The X6 "Thermostabilizing" Domains of Xylanases Are Carbohydrate-Binding Modules: Structure and Biochemistry of the *Clostridium thermocellum* X6b Domain[†,‡]

Simon J. Charnock,[§] David N. Bolam,[||] Johan P. Turkenburg,[§] Harry J. Gilbert,[||] Luis M. A. Ferreira,[⊥] Gideon J. Davies,*[,§] and Carlos M. G. A. Fontes[⊥]

Department of Chemistry, Structural Biology Laboratory, University of York, Heslington, York YO10 5DD, U.K., Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K., and CIISA-Faculdade de Medicina Veterinária, Rua Gomes Freire, 1199 Lisboa Codex, Portugal

Received December 8, 1999; Revised Manuscript Received January 31, 2000

ABSTRACT: Many polysaccharide-degrading enzymes display a modular structure in which a catalytic module is attached to one or more noncatalytic modules. Several xylanases contain a module of previously unknown function (termed "X6" modules) that had been implicated in thermostability. We have investigated the properties of two such "thermostabilizing" modules, X6a and X6b from the *Clostridium thermocellum* xylanase Xyn10B. These modules, expressed either as discrete entities or as their natural fusions with the catalytic module, were assayed, and their capacity to bind various carbohydrates and potentiate hydrolytic activity was determined. The data showed that X6b, but not X6a, increased the activity of the enzyme against insoluble xylan and bound specifically to xylooligosaccharides and various xylans. In contrast, X6a exhibited no affinity for soluble or insoluble forms of xylan. Isothermal titration calorimetry revealed that the ligand-binding site of X6b accommodates approximately four xylose residues. The protein exhibited K_d values in the low micromolar range for xylotetraose, xylopentaose, and xylohexaose; 24 μM for xylotriose; and 50 μM for xylobiose. Negative ΔH and ΔS values indicate that the interaction of X6b with xylooligosaccharides and xylan is driven by enthalpic forces. The three-dimensional structure of X6b has been solved by X-ray crystallography to a resolution of 2.1 Å. The protein is a β-sandwich that presents a tryptophan and two tyrosine residues on the walls of a shallow cleft that is likely to be the xylan-binding site. In view of the structural and carbohydrate-binding properties of X6b, it is proposed that this and related modules be re-assigned as family 22 carbohydrate-binding modules.

The plant cell wall consists mainly of a complex mixture of polysacccharides, primarily pectins, cellulose, and hemicellulose (*1*). In most plant material, xylan is the major hemicellulose component. This polysaccharide comprises a backbone of β-1,4-linked xylopyranose units, which can be decorated at the 2′ and/or 3′ carbons with arabinose, 4-methylglucuronic acid, or acetate moieties (*2*). The nature and the extent of xylan substituents varies between plants, with softwoods and cereals containing high levels of arabinoxylan, while hardwood hemicellulose consists mainly of glucuronoxylan (*3*). The glycosidic bonds in the xylan backbone are hydrolyzed by endo-β-1,4-xylanases (EC 3.2.1.8) (*4*). Currently, the primary structures of over 150 xylanases have been deduced from gene or cDNA sequences (*5*). Primary structure and hydrophobic cluster analyses have shown that xylanases can be classified into two glycoside hydrolase families, 10 and 11 (*5*). The catalytic modules of family 10 enzymes have an $(\alpha/\beta)_8$ barrel structure (*6*) and an approximate M_r of 40 000 and hydrolyze various aryl-β-glycosides in addition to their natural substrates xylooligosaccharides and xylan (*7*, *8*). In contrast, family 11 xylanases are smaller (M_r of 20 000) β-stranded molecules (*9*) that exhibit absolute specificity for β-1,4-linked xylose-derived substrates (*10*).

It is now widely established that xylanases, in common with the majority of plant cell wall hydrolases, often exhibit a modular structure comprised of catalytic domains linked to one or more noncatalytic modules (NCM)[1] (*11*). The majority of the characterized NCMs are carbohydrate-binding modules (CBM) (*12*, *13*). On the basis of primary structure comparisons, CBMs have been grouped into 21 different families (*14*). Thus far, the majority of the well-characterized

[†] This work was funded by the Biotechnology and Biological Sciences Research Council and Fundação para a Ciência e Tecnologia project PraxisXXI/C/CVT/11041/98. G.J.D. is a Royal Society University Research Fellow.

[‡] Coordinates for the structure described in this paper have been deposited with the Brookhaven Protein Data Bank under Accession No. 1DYO.

* Corresponding author telephone: +44-(0)1904-432596; fax: +44-1904-410519; e-mail: davies@ysbl.york.ac.uk.

[§] University of York.

[||] University of Newcastle upon Tyne.

[⊥] CIISA-Faculdade de Medicina Veterinária.

[1] Abbreviations: ANDE, affinity nondenaturing gel electrophoresis; CBM, carbohydrate-binding module; CD, catalytic domain; ESRF, European Synchrotron Radiation Facility; ITC, isothermal titration calorimetry; IPTG, isopropyl β-thiogalactopyranoside; MAD, multiple wavelength anomalous dispersion; NCM, noncatalytic module; NCS, noncrystallographic symmetry; TD, thermostabilizing domain; XBM, xylan-binding modules.

10.1021/bi992821q CCC: $19.00
Published on Web 04/07/2000

modules interact with cellulose (*12*, *13*, *15*), although CBMs that bind other polysaccharides such as xylan, chitin, and starch have also been described (*16*−*18*). The 3D structures of several CBMs are known. Those that exhibit affinity for crystalline cellulose contain surface aromatic residues that form a planar hydrophobic surface that is implicated in binding to the ligand (*19*−*22*). In contrast, domains that bind to single polysaccharide chains display an open cleft, such as the CBM4 from *Cellulomonas fimi* CenC and the $CBM2_b$ from *C. fimi* Xyn11A, that bind single cellulose and xylan chains, respectively (*23*, *24*). The role of the cellulose- and xylan-binding CBMs is to potentiate the catalytic activity of cellulases and hemicellulases against complex substrates. The mechanism by which they mediate this effect varies. For example, *Pseudomonas* $CBM2_a$s and CBM10s enhance the activity of plant cell wall hydrolases by bringing the enzymes close to their substrates (*25*−*27*). In contrast, the corresponding *Cellulomonas* modules aid cellulose hydrolysis not only by proximity but also by increasing the accessibility of the macromolecule through dispersion of the the polysaccharide chains (*28*, *29*).

A large number of NCMs remain uncharacterized and have been referred to as "X modules" (*14*). NCMs from a family designated "X6" are frequently found adjacent to family 10 and family 11 xylanase CDs. Deletion of these modules often leads to a reduction in the structural integrity of the protein at elevated temperatures; hence, they have been termed "thermostabilizing domains" to reflect this destabilization (*30*−*32*). X6 modules were initially identified in enzymes from thermophilic microorganisms, but they are now known to be present in mesophilic prokaryotes (*33*, *34*) and eukaryotes, including *Arabidopsis thaliana*. The occurrence of the X6 module in nonthermophilic organisms, coupled with the fact that they are frequently found in arrays of up to four X6 repeats, hints that the apparent thermostabilizing properties of these modules may simply be related to the physical consequences of protein truncation.

