at room temperature and dissolved in TE.

DNA Amplification and Cloning

A gene fragment encoding 149 amino acid residues, 126 of which are the C-terminal CBM36 module from *P. polymyxa* xylanase 43A (GenBank accession number X57094), was obtained and amplified by standard PCR procedures using *P. polymyxa* genomic DNA preparation as a template. Oligonucleotide primers were designed to have 5′ NheI and 3′ HindIII restriction sites. The resulting PCR fragments were digested with NheI and HindIII and ligated into pET28a, previously digested with the same restriction enzymes, to give pET-CBM36. DNA was sequenced by the NAPS Unit (Biotechnology Laboratory, University of British Columbia) using the AmpliTaq dye termination cycle sequencing protocol and an Applied Biosystems Model 377 sequencer. The encoded polypeptides comprised a hexahistidine (H_6) tag fused with a thrombin cleavage site to the 126 amino acids of CBM36 module, respectively.

Protein Purification

Overnight cultures of *E. coli* strain BL21 DE3/pET-CBM36 were diluted 100-fold in TYP supplemented with 50 μg kanamycin/ml and grown at 30°C to a cell density (A_{600nm}) of ~0.3. Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.1 mM. Incubation was continued for a further 6 hr at 30°C. 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 by gentle mixing in 20 mM Tris HCl (pH 8.0) with 500 mM NaCl. Cells were ruptured by two passages through a French pressure cell (21,000 lb/in^2, 147,746 kPa) or by sonication, and cell debris was removed by centrifugation for 30 min at 27,000 × g and 4°C. CBM36 was purified from the clarified cell extract by Ni^{2+} affinity chromatog-

