

# Substrate Binding by the Catalytic Domain and Carbohydrate Binding Module of *Ruminococcus flavefaciens* FD-1 Xyloglucanase/Endoglucanase

Christopher D. Warner,<sup>†</sup> Gulden Camci-Unal,<sup>‡</sup> Nicola L. B. Pohl,<sup>†,‡,||</sup> Clark Ford,<sup>§</sup> and Peter J. Reilly<sup>\*,†</sup>

Departments of <sup>†</sup>Chemical and Biological Engineering, <sup>‡</sup>Chemistry, and <sup>§</sup>Food Science and Human Nutrition and the <sup>||</sup>Plant Sciences Institute, Iowa State University, Ames, Iowa 50011, United States

## Supporting Information

**ABSTRACT:** Binding and thermodynamic properties of a carbohydrate binding module (CBM) and a glycoside hydrolase family 44 xyloglucanase/endoglucanase catalytic domain (CD) from *Ruminococcus flavefaciens*, both when separate and when linked to each other, have been quantified when binding various  $\beta$ -1,4-linked glucans and xylans. The three constructs bind cellotetraose, cellopentaose, and cellohexaose with association constants that increase with chain length. The CBM does not bind xylo-tetraose, xylo-pentaose, or xylo-hexaose. The CD appears to bind carboxymethylcellulose (CMC) and xylan only weakly, while the CBM and the CD/CBM bind them much more strongly than they bind the celooligosaccharides. CMC is bound to a much greater degree than is xylan. Association constants for the celooligosaccharides are in the order  $\text{CBM} \ll \text{CD} < \text{CD/CBM}$ , while those on CMC and xylan are  $\text{CD} \ll \text{CBM} \ll \text{CD/CBM}$ . A synergistic effect was observed for the association constants of cellopentaose and cellohexaose with the CD/CBM when compared to the CD and CBM alone. Binding of all ligands by all three constructs is energetically favorable, enthalpy-driven, and subject to enthalpy–entropy compensation.

## INTRODUCTION

Carbohydrate binding modules (CBMs) are noncatalytic proteins that bind oligo- and polysaccharides. They are attached to or associated with carbohydrate-active catalytic domains (CDs), and they are classified based upon their binding properties as either type A (surface binding), type B (glycan chain binding), or type C (small sugar binding). CBMs have three primary roles: to increase local carbohydrate concentrations by the proximity effect, to target specific carbohydrates in complex structures such as plant cell walls, or to disrupt or modify carbohydrate structures.<sup>1,2</sup>

CBMs are classified based upon amino acid sequence similarity into families. In addition, it is assumed that all members of a family have roughly the same tertiary structure. At present the Carbohydrate-Active enZymes database (CAZy) lists over 60 CBM families along with over 100 unclassified CBM sequences.<sup>3</sup> Of those families having known tertiary structures, the majority have  $\beta$ -sandwich folds, while some families have  $\beta$ -trefoil, hevein, OB,  $\beta$ -barrel, and lectin-like folds.<sup>1,3</sup>

*Ruminococcus flavefaciens* is a species of ruminant bacteria highly active against plant cell walls. Two of its strains, *R. flavefaciens* 17 and *R. flavefaciens* FD-1, have been extensively studied.<sup>4–7</sup> Both strains produce cellulosomes, extracellular multiprotein complexes anchored to their cell membranes that break down cellulose and hemicelluloses. Two common features of the cellulosomes produced by each strain are the large number of different proteins, among them hydrolases, cohesins, dockerins, and scaffoldins, incorporated into them and the lack of CBMs attached to their scaffoldin domains,<sup>8</sup> unlike the structures of many cellulosomes. The latter factor makes the attachment of CBMs to CDs functionally necessary to digest insoluble polysaccharides.<sup>9–14</sup>

*R. flavefaciens* 17 endoglucanase B (EndB) has a 150-residue CBM attached to the C-terminal end of its CD. This CBM binds Avicel but is otherwise relatively unstudied.<sup>15</sup> It is GenBank accession CAC83072.1 and UniProt accession Q934F9, but it is unclassified in the CAZy database. *R. flavefaciens* FD-1 produces a xyloglucanase/endoglucanase (XG/EG) (CelB) highly similar to *R. flavefaciens* 17 EndB. The CD sequence of strain FD-1 was published many years ago;<sup>7</sup> its CBM sequence was published recently and does not appear separately in the CAZy database.<sup>8</sup> The CDs of both strains are part of glycoside hydrolase family 44 (GH44).<sup>3,15</sup>

No tertiary structure of the FD-1 CBM has been published; however, a preliminary NMR study suggests that it is a  $\beta$ -sandwich like many other CBMs.<sup>16</sup> Four tertiary structures of GH44 CDs, from *Clostridium thermocellum* F1 (PDB 2E4T),<sup>17</sup> an uncultured bacterium (PDB 3FW6),<sup>18</sup> *Clostridium acetobutylicum* ATCC 824 (PDB 3IK2),<sup>19</sup> and *Paenibacillus polymyxa* GS01 (PDB 2YIH),<sup>20</sup> are available. All have  $(\beta, \alpha)_8$  barrel and Greek key domains. A multiple sequence alignment of the amino acid sequences of the FD-1 CD and the four GH44 members with crystal structures is shown in Supporting Information Figure 1.