The recent discovery of true xylan-binding modules (XBMs) in several xylanases (*16*, *18*) suggests that other noncatalytic modules of previously undetermined function may also play a role in xylan binding. *Clostridium thermocellum* Xyn10B (formerly xylanase Y) comprises two X6 modules, flanking a family 10 glycoside hydrolase CM, together with a dockerin and a C-terminal family 1 esterase, Figure 1 (*30*). In this paper, we describe in detail the biochemical properties of the X6b module. X6b, expressed as a discrete protein, binds tightly to xylooligosaccharides and various substituted xylans but not to numerous other (nonxylo) oligo- and polysaccharides. In harness with the catalytic core, X6b potentiates the enzyme activity on insoluble xylans. The 3D structure of the X6b module has been determined by X-ray crystallography at a resolution of 2.1 Å using selenomethionyl-derived protein and the multiple-wavelength anomalous dispersion (MAD) method. The structure of X6b is a classic β-jelly roll, similar to that seen in many lectin families. A shallow surface groove contains three highly conserved aromatic residues that are implicated in xylan binding. We propose that the X6 domains are reclassified as carbohydrate-binding module family 22 to reflect their true biochemical properties.

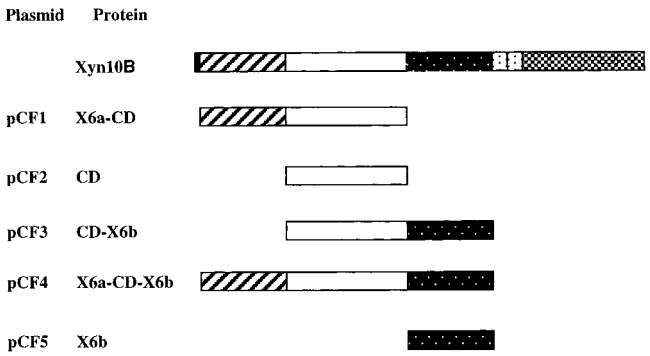

FIGURE 1: Molecular architecture of Xyn10B and its derivatives. Domains present in Xyn10B are as follows: X6a (white bars with black stripes), X6b (black bars with white dots), the xylanase CD module (white bars), the dockerin sequence (white bars with black dots), a family 1 polysaccharide esterase homology module (checkered bars), and the signal peptide (solid black bars).

MATERIALS AND METHODS

Bacteria, Vectors, and Growth Media. Escherichia coli strains used in this study were JM83 (*35*), BL21(DE3), and B834(DE3) (*36*). Recombinant *E. coli* was cultured in Luria broth supplemented with 100 μg/mL ampicillin at 37 °C. To induce gene expression in strains containing derivatives of the gene encoding Xyn10B (*xyn10B*) cloned into pET21a, liquid cultures were grown to mid-log phase ($A_{550} = 0.4$), after which time isopropyl β-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM, and the bacteria were incubated for a further 4 h. The vectors used in this study were pET21a (Novagen) and pGEM-T (Promega).

Production of Truncated Derivatives of xyn10B and Protein Purification. To express truncated derivatives of *xyn10B*, the appropriate DNA sequences were amplified using PCR from *C. thermocellum* (strain YS) genomic DNA. The reactions were performed using PfuTurbo DNA polymerase (Stratagene) as described previously (*30*). The resulting PCR fragments were cloned into pCR-Blunt as described in the Promega protocol. The primers, containing *Nde*I and *Xho*I restriction sites, used to construct the *xyn10B* derivatives were as follows:

pCF1, 5'CTCGCTAGCGATTATGAAGTGGTTCATG3'
5'CACCTCGAGGGCCGGATTGTTACCGTC 3'

pCF2, 5'CTCGCTAGCGCAAACTATTTCAGAGTTG3'
5'CACCTCGAGGGCCGGATTGTTACCGTC 3'

pCF3, 5'CTCGCTAGCGCAAACTATTTCAGAGTTG3'
5'CACCTCGAGTCCTTCGATTACAGTTCC 3'

pCF4, 5'CTCGCTAGCGATTATGAAGTGGTTCATG3'
5'CACCTCGAGTCCTTCGATTACAGTTCC 3'

pCF5, 5'CTCGCTAGCAAACCGGAAGAGCCGGATG 3'
5'CACCTCGAGTCCTTCGATTACAGTTCC 3'

Truncated derivatives of Xyn10B were purified by metal-chelate affinity chromatography from cell-free extracts, derived from 200-mL cultures, using His Trap columns (Pharmacia) as described previously (*30*). Fractions recovered from the column were analyzed by SDS−PAGE (*37*), and those containing the appropriate purified polypeptides were dialyzed against 2 × 1000 mL vol of 50 mM sodium phosphate buffer, pH 7.4.

Enzyme and Protein Assays. All enzyme assays were performed at 60 °C in 50 mM sodium phosphate buffer, pH 7.4, using 0.2% of the appropriate plant polysaccharide, unless otherwise stated (*16*). Soluble and insoluble xylan were prepared from oat spelt xylan as described previously (*38*). Reducing sugar was measured with the dinitrosalicylic acid reagent (*39*). Protein concentration was determined by measuring either the A_{280} or the Lowry method, with BSA as standard.

Ligand Binding Studies. Affinity of the derivatives of Xyn10B for insoluble xylan and cellulose was measured as described previously (*16*). The capacity of X6b to bind different soluble polysaccharides was evaluated by affinity nondenaturing gel electrophoresis (ANDE). ANDE was performed by preparing continuous polyacrylamide gels consisting of 7.5% (w/v) acrylamide in 25 mM Tris/250 mM glycine buffer, pH 8.3. To one of the gels, 0.1% of the appropriate polysaccharide was added prior to polymerization. Approximately 5 μg of target proteins and BSA as a nonbinding negative control were mixed with Bromophenol Blue and subjected to electrophoresis at room temperature with a current of 10 mA/gel for 1−2 h. Proteins were visualized by staining with Coomassie Blue. Isothermal titration calorimetry (ITC) was performed as described previously (*24*), except that the concentration of protein was 50 μM, the concentration of the stock ligand solutions was 3 mM.