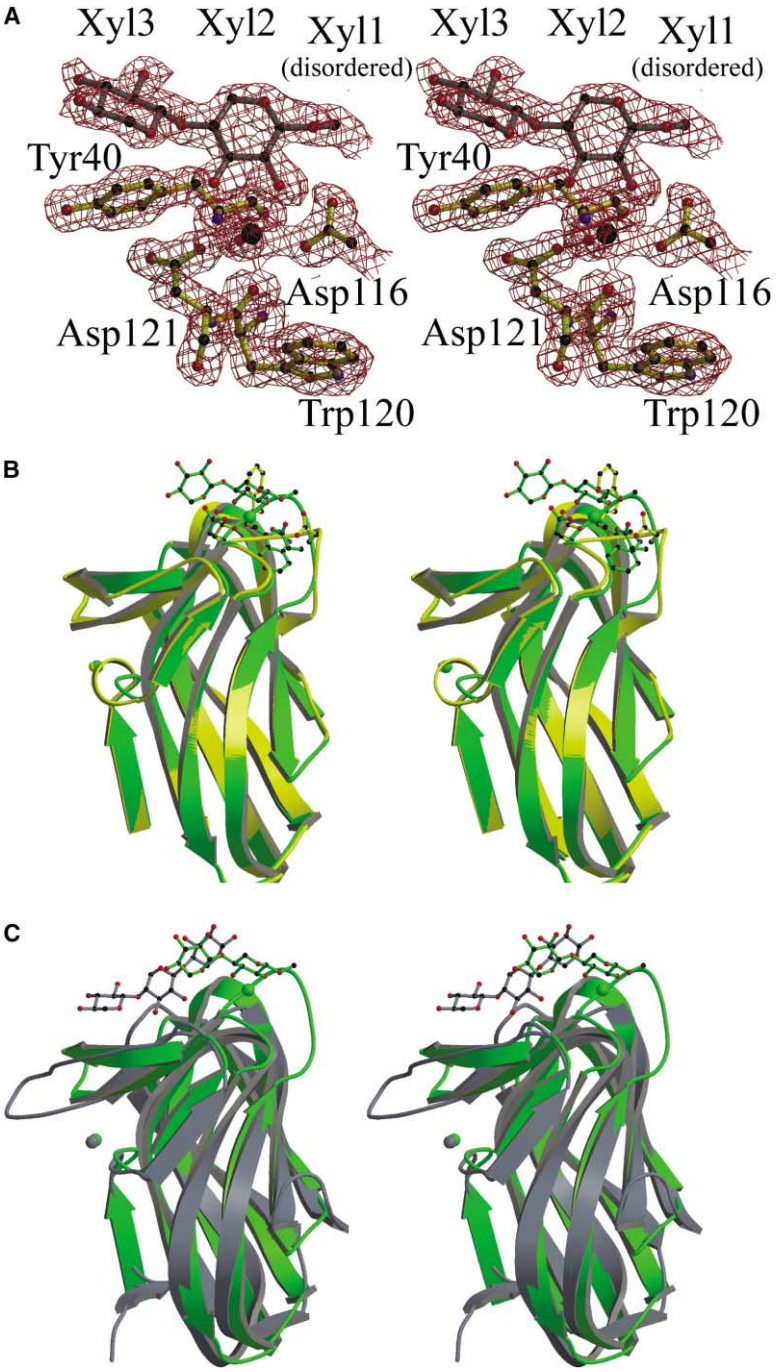

Figure 8. Ligand Binding and Conformational Change Associated with CBM36

(A) Observed electron density for the Ca^{2+}/xylotriose bound to CBM36. The first xylose at the reducing end is disordered, and only O4 and C4 have been modeled. The map shown is a maximum-likelihood and σ_A (Read, 1986) weighted $2F_o - F_c$ synthesis at 0.5 electrons/Å^3.

(B) Overlap of native (yellow) and xylotriose-bound (green) CBM36 molecules showing the conformational changes associated with the binding of Ca^{2+}/xylotriose compared to the native structure (in the presence of 1.6 M $MgSO_4$).

(C) Overlap of the CBM36 complex (this work, green) with the xylotriose complex of the *Clostridium stercorarium Cs*CBM6-3 (Boraston et al., 2003b). The structural similarity, including complete conservation of a "structural" calcium site, is clear, as is the similar "cleft A" location of ligand.

These figures, all in divergent (wall-eyed) stereo were drawn with BOBSCRIPT (A) (Esnouf, 1997) and MOLSCRIPT (B and C) (Kraulis, 1991).

raphy using His-Bind resin (Novagen, Madison, WI) as a column matrix. Standard manufacturer recommended procedures were used.

Seleno-methionine-labeled CBM36 was produced using a protocol described elsewhere (Notenboom et al., 2001) using *E. coli* B834/DE3 as the expression host. The labeled CBM36 was purified by IMAC as above, except that all buffers contained 1 mM β-mercaptoethanol (βME). Purified CBM36 was exchanged into distilled water, or distilled water with 1 mM βME for seleno-methionine-labeled CBM36, and concentrated in a stirred ultrafiltration unit (Amicon, Beverly, MA) on a filter with an exclusion limit of ~1 kDa (Filtron, Northborough, MA).

Determination of Protein Concentration

The concentrations of purified CBM36 were determined by UV absorbance (280 nm) using a calculated molar extinction coefficient (Mach et al., 1992) of 13,900 $M^{-1}cm^{-1}$.

Binding Studies

Affinity electrophoresis was performed as described previously (Tomme et al., 2000). Isothermal titration calorimetry (ITC) was performed using a MCS ITC (MicroCal, Inc., Northampton MA). All samples were pH 7.5 in 50 mM Na-phosphate buffer or HEPES buffer, the latter being supplemented with 2 mM $CaCl_2$. Samples were filtered and degassed extensively prior to use. Titrations were per-

Table 2. Structure Statistics

	CBM36-SeMet	CBM36 Native	CBM36-Xylotriose
Diffraction Data			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions (Å)	a = 30.7 b = 40.6, c = 82.4	a = 30.5 b = 40.2, c = 81.2	a = 38.6, b = 52.1, c = 54.8
Resolution (Å) (outer shell)	30–1.0 (1.02–1.00)	25–0.8 (0.84–0.8)	30–1.5 (1.55–1.50)
R_{merge}	0.070 (0.42)	0.057 (0.37)	0.043 (0.094)
Completeness (%)	98 (77)	97 (97)	99 (97)
$I/\sigma I$	25 (2)	12 (3)	27 (12)
Multiplicity	5 (2)	4 (2)	5 (3)
Radiation source	SRS PX9.6	ESRF ID14-EH3	ESRF ID14-EH4
Detector	ADSC Quantum-4 CCD	MAR 165 mm CCD	ADSC Quantum-4 CCD
Refinement			
R_{cryst}		0.11	0.155
R_{free}		0.13	0.20
Rms deviation of 1–2 bonds (Å)		0.024	0.011
Rms angle deviation (°)		1.6	1.46
Number of protein atoms[a]		1160	892
Number of solvent water		206	160
Number of metal ions		1 × Ca^{2+}, 1 × Mg^{2+}	2 × Ca^{2+}
Mean B protein (Å^2)		9	14
B value metal ions (Å^2)		Ca^{2+} 6; Mg^2 6	15/11
Mean B xylotriose (Å^2)			28
PDB code		1W0N	1UX7