Both *R. flavefaciens* FD-1 XG/EG and a GH44 EG from *C. acetobutylicum* ATCC 824 are active on cellotetraose, cellopentaose, cellohexaose, carboxymethylcellulose (CMC), birchwood and larchwood xylan, lichenan, and Avicel, but not

**Special Issue:** L. T. Fan Festschrift

**Received:** December 20, 2011

**Revised:** April 20, 2012

**Accepted:** April 24, 2012

Table 1. Thermodynamic Parameters for CBM–Ligand Binding

	$K_w$ , M <sup>-1</sup>	$\Delta H$ , kJ/mol	$\Delta S$ , J/mol-K	$\Delta G$ , kJ/mol	$n$
cellotetraose <sup>a</sup>	635 ± 175	-0.899 ± 0.088	50.6	-16.0	3.53
cellopentaose <sup>a</sup>	972 ± 97	-4.81 ± 0.50	41.1	-17.1	4.23
cellohexaose <sup>a</sup>	1620 ± 80	-13.6 ± 1.1	16.0	-18.3	1.40
CMC <sup>a</sup>	63,600 ± 9200	-40.0 ± 3.0	-42.2	-27.4	0.871
xylan <sup>a</sup>	16,800 ± 2000	-45.6 ± 3.8	-71.9	-24.1	1.89

<sup>a</sup>Single-site model.

on cellobiose, cellotriose, mannan, or pullulan.<sup>19,21</sup> Addition of the *R. flavefaciens* FD-1 XG/EG CBM to its CD increases catalytic efficiencies on both CMC and birchwood xylan. Both family members show evidence of asymmetric hydrolysis and disproportionation of reaction products upon incubation with cellopentaose and cellohexaose.<sup>19,21,22</sup> Cellotetraose is digested to cellotriose and glucose, with no cellobiose observed. Digestion of cellopentaose produces cellotetraose, cellotriose, cellobiose, and glucose. The distribution of cellotriose, cellobiose, and glucose varies depending on the family member and/or presence of the CBM. In each case cellotetraose was present in larger amounts than glucose, its coupled hydrolysis reaction product. Cellohexaose digestion yields much more cellotetraose than cellobiose and some glucose and cellotriose, along with a low and transient amount of cellopentaose.

A reaction model consisting of hydrolysis and transglycosylation reactions was developed and optimized to account for asymmetric hydrolysis and disproportionation of reaction products.<sup>22</sup> The model fit the cellotetraose, cellopentaose, and cellohexaose digest data sets of *C. acetobutylicum* ATCC 824 GH44 EG (with linked CD and CBM) and *R. flavefaciens* FD-1 XG/EG CD and CD/CBM. Presence of transglycosylation products longer than cellopentaose and cellohexaose at levels below their detection limits was predicted with the model, and this was consistent with their absence in capillary electrophoresis chromatographs.

This article presents the ligand specificities, binding constants, and thermodynamic properties of the CBM, CD, and CD/CBM produced by *R. flavefaciens* FD-1 XG/EG, the first such investigation of these domains.