Production of Seleno-L-methionine Containing X6b. The methionine auxotroph *E. coli* B834(DE3) transformed with pCF5 was cultured at 37 °C in 1 L of a liquid growth medium that comprised 2 mM $MgSO_4$; 37 mM NH_4Cl; 44 mM KH_2PO_4; 85 mM Na_2HPO_4; 0.09 mM $FeSO_4$; 22.2 mM glucose; 40 mg each of the 20 common L-amino acids (except methionine, which was substituted with seleno-L-methionine, Sigma); 1 mg each of riboflavin, niacinamide, pyridoxine monohydrochloride, and thiamine; and 50 μg/mL ampicillin until the A_{550} reached 0.4, at which point expression of the recombinant protein was induced with the addition of 1 mM IPTG. Incubation was continued for a further 12 h, after which time the cells were collected by centifugation and sonicated in 25 mM Na-Hepes buffer, pH 7.5, containing 10 mM imidazole, 1 mM β-mercaptoethanol, and 1 M NaCl. The cell-free extract was centrifuged at 30 000*g*, 4 °C for 1 h and loaded onto a 5-mL chelating Sepharose (FF) column (Pharmacia). Bound proteins were washed with 50 mL of sonication buffer and then eluted with a stepped gradient of imidazole (25−150 mM). SDS−PAGE was used to identify fractions containing purified X6b, which were subsequently exchanged into 10 mM DTT using PD-10 Sephadex G-25 M gel filtration columns (Pharmacia) and concentrated to a volume of 100 μL with 10-kDa centifugation membranes (Pall Filtron). Using this procedure, the samples were washed with 10 mM DTT a further two times, and the protein concentration was finally adjusted to 50 mg/mL. MALDI-TOF mass spectrometric analysis confirmed the identity of the polypeptide and indicated that the N-terminal seleno-L-methionine residue had been processed by the host bacterium.

Crystallization, Data Collection, and Refinement. Crystals of X6b were grown by vapor phase diffusion using the hanging-drop method. The protein concentration was 50 mg/mL in 0.1 M NaAc buffer, pH 4.6, containing 10 mM DTT, 25% (v/v) glycerol, and 12% (w/v) PEG 8000 as the precipitant. Crystals were grown from microseeds for 1 day. A rayon-fiber loop was used to transfer a single crystal directly into liquid nitrogen. Preliminary X-ray diffraction analysis revealed that they belong to space group $P6_122$ or $P6_522$, with unit-cell dimensions $a = b = 90.08$ and $c = 207.71$ Å and with one or two molecules in the asymmetric unit.

Table 1: MAD Data Collection and Processing Statistics[a]

	peak	inflection	remote
wavelength (Å)	0.9797	0.9800	0.9386
resolution (Å)	12−2.5	12−2.5	12−2.5
redundancy	6.5 (6.3)	5.2 (4.9)	5.4 (5.2)
completeness (%)	94.9 (94.9)	85.9 (84.9)	90.2 (89.7)
$I/\sigma I$	30 (19)	26 (17)	25 (17)
R_{merge} (%)	7.5 (11.3)	6.8 (9.5)	7.3 (10.8)

[a] Highest resolution shell (in parentheses): 2.54−2.5A.

MAD X-ray Diffraction Data Collection and Processing. A three-wavelength MAD experiment was conducted on beamline ID14-4 at the European Synchrotron Radiation Facility (ESRF) at Grenoble using an ADSC Quantum-4 CCD detector. A single crystal of X6b was flash-cooled in the home laboratory and tested for diffraction quality. The crystal was preserved and transported to the ESRF, where it was remounted on the single-axis goniometer. The wavelengths for the MAD experiment were chosen by scanning through the absorption edge of the Se-X6b crystal. Data were then collected at the minimum f', the maximum f'', and a reference wavelength at an energy above the absorption edge, Table 1. After indexing an initial diffraction image using the program package HKL2000 (*40*), the program STRATEGY (*41*) was used to determine the optimal φ range to collect complete anomalous data using a minimal oscillation sweep. A total of 82 images with 1° oscillation was collected at each of the three wavelengths. To increase the multiplicity and thus the reliability of the data, a further 40 images were collected at each of the three wavelengths. Data were processed using DENZO/SCALEPACK as part of the HKL2000 suite of programs. Data for each wavelength were scaled in SCALEPACK, but multiple observations were not merged at this stage.

Location of Se positions and Phasing. The three unmerged data sets were input to SOLVE (*42*) and scaled and merged internally by this program. The space group ambiguity ($P6_1$-22 vs $P6_522$) was resolved by running SOLVE in both space groups and using the standard deviation of the local rms electron density to determine the correct enantiomorph. Automated Patterson searching, using data to 3.0 Å, readily located six Se positions, corresponding to two molecules of X6b each with three selenomethionine residues in space group $P6_122$. The Se positions were refined, and phases were calculated in SOLVE. These phases were used as a starting set for phase improvement in DM (*43*). To employ noncrystallographic symmetry averaging, the operator relating the two molecules in the asymmetric unit was derived from the positions of the six Se atoms. Two sets of three atoms, each forming nearly identical, irregular triangles could be identified. A least-squares fit of one set onto the other revealed the rotational and translational operators. These were then input to DM for NCS averaging with automatic mask determination.

Table 2: Binding of Xyn10B Truncated Derivatives to Xylan and Avicel[a]

enzyme	relative binding (%)	
	insoluble xylan	Avicel
X6a−CD	10 ± 4[c]	ND[b]
CD	12 ± 4	ND
CD−X6b	88 ± 2	8 ± 3
X6b	91 ± 5	11 ± 7
X6a−CD−X6b	84 ± 6	ND

[a] Purified proteins (0.5 g/L) were incubated with the same volume of 5% (w/v) insoluble polysaccharide. The quantity of protein that bound to the polymers was evaluated by determining protein concentration in the filtrate. [b] ND, not determined. [c] Results are means ± SEM (n = 3).

Model Building and Refinement. The electron density map calculated after DM displayed extensive well-defined regions revealing continuous stretches of main chain density, with unambiguous density for carbonyl oxygen atoms and side chains. A model comprising approximately 150 amino acids was built from the initial map using the X-AUTOFIT module in Quanta (Molecular Simulations Inc., San Diego, CA). This model was refined using the CCP4 program REFMAC (*44*), with the phases from DM included as experimental restraints.

High-Resolution Data and Refinement. Data to 2.1 Å resolution were collected on a Se−Met crystal on beamline BW7B (λ = 0.8445 Å) at the EMBL Hamburg outstation at the DESY synchrotron site (previous 2.3 Å data on a native, normal methionine sample showed that the structures are essentially identical; data not shown). A single crystal was grown, as described above, but with the addition of 10 mM $CaCl_2$ (a single structural Ca^{2+} ion had been identified during the 2.5 Å refinement). The crystal was mounted, essentially as described above, and 57.6 degrees of data were collected, with an oscillation angle of 0.6° per image with a MAR Research MAR345 image-plate detector. Data were processed and reduced as described previously in this paper. Refinement of the 2.5 Å model was continued using these data. Tight NCS restraints were used initially and gradually relaxed as deemed appropriate from the behavior of the cross-validation (R_{free}) subset of reflections (5%).