[a] The apparent discrepancy in the number of protein atoms reflects different modeling of disorder in the two structures.

formed at 25°C by injecting 5–10 μl aliquots of sugar solution into the ITC sample cell (volume = 1.3528 ml) containing 100–200 μM CBM36. For xylopentaose and xylohexaose, this maintained C values ($C = n \times K_a \times A$; where n is the stoichiometry, K_a is the association constant, and A is the CBM concentration) greater than 1. Due to constraints on protein concentration and quantities, as well as the low affinity, the C value for xylotetraose could not be greater than one. In this instance, the stoichiometry (n value) was fixed at unity during the analysis. Analysis was performed essentially as described by Wiseman et al. (1989).

All UV absorption experiments were performed on a Cary 300 Bio UV-VIS spectrophotometer with a cell block maintained at 25°C. Unless otherwise stated, all samples were in 50 mM Tris (pH 8.0) with 10 mM $CaCl_2$. Scans were collected with a 0.2 nm interval, a 12 nm/min scan rate, and 2.0 nm spectral bandwidth. DMSO-perturbed difference spectra of 50 μM N-acetyltryptophan and 100 μM N-acetyltyrosine were collected as described (Boraston et al., 2000). The UV difference spectrum of xylotriose-perturbed CBM36 was collected by taking a baseline with 400 μl of 70 μM CBM36 and then adding ~2 mg of powdered xylotriose (final xylotriose concentration of ~12 mM), mixing, and equilibrating by inversion and then rescanning. This sample was then used for a baseline and then was perturbed by the addition of 3 mg of EDTA to collect a UV difference spectrum of xylotriose-ligated CBM36 perturbed by EDTA. UV difference spectrum of EDTA-perturbed CBM36 was collected in a similar manner except that the xylotriose was replaced by 3 mg of EDTA.

Quantitative binding studies were performed using 2 ml of 70 μM CBM36 in a quartz cuvette with continuous stirring. Additions of a 50 mM solution of xylotriose were done with a micrometer driven syringe pump. The peak-to-trough height at 287 nm and 281 nm was used to monitor binding. The isotherms were fitted as described previously (Boraston et al., 2000).

Crystallization of CBM36

CBM36 was prepared for crystallization by 6–12 hr treatment with thrombin at room temperature in 25 mM Tris HCl (pH 8.0) to remove the N-terminal H_6 tag. The free tag was removed by passing the cleavage reaction over an IMAC column. The flow through contained polypeptides lacking the H_6 tag. These were extensively dialyzed against distilled water and concentrated in stirred ultrafiltration unit. Crystals of native CBM36 (14.9 mg/ml) were grown using the vapor diffusion technique from hanging drops in 1.6 M magnesium sulfate and 0.1 M MES (pH 6.5). Crystals of selenomethionine CBM36 were grown under the same conditions but required streak seeding with native CBM36 crystals to initiate crystal growth.

Structure Determination and Refinement

A single crystal of SeMet CBM36 was mounted in a Rayon fiber loop, and X-ray data to 1.0 Å were collected at the Daresbury SRS on beamline PX9.6 at a wavelength of 0.87 Å (Table 2). Data were processed and reduced with the HKL suite (Otwinowski and Minor, 1997); all further computing used the CCP4 suite (CCP4, 1994) unless otherwise stated. The position of the single Se site was determined by manual inspection of the anomalous difference Patterson (Figure 6), and this position was used as a starting seed for direct methods phasing using ACORN (Foadi et al., 2000). Ab initio phases were determined (approximately 1 min elapsed time on a MAC G5) and converged, after 51 cycles, with a correlation value for the medium E values of 0.45. The initial ACORN-generated electron density map (F_{obs}, PHI_{ACORN} FOM_{ACORN}), at 1.0 Å, has a mean main chain map correlation of 0.91 and side chain/water correlation of 0.71 with the comparable map calculated from the refined coordinates below.