RESULTS AND DISCUSSION

Binding of X6b to Polysaccharides. To investigate the biochemical properties of the X6 domains in *C. thermocellum* Xyn10B, truncated forms of the xylanase containing a C-terminal His tag were expressed. The enzyme derivatives consisted of the following modules: X6a−CD, CD, CD−X6b, X6a−CD−X6b, and X6b (Figure 1). The five recombinant proteins were purified to apparent homogeneity, as judged by SDS−PAGE (data not shown). The capacity of the purified proteins to bind insoluble xylan and cellulose (both acid swollen and Avicel) was studied by incubating the proteins with the polysaccharides and determining the protein concentration in the filtrate (unbound protein). The results show that all proteins containing X6b bound the hemicellulose, whereas polypeptides lacking this module were unable to adhere to this polymer (Table 2). None of the five proteins exhibited significant affinity for cellulose.

A semiquantitative assessment of the capacity of X6b to bind to a series of different soluble polysaccharides was evaluated by ANDE (Figure 2). The migration of X6b was significantly retarded by the inclusion of oat spelt xylan in the gel, while the electrophoretic mobility of the control protein BSA was not affected by the presence of hemicellulose. The data show that X6b bound to various xylans and barley β-glucan; interacted extremely weakly with a soluble form of cellulose; and did not associate with arabinan, β-1,4-galactan, pectin (polygalacturonic acid and rhamnogalacturan), carob galactomannan, or amylose (Table 3). These results indicate that X6b is a specific xylan-binding module, while X6a did not bind to any of the ligands tested.

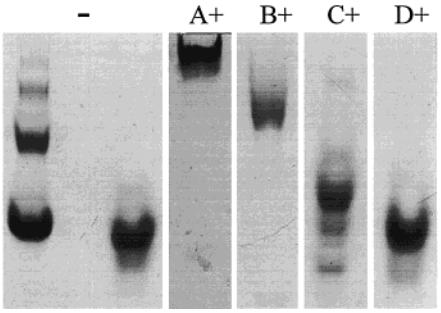

FIGURE 2: Affinity nondenaturing gel electrophoresis of X6b. X6b and BSA were electrophoresed in 7.7% nondenaturing polyacrylamide gels containing no polysaccharide (lanes 1 and 2, respectively), 0.1% oat spelt xylan (lane A+), 0.1% barley β-glucan (lane B+), hydroxyethylcellulose (lane C+), and apple pectin (lane D+). Proteins were visualized by Coomassie Blue staining.

Table 3: Semiquantitive Affinities of X6b for Various Plant Cell Wall Polysaccharides[a]

polysaccharide	binding to X6b
oat spelt xylan	+++
wheat arabinoxylan	+++
rye arabinoxylan	+++
hydroxyethyl cellulose	+
carboxymethyl cellulose	−
polygalacturonic acid	−
rhamnogalaturonan	−
carob galactomannan	−
β-1,4-galactan	−
α-1,5-arabinan	−
β-glucan	++

[a] X6b was subjected to ANDE in polyacrylamide gels containing different plant cell wall polysaccharides. Maximum, significant, slight, and no retardation of the protein by the polysaccharide is indicated by +++, ++, +, and −, respectively.

Thermodynamics of Xylooligosaccharides and Xylan Binding to X6b. ITC was used to quantify the affinity of X6b for xylooligosaccharides, xylan, β-glucan, and cellohexaose and to determine the thermodynamics of ligand binding (Figure 3). The data show that X6b interacts strongly with xylooligosaccharides with a degree of polymerization of 2 or greater. The protein exhibits similar binding constants for xylotetraose, xylopentaose, and xylohexaose (Table 4). These data suggest that X6b has a ligand-binding region that can accommodate four xylose moieties. The stoichiometry for all ligands bound to X6b at saturation was approximately 1:1, indicating that there was only one ligand-binding site per protein molecule. This is similar to the stoichiometry of ligand binding of CBM4 from *C. fimi* CenC (*45*) but is in contrast to the starch-binding domain of *Aspergillus* glucoamylase, which accommodates two amylose chains (*46*).

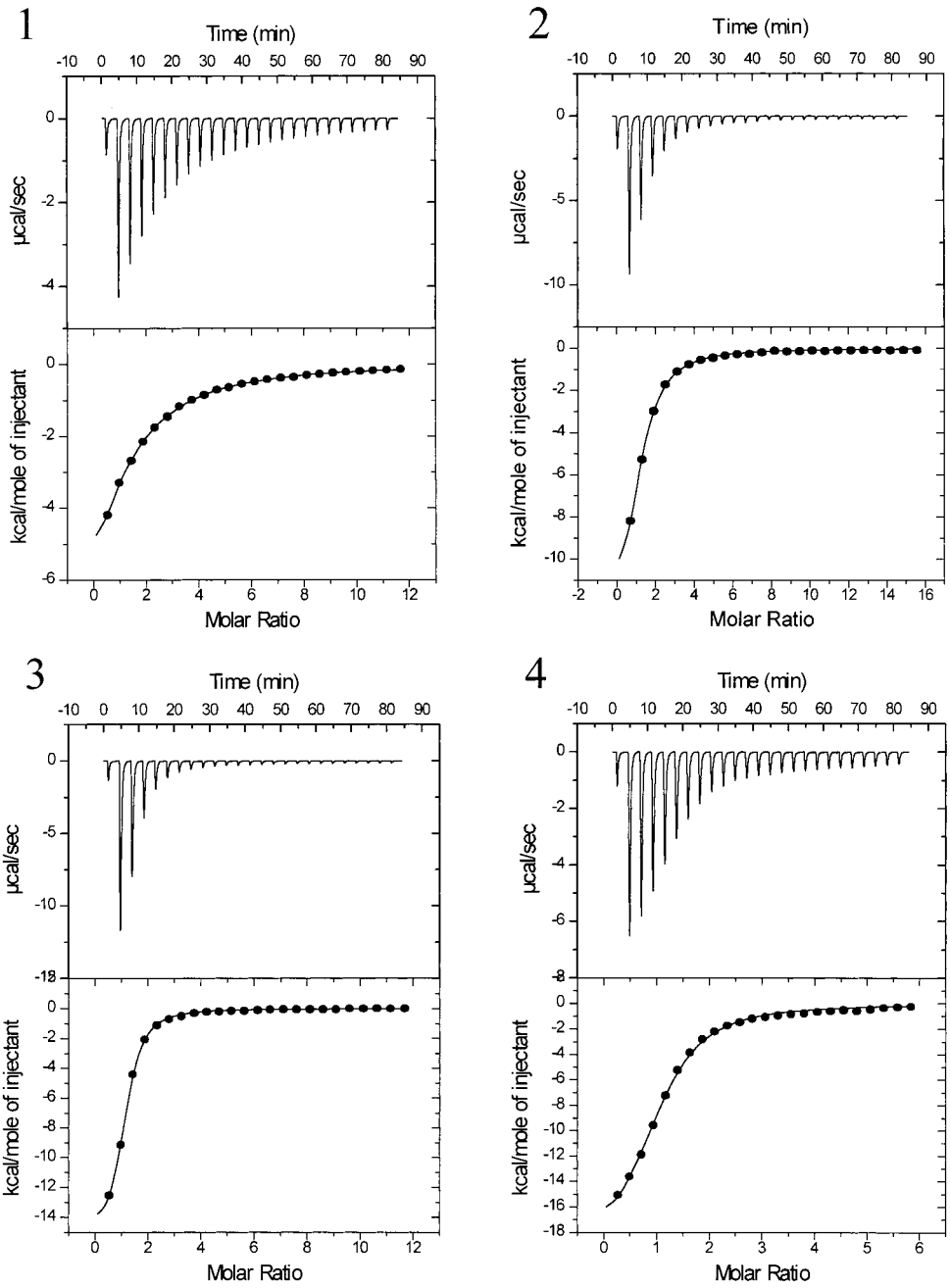

FIGURE 3: ITC analysis of the binding of xylan and xylooligosaccharides to X6b. X6b at a concentration of 50 μM was titrated with (1) xylobiose, (2) xylotriose, (3) xylotetraose, and (4) rye arabinoxylan. The top half of each experiment shows the calorimetric titration of X6b with ligand, and the lower half displays the integrated injection heats from the upper panel, corrected for control dilution heats. The solid line is the curve of best fit that was used to derive parameters n (stoichiometry), K_d, and ΔH. All titrations were performed in 50 mM sodium phosphate buffer, pH 7.0, at 25 °C.

To investigate the thermodynamics of binding, the association constant of X6b for its ligands (K_a) and the change in enthalpy (ΔH) were used to determine the change in Gibbs free energy (ΔG) and $T\Delta S$ for each ligand-binding event. The data (Table 3) showed that the binding of each oligosaccharide may be fitted by a one-binding site model. The ΔH and $T\Delta S$ values for ligand binding were always negative, and thus the interaction of X6b with carbohydrates is exothermic. The change in enthalpy makes a positive contribution to ligand binding, while the increase in entropy has a detrimental influence on the interaction of X6b with soluble oligosaccharides. The thermodynamics of X6b binding to xylooligosaccharides is similar to other proteins that bind soluble polymeric ligands (*47*) but is in contrast with CBMs that interact with insoluble crystalline cellulose, in which entropic forces dominate the binding event (*48*).

The interaction of proteins with carbohydrates may be mediated by hydrophobic interactions, such as those involving sugar rings and aromatic amino acids, and/or hydrogen bonds between hydroxyl groups of the sugar and the polar amino acids. It has been proposed that very large negative heat capacity (ΔC_p) values are associated with hydrophobic stacking interactions and result from the dehydration of aromatic side chains in the protein and the pyranose rings of the carbohydrate (*49*). The ΔC_p for the interaction of X6b with xylotetraose was determined by measuring the ΔH at 25 and 35 °C and gave a value of +60 cal mol^{-1} K^{-1}. Such a small positive value indicates a mixed binding mode

Table 4: Thermodynamics of X6b Binding to Various Carbohydrate Ligands As Determined by Isothermal Titration Calorimetry

	K_a (M^{-1})	$\Delta G°$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$T\Delta S°$ (kcal mol^{-1})	n
xylose	NB[a]				
xylobiose	$1.7 (\pm 0.08) \times 10^4$	$-5.76 (\pm 0.02)$	$-12.28 (\pm 0.07)$	$-6.52 (\pm 0.09)$	1.00
xylotriose	$4.2 (\pm 0.20) \times 10^4$	$-6.30 (\pm 0.03)$	$-14.97 (\pm 0.05)$	$-8.67 (\pm 0.08)$	1.04
xylotetraose	$1.4 (\pm 0.05) \times 10^5$	$-7.09 (\pm 0.02)$	$-15.65 (\pm 0.07)$	$-8.56 (\pm 0.09)$	1.00
xylopentaose	$2.0 (\pm 0.04) \times 10^5$	$-7.23 (\pm 0.01)$	$-15.64 (\pm 0.01)$	$-8.41 (\pm 0.02)$	1.02
xylohexaose	$1.5 (\pm 0.07) \times 10^5$	$-7.07 (\pm 0.03)$	$-16.17 (\pm 0.02)$	$-9.10 (\pm 0.05)$	1.02
oat spelt xylan	$7.6 (\pm 0.60) \times 10^4$	$-6.65 (\pm 0.05)$	$-20.12 (\pm 0.07)$	$-13.47 (\pm 0.12)$	1.03
wheat arabinoxylan	$8.0 (\pm 1.00) \times 10^4$	$-6.68 (\pm 0.07)$	$-16.97 (\pm 0.09)$	$-10.29 (\pm 0.16)$	1.03
rye arabinoxylan	$1.1 (\pm 0.05) \times 10^5$	$-6.82 (\pm 0.03)$	$-17.42 (\pm 0.03)$	$-10.60 (\pm 0.06)$	1.02
barley β-glucan	$7.8 (\pm 0.70) \times 10^2$	$-3.94 (\pm 0.05)$	$-35.58 (\pm 0.10)$	$-31.64 (\pm 0.15)$	1.00[b]
cellohexaose	$7.8 (\pm 0.30) \times 10^2$	$-3.94 (\pm 0.02)$	$-15.37 (\pm 0.10)$	$-11.43 (\pm 0.12)$	1.00[b]

[a] NB indicates no binding detected. [b] Value was set at 1.00 as the binding was too weak to reach saturation during the titration, thus preventing accurate determination of n from the regressed calorimetry data.

Table 5: Molar Activities of Xyn10B Truncated Derivatives for Soluble and Insoluble Xylan

	enzyme activity (k_{cat}[a])	
enzyme	soluble xylan	insoluble xylan
X6a−CD	1289 ± 88[b]	25.2 ± 1.2
CD−X6b	1359 ± 95	68.5 ± 4.6
X6a−CD−X6b	1145 ± 56	72.9 ± 2.3
CD	1445 ± 87	28.2 ± 3.2
CD+X6b	1376 ± 45	29.8 ± 3.6

[a] k_{cat}, molecules of product s^{-1}. Final concentration of substrates was 5 and 10 g/L for soluble and insoluble xylan, respectively. [b] Results are means ± SEM ($n = 3$).

involving both hydrophobic interactions and hydrogen bonding, similar to that seen with CBM4 from *C. fimi* CenC (*45*).

Xylanase Activity of Xyn10B Derivatives. To evaluate the effect of X6a and X6b on the catalytic activity of Xyn10B, the molar activities of purified CD, X6a−CD, X6a−CD−X6b, and CD−X6b on soluble and insoluble xylan were determined. The results show that all enzymes have similar activities toward soluble xylan, suggesting that deletion of X6a and X6b from Xyn10B did not affect the function of the xylanases against soluble substrates (Table 4). In contrast, CD and X6a−CD showed a decreased capacity to hydrolyze insoluble xylan as compared with CD−X6b and X6a−CD−X6b, demonstrating that X6b enhances the catalytic properties of the xylanase against these insoluble substrates.

To assess whether X6b potentiates the activity of the Xyn10B CD by disrupting the structural integrity of insoluble xylan, the capacity of CD to hydrolyze the polymer when mixed with equimolar concentrations of X6b was assessed. The results (Table 5) show that X6b does not significantly enhance the catalytic activity of the CD in trans. These data suggest that X6b potentiates the catalytic activity of the xylanase by promoting a close interaction between the enzyme and its substrate.