REFMAC/ARP/wARP (Murshudov et al., 1997; Perrakis et al., 1999) with the warp-N-trace option was used to automatically trace the CBM36 sequence into electron density. Optimal chain-tracing was achieved only when the resolution was cut to 1.5 Å for the ARP/wARP job. Following automatic docking of 119 residues, 11 internal residues left unbuilt by ARP/wARP were added manually using QUANTA (Accelrys, San Diego, CA), and a single Ca^{2+} ion, one Mg^{2+} ion, one sulfate ion, and the selenium atom were included. Maximum-likelihood refinement including anisotropic refinement of atomic displacement parameters, at 1.0 Å, with REFMAC resulted in a preliminary model (residues 3–121 inclusive, one Ca^{2+}, one Mg^{2+}, one SO_4^{2-}, 258 waters with B < 42 Å^2, with $R_{cryst} = 0.15$, $R_{free} = 0.17$) that was refined no further and instead used as the starting model for refinement of the 0.8 Å native structure.

Native data were collected on ESRF beamline ID14-EH3 using a MAR 165 mm CCD as detector (Table 2). The atomic resolution data were collected through the use of a "swung-out" detector. Data were processed and reduced using MOSFLM/SCALA (Leslie, 1992) (Table 2). The preliminary SeMet CBM36 model, described above,

was used as the starting point for refinement of the native structure at 0.8 Å resolution.

CBM36 was cocrystallized in the presence of xylotriose from 25% (w/v) polyethylene glycol 2000 monomethyl ether, 300 mM Na-acetate, and Mes 6.5. Data for the complex, collected on ESRF beamline ID14-EH4 to 1.5 Å resolution, were processed and reduced using the HKL suite (Otwinowski and Minor, 1997). The structure was solved by molecular replacement using the refined native coordinates (with metal ions and waters removed) with AMoRe (Navaza and Saludijan, 1997) using the program-assigned default values and refined with REFMAC, as above.

Acknowledgments

Professors Tony Warren and Douglas Kilburn (UBC) are thanked for their continued support. This work was funded by the BBSRC, and S.J. thanks the University of York for funding her Ph.D. G.J.D. is a Royal Society University Research Fellow and recognizes the support of the Peter Wall Institute for Advanced Studies (UBC) for supporting him during the initial phase of this work.

Received: February 26, 2004
Revised: April 15, 2004
Accepted: April 18, 2004
Published: July 13, 2004

References

Ali, M.K., Hayashi, H., Karita, S., Goto, M., Kimura, T., Sakka, K., and Ohmiya, K. (2001). Importance of the carbohydrate-binding module of *Clostridium stercorarium* Xyn10B to xylan hydrolysis. Biosci. Biotechnol. Biochem. *65*, 41–47.

Blake, J.D., and Richards, G.N. (1971). An examination of some methods for fractionation of plant hemicelluloses. Carbohydr. Res. *17*, 253–268.

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.

Bolam, D.N., Hefang, X., Pell, G., Hogg, D., Galbraith, G., Henrissat, B., and Gilbert, H.J. (2004). X4 modules represent a new family of carbohydrate-binding modules that display novel properties. J. Biol. Chem., in press.

Boraston, A.B., McLean, B.W., Kormos, J.M., Alam, M., Gilkes, N.R., Haynes, C.A., Tomme, P., Kilburn, D.G., and Warren, R.A.J. (1999). Carbohydrate-binding modules: diversity of structure and function. In Recent Advances in Carbohydrate Bioengineering, B. Svensson, ed. (Cambridge, UK: Royal Society of Chemistry), pp. 202–211.

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.

Boraston, A.B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D.R., Kilburn, D.G., and Davies, G.J. (2002). Differential oligosaccharide recognition by evolutionarily-related beta-1,4 and beta-1,3 glucan-binding modules. J. Mol. Biol. *319*, 1143–1156.

Boraston, A.B., Kwan, E., Chiu, P., Warren, R.A., and Kilburn, D.G. (2003a). Recognition and hydrolysis of noncrystalline cellulose. J. Biol. Chem. *278*, 6120–6127.

Boraston, A.B., Notenboom, V., Warren, R.A.J., Kilburn, D.G., Rose, D.R., and Davies, G.J. (2003b). Structure and ligand binding of carbohydrate-binding module CsCBM6–3 reveals similarities with fucose-specific lectins and "galactose-binding domains." J. Mol. Biol. *327*, 659–669.

Carrard, G., Koivula, A., Soderlund, H., and Beguin, P. (2000). Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. Proc. Natl. Acad. Sci. USA *97*, 10342–10347.

CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

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.