Three-Dimensional Structure of X6b. Crystals of the X6b domain are in space group $P6_122$ unit-cell dimensions $a = b = 90.08$ and $c = 207.71$ Å. There are two molecules of X6b in the asymmetric unit. Using MAD data collected at the ESRF on beamline ID14-4, the structure was solved by the MAD method using a selenomethionyl form of the protein (see Materials and Methods for details). Phasing in SOLVE gave a map with a mean figure of merit of 0.56. This was improved through a combination of solvent flattening and noncrystallographic symmetry averaging using DM. The chain was initially traced into this 2.5 Å map. Later refinement used higher resolution (2.1 Å) data collected at the EMBL Hamburg Outstation.

The 2.1 Å data are 98% complete to 2.12 Å resolution. The R_{merge} ($\sum_{hkl}\sum_i|I_{hkli} - \langle I_{hkl}\rangle|/\sum_{hkl}\sum_i\langle I_{hkl}\rangle$) is 0.074, with a mean $I/\sigma I$ of 19.2 and a mean multiplicity of observation of 5.1 observations per reflection. In the outer resolution shell, the data are 97% complete, with an R_{merge} of 0.43, a mean $I/\sigma I$ of 3.4, and a mean multiplicity of 4.6. The final model of X6b consists of two molecules in the asymmetric unit each consisting of residues 5−160 and a single calcium ion. 304 water molecules with B values <65.0 Å^2 have also been modeled. The final crystallographic R value is 0.19, with an R_{free} of 0.25. The model shows deviations from stereochemical target values of 0.012 and 0.034 Å for the 1−2 bond and 1−3 angle distances, respectively. A paucity of crystal packing interactions is reflected in high mean temperature factors for the two molecules of X6b in the asymmetric unit of 41 and 40 Å^2 for the A and B molecules, respectively.

Residues 5−160 of the overexpressed X6b construct could be traced in both molecules. The structure forms a classical lectin-like β-jelly roll (Figure 4), predominantly consisting of four major antiparallel β-stands on both of the two faces. This fold displays most similarity to the *Bacillus thuringiensis* insecticidal toxin (PDB code 1ciy) and the lectin "wing" domain of the *Vibrio cholerae* neuraminidase (PDB code 1kit), with rms deviations of 4.0 (127 matched Cα atoms) and 3.5 Å (128 matched Cα atoms), respectively, when calculated using DALI (*50*). A single, structural calcium ion is found at the end of strand β-1. It displays typical hepta-coordination and is liganded to a single water molecule; the main-chain carbonyl oxygens of residues Thr 14, Lys 39, and Ser 43; the OE2 oxygen of Glu 16; and a bidentate interaction with both side-chain oxygens of Asp 149 (Figure 5).

X6b displays a shallow surface groove that runs along the whole of the concave face of the jelly roll, and we would imagine, by analogy with other lectin domains, that this forms the ligand-binding site. This groove is approximately 20 Å long, corresponding to the length of a xylotetraose moiety. This is entirely consistent with the observation that the protein exhibits maximum affinity for oligosaccharides containing four xylose units and supports the view that the cleft in X6b constitutes the ligand-binding site. The surface of the putative ligand-binding cleft contains three aromatic residues, Trp53, Tyr103, and Tyr134, which are all highly conserved in CBM family 22b (Figure 6). Surface aromatic residues have been implicated in ligand binding in several

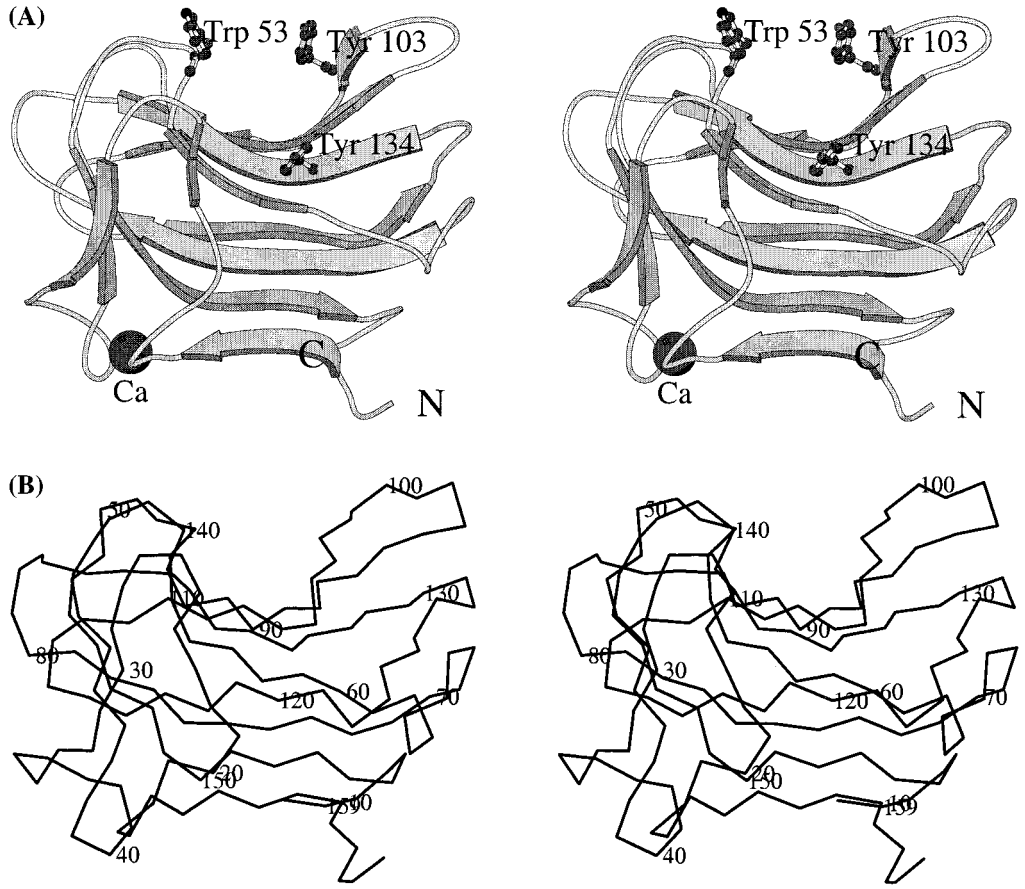

FIGURE 4: Structure of X6b. (A) Divergent (wall-eyed) representation of X6b. The three aromatic residues lining the putative binding cleft are shown in ball-and-stick representation, and the calcium ion is shown as a shaded sphere. (B) α-carbon backbone trace with every 10th residue numbered. This figure was drawn with the MOLSCRIPT program (*57*).

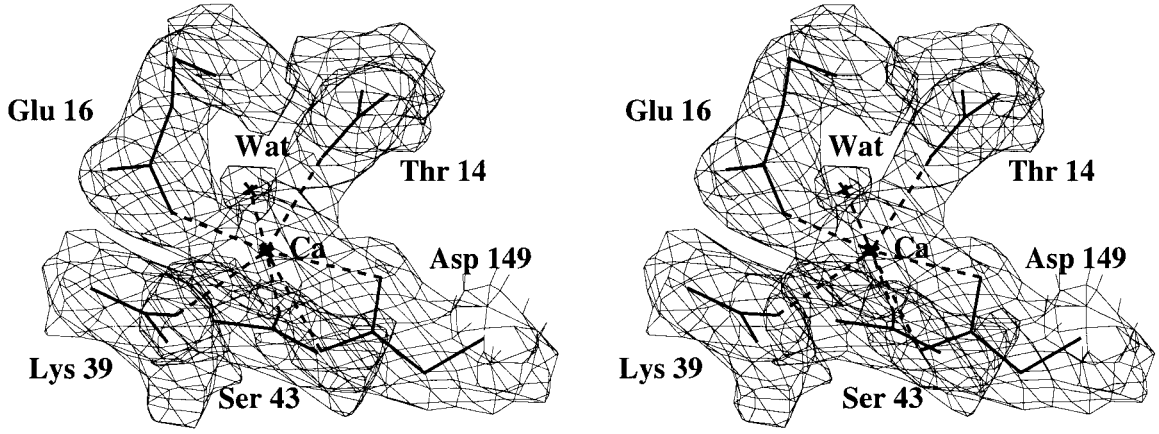

FIGURE 5: Electron density for the single structural calcium ion of X6b. Map shown is a maximum likelihood/σ_A weighted $2F_{obs} - F_{calc}$ synthesis contoured at a level of 0.38 electrons/Å^3 and is in divergent (wall-eyed) stereo.

different CBM families, via face-to-face stacking interactions with the sugar rings of the carbohydrate (*24*, *51*−*53*). It seems likely that since Trp 53, Tyr 103, and Tyr 134 are all exposed, highly conserved, and found in the groove that these residues are directly involved in ligand binding in X6b. Further support of the importance of these aromatic amino acid residues in ligand binding is illustrated by the observation that they are absent from the X6a module, which displays no capacity to bind xylan. Interestingly, X6b, in common with the XBM from *C. fimi* Xyl11A (24), can bind to the heavily substituted xylan chains of wheat and rye arabinoxylan. Studies on the structure of these carbohydrates have shown that they are substituted at both the 2′ and the 3′ positions and have arabinose:xylose ratios of 50:50 and 60:40, respectively. Thus, it is apparent that the binding site of X6b can accommodate a high degree of xylan backbone substitution.

The X6 modules had previously been described as thermostabilizing domains on account of their apparent ability to confer stability to the catalytic module. Truncation of Xyl10B leads to a decrease in the stability of the CM. Deletion of X6a is particularly destabilizing (30). It is difficult to say, however, that these modules confer true thermostability or whether the truncated proteins are simply

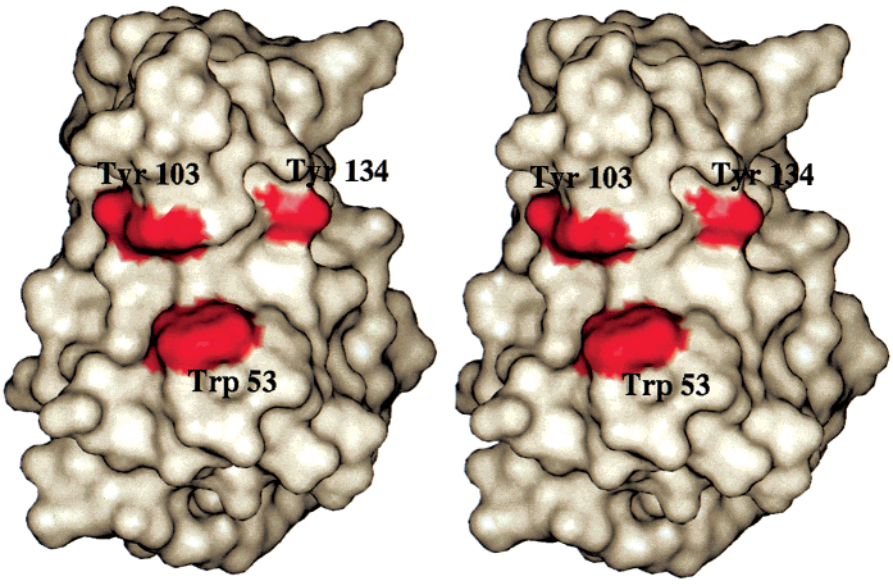

FIGURE 6: Surface representation of X6b. Three surface aromatics mentioned in the text are indicated.

less stable due to disruption of their structure and exposure of previously interacting surfaces. We propose that the primary function for many of the X6 modules will be xylan binding and that they play a role in the potentiation of catalytic activity on insoluble substrates. NCMs are frequently separated from the CMs by flexible, sometimes glycosylated, linker regions (*54*). A short linker sequence is present between X6b and the CM. It is likely that the flexibility provided by this linker is important for the potentiation of catalytic activity on insoluble xylans. X6a, however, is attached with no obvious linker sequence. The observation that removal of X6a leads to a reduction in thermostability of the truncated enzymes is consistent with a closer interaction of the CBM with the X6a as a direct result of this lack of a flexible linker sequence.

Close interaction between a CBM and the CM has been described in family 9 cellulases in which the CM is joined directly to $CBM3_C$ (*55*). This CBM does not bind tightly to cellulose but instead feeds single cellulose chains into the active site of the cellulase, facilitating the processive action of the enzyme against crystalline substrates (*56*). The role of X6a remains unclear. Deletion of X6a is certainly destabilizing, but we cannot rule out a role in binding, perhaps of an as yet undefined polymer. It may even play a role similar to CMB3c (*55*), but this would seem unlikely given the substituted nature of the substrate. Further experiments to investigate the function of X6a will be performed.

The data presented here show that the X6b module of *C. thermocellum* Xyn10B is a xylan-binding module. The 3D structure reveals a shallow cleft, which we propose, by virtue of sequence conservation and analogy with lectins, is the ligand-binding site. The length of this cleft is consistent with the observed binding affinities for xylooligosaccharides. These data strongly suggest that the X6 thermostabilizing domains are reclassified as a new family of carbohydrate-binding modules: family CBM22 in the noncatalytic module classification of Couthino and Henrissat (*14*). Further studies regarding the biological function of CBM family 22 proteins and their interaction with their CMs will be evaluated.

ACKNOWLEDGMENT

We would like to thank the staff of the EMBL Hamburg and the ESRF for provision of data collection facilities and the EU and ESRF for financial assistance for data collection. Collaboration between the Universities of York and Newcastle was assisted by the provision of a BBSRC Structural Biology Centre award to the York Structural Biology Laboratory.