Czjzek, M., Bolam, D.N., Mosbah, A., Allouch, J., Fontes, C.M.G.A., Ferreira, L.M.A., Bornet, O., Zamboni, V., Darvon, H., Smith, N.L., et al. (2001). The location of the ligand-binding site of carbohydrate-binding modules that have evolved form a common sequence is not conserved. J. Biol. Chem. *276*, 48580–48587.

Din, N., Damude, H.G., Gilkes, N.R., Miller, R.C., Jr., Warren, R.A., and Kilburn, D.G. (1994). C1-Cx revisited: intramolecular synergism in a cellulase. Proc. Natl. Acad. Sci. USA *91*, 11383–11387.

Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced colouring capabilities. J. Mol. Graph. Model. *15*, 132–134.

Foadi, J., Woolfson, M.M., Dodson, E., Wilson, K.S., Yao, J.X., and Zheng, C.D. (2000). A flexible and efficient procedure for the solution and phase refinement of protein structures. Acta Crystallogr. *D56*, 1137–1147.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller, R.C., Jr., and Warren, R.A. (1991). Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. Microbiol. Rev. *55*, 303–315.

Gosalbes, M.J., Perez-Gonzalez, J.A., Gonzalez, R., and Navarro, A. (1991). Two beta-glycanase genes are clustered in *Bacillus polymyxa*: molecular cloning, expression, and sequence analysis of genes encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase. J. Bacteriol. *173*, 7705–7710.

Hall, J., Black, G.W., Ferreira, L.M., Millward-Sadler, S.J., Ali, B.R., Hazlewood, G.P., and Gilbert, H.J. (1995). The non-catalytic cellulose-binding domain of a novel cellulase from *Pseudomonas fluorescens* subsp. cellulosa is important for the efficient hydrolysis of Avicel. Biochem. J. *309*, 749–756.

Henrissat, B., and Davies, G.J. (2000). Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol. *124*, 1515–1519.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Laine, R.A. (1994). A calculation of all possible oligosaccharide isomers both branched and linear yields structures for a reducing hexasaccharide: the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems. Glycobiology *4*, 759–767.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. In Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography, Volume 26 (Warrington, UK: Daresbury Laboratory).

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.

Maglione, G., Matsushita, O., Russell, J.B., and Wilson, D.B. (1992). Properties of a genetically reconstructed *Prevotella ruminicola* endoglucanase. Appl. Environ. Microbiol. *58*, 3593–3597.

McCartney, L., Gilbert, H.J., Bolam, D.N., Boraston, A.B., and Knox, J.P. (2004). Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. Anal. Biochem. *326*, 49–54.

McLean, B.W., Boraston, A.B., Brouwer, D., Sanaie, N., Fyfe, C.A., Warren, R.A., Kilburn, D.G., and Haynes, C.A. (2002). Carbohydrate-binding modules recognize fine substructures of cellulose. J. Biol. Chem. *277*, 50245–50254.

Morales, P., Sendra, J.M., and Perez-Gonzalez, J.A. (1995). Purification and characterization of an arabinofuranosidase from *Bacillus polymyxa* expressed in *Bacillus subtilis*. Appl. Microbiol. Biotechnol. *44*, 112–117.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. Acta Crystallogr. *D53*, 240–255.

Navaza, J., and Saludijan, P. (1997). AMoRe: an automated molecular replacement program package. Methods Enzymol. *276*, 581–594.

Notenboom, V., Boraston, A.B., Kilburn, D.G., and Rose, D.R. (2001). Crystal structures of the family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10a in native and ligand-bound forms. Biochemistry *40*, 6248–6256.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. *A42*, 140–149.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (New York: Cold Spring Harbor Laboratory Press).

Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988). Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. Eur. J. Biochem. *170*, 575–581.

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.

Tomme, P., Boraston, A., Kormos, J.M., Warren, R.A., and Kilburn, D.G. (2000). Affinity electrophoresis for the identification and characterization of soluble sugar binding by carbohydrate-binding modules. Enzyme Microb. Technol. *27*, 453–458.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal. Biochem. *179*, 131–137.

Zverlov, V.V., Volkov, I.Y., Velikodvorskaya, G.A., and Schwarz, W.H. (2001). The binding pattern of two carbohydrate-binding modules of laminarinase Lam16A from *Thermotoga neapolitana*: differences in beta-glucan binding within family CBM4. Microbiol. *147*, 621–629.

Accession Numbers

Native and xylotriose-bound coordinates and structure factor amplitudes have been deposited with the PDB (1W0N and 1UX7, respectively).