REFERENCES

1. Brett, C., and Walden, K. (1990) *Physiology and Biochemistry of Plant Cell Walls*, Unwin and Hyman, London.
2. Gilbert, H. J., and Hazlewood, G. P. (1993) *J. Gen. Microbiol. 139*, 187−194.
3. Coughlan, M. P., and Hazlewood, G. P. (1993) *Biotechnol. Appl. Biochem. 17*, 259−289.
4. Biely, P. (1985) *Trends Biotechnol. 3*, 286−290.
5. Henrissat, B., and Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
6. Harris, E., Jenkins, J. A., Connerton, I., Cummings, N., Lo Leggio, L., Scott, M., Hazlewood, G. P., Laurie, J. I., Gilbert, H. J., and Pickersgill, R. W. (1994) *Structure 2*, 1107−1114.
7. Charnock, S. J., Lakey, H. J., Virden, R., Hughes, N., Sinnott, M. L., Hazlewood, G. P., Pickersgill, R., and Gilbert, H. J. (1997) *J. Biol. Chem. 272*, 2942−2951.
8. Charnock, S. J., Spurway, T. D., Xie, H., Beylot, H.-M., Virden, R., Warren, R. A. J., Hazlewood, G. P., and Gilbert, H. J. (1998) *J. Biol. Chem. 273*, 32187−32199.
9. Törrönen, A., Harkki, A., and Rouvinen, J. (1994) *EMBO J. 13*, 2493−2501.
10. Gilbert, H. J., Hazlewood, G. P., Laurie, J. I., Orpin, C. G., and Xue, G. P. (1992) *Mol. Microbiol. 6*, 2065−2072.
11. Hall, J., and Gilbert, H. J. (1988) *Mol. Gen. Genet. 213*, 112−117.
12. Tomme, P., Warren, R. A. J., and Gilkes, N. R. (1995) *Adv. Microb. Physiol. 37*, 1−81.
13. Ferreira, L. M. A., Durrant, A. J., Hall, J., Hazlewood, G. P., and Gilbert, H. J. (1990) *Biochem. J. 269*, 261−264.
14. Coutinho, P. M., and Henrissat, B. (1999) in *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B. and Svensson, B., Eds.) pp 3−12, The Royal Society of Chemistry, Cambridge.
15. Gilkes, N. R., Jervis, E., Henrissat, B., Tekant, B., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. (1992) *J. Biol. Chem. 267*, 6743−6749.

16. Fernandes, A. C., Fontes, C. M. G. A., Gilbert, H. J., Hazlewood, G. P., Fernandes, T. H., and Ferreira, L. M. A. (1999) *Biochem. J. 342*, 105−110.
17. Blaak, H., and Schrempf, H. (1995) *Eur. J. Biochem. 229*, 132−139.
18. 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.
19. Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Petterson, G., Knowles, J., and Gronenborn, A. M. (1989) *Biochemistry 28*, 7241−7257.
20. Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.
21. Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) *Biochemistry 36*, 16074−16086.
22. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
23. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
24. Simpson, P. J., Bolam, D. N., Hazlewood, G. P., Gilbert, H. J., and Willamson, M. P. (1999) *Structure 7*, 853−864.
25. Gill, J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Willamson, M. P., Hazlewood, G. P., and Gilbert, H. J. (1999) *Biochem. J. 342*, 473−480.
26. Bolam, D. N., Ciruela, A., Rixon, J. E., Hazlewood, G. P., and Gilbert, H. J. (1998) *Biochem. J. 331*, 775−781.
27. Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Theodorou, M. K., and Gilbert, H. J. (1996) *Biochem. J. 319*, 515−520.
28. Din, N., Gilkes, N. R., Tekant, B., Miller, R. C., Jr., and Warren, R. A. J. (1991) *Bio/Technology 9*, 1096−1099.
29. Din, N., Damude, H. G., Gilkes, N. R., Miller, R. C., Jr., and Warren, R. A. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11383−11387.
30. Fontes, C. M. G. A., Hazlewood, G. P., Morag, E., Hall, J., Hirst, B. H., and Gilbert, H. J. (1995) *Biochem. J. 307*, 151−158.
31. Lee, Y. E., Lowe, S. E., Henrissat, B., and Zeikus, J. G. (1993) *J. Bacteriol. 175*, 5890−5898.
32. Winterhalter, C., Heinrich, P., Candussio, A., Wich, G., and Liebl, W. (1995) *Mol. Microbiol. 15*, 431−444.
33. Clarke, J. H., Davidson, K., Gilbert, H. J., Fontes, C. M. G. A., and Hazlewood, G. P. (1996) *FEMS Microbiol. Lett. 139*, 27−35.
34. Zhang, J. X., Martin, J., and Flint, H. J. (1994) *Mol. Gen. Genet. 245*, 260−264.
35. Norrander, J., Kempe, T., and Messing, J. (1983) *Gene 26*, 101−106.
36. Studier, F. W., and Moffat, B. A. (1986) *J. Mol. Biol. 189*, 113−130.
37. Laemmli, U. K. (1970) *Nature (London) 227*, 680−685.
38. Sun, J. L., Sakka, K., Karita, S., Kimura, T., and Ohmiya, K. (1998) *J. Ferment. Bioeng. 85*, 63−68.
39. Miller, G. L. (1959) *Anal. Chem. 31*, 426−428.
40. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology: Macromolecular Crystallography, Part A* (Carter, C. W. J., and Sweet, R. M., Eds.) pp 307−326, Academic Press, London.
41. Ravelli, R. B. G., Sweet, R. M., Skinner, J. M., Duisenberg, A. J. M., and Kroon, J. (1997) *J. Appl. Crystallogr. 30*, 551−554.
42. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D55*, 849−861.
43. Cowtan, K. D., and Main, P. (1996) *Acta Crystallogr. D49*, 148−157.
44. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystrallogr. D53*, 240−255.
45. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
46. Sorimachi, K., Le Gal-Coëffet, M. F., Willamson, G., Archer, D. B., and Williamson, M. P. (1997) *Structure 5*, 647−661.
47. Chervenak, M. C., and Toone, E. J. (1995) *Biochemistry 34*, 5685−5695.
48. 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.
49. Livingstone, J. R., Spolar, R. S., and Record, M. T. (1991) *Biochemistry 30*, 4237−4244.
50. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
51. Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Orosz, L., and Gilbert, H. J. (1998) *FEBS Lett. 429*, 312−316.
52. Linder, M., Mattinen, M.-L., Kontteli, M., Lindberg, G., Ståhlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G., and Annila, A. (1996) *Protein Sci. 4*, 1056−1064.
53. Din, N., Forsythe, I. J., Burtnick, L. J., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. (1994) *Mol. Microbiol. 11*, 11383−11387.
54. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1991) *Microbiol. Rev. 55*, 303−315.
55. Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A (1997) *Nat. Struct. Biol. 4*, 810−817.
56. Irwin, D., Shin, D.-H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A., and Wilson, D. B. (1998) *J. Bacteriol. 180*, 1709−1714.
57. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI992821Q