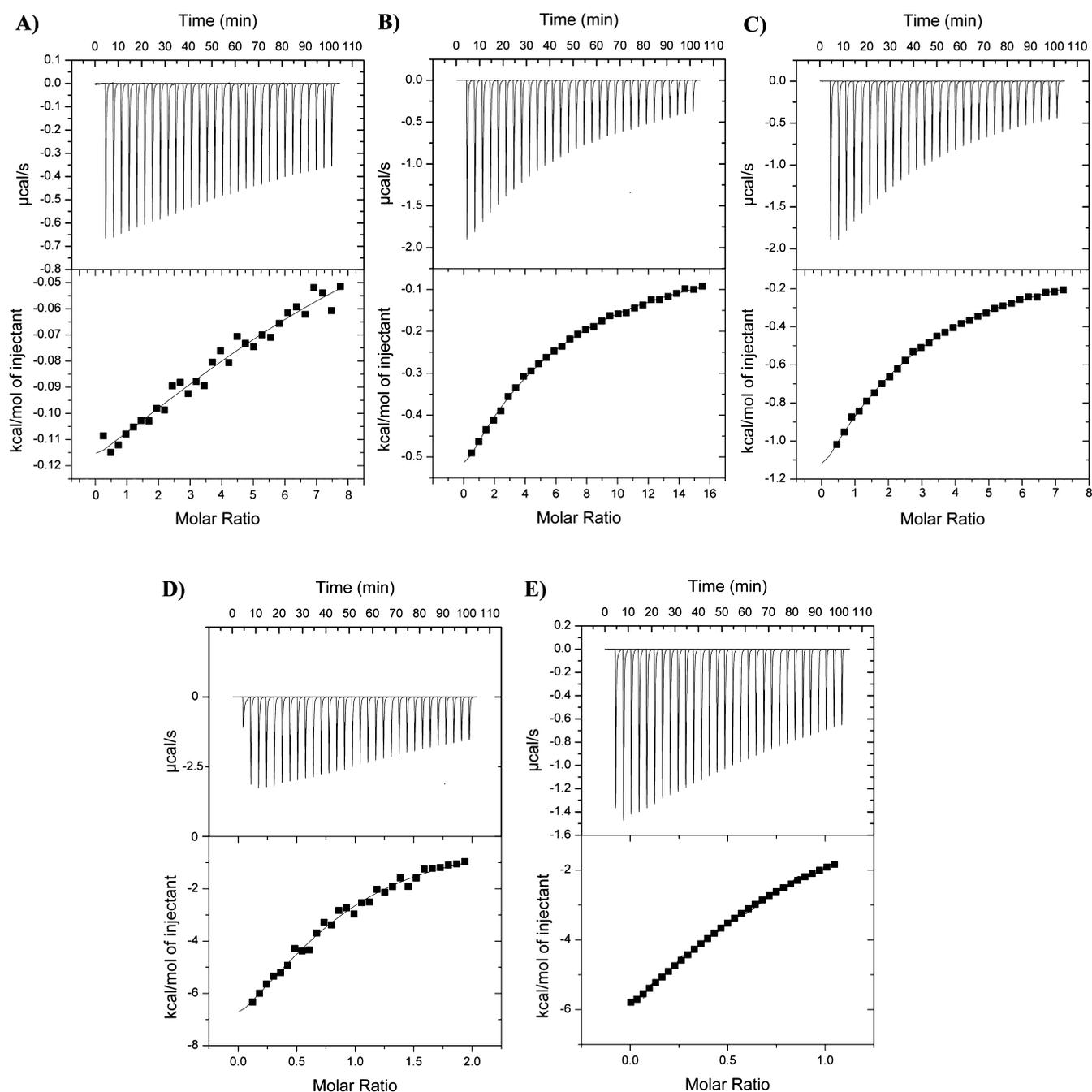
## EXPERIMENTAL SECTION

**Gene Construction.** Construction of genes encoding the CD and CD/CBM was described by Warner et al.<sup>21</sup> The pBAW101 plasmid, donated by Professor Bryan White of the University of Illinois at Urbana–Champaign, was used to amplify the gene fragment encoding the CBM using the polymerase chain reaction (PCR) with *Taq* polymerase (forward primer 5'-AATACATATGGCAGGAGGTTTTGATATG-3', reverse primer 5'-TCGACATATGCTACTCAGTCTTCTCGGTTAC-3'). The PCR products were cloned into pGEM-T Easy vectors (Promega, Madison, WI) and sequenced at the Iowa State University DNA Facility. The gene fragments were then ligated into the NdeI site of pET-14b (Novagen, San Diego, CA) and transformed into *E. coli* BL-21 (DE3). Stratagene's (San Diego, CA) QuikChange II site-directed mutagenesis kit was used to mutate the catalytic acid of CelB (forward primer 5'-CAGCCTTGATAAC-CAGCCTGTTCTCTGGAACG-3', reverse primer 5'-CGTTCAGAGAACAGGCTGGTTATCAAGGCTG-3') and/or to insert a stop codon to remove the CBM from the CD (forward primer 5'-GGTAACCGAGAAGACTGAGTAATT-CAAGGATCCTTCTTC-3', reverse primer 5'-GAAGAAG-

GATCCTTGAATTACTCAGTCTTCTCGGTTACC-3'). The mutated plasmids were transformed into *E. coli* XL-Blue and sequenced to confirm the mutated DNA sequence. The plasmid was then transformed into *E. coli* BL-21 (DE3) using electroporation. This resulted in three clones, each of which was used to produce a specific construct: CelB E210Q (the CelB CD with mutated catalytic acid and CBM); CelB E210Q/E551- (the CelB CD with mutated catalytic acid but no CBM); and the isolated CBM.

**Protein Production and Purification.** Production and purification of the CD and CD/CBM were described by Warner et al.<sup>21</sup> The isolated CBM was produced and purified similarly.

**Isothermal Calorimetry.** Isothermal calorimetry (ITC) was performed using a MicroCal (GE Healthcare, Northampton, MA) VP-ITC calorimeter. Cello- and xylooligosaccharides were purchased from Associates of Cape Cod (East Falmouth, MA), while low-viscosity CMC [(cellulose derivatized mainly with 2-O- and 6-O-linked carboxyl groups, but also with 3-O- and terminal 1-O- and 4-O-linked carboxyl groups, averaging 0.6 to 0.95 groups per glucopyranosyl residue) (product C-5678, lot 065K0111) and larchwood xylan [ $\beta$ -D-xylopyranosyl-(1→4)]<sub>n</sub>- $\beta$ -D-xylose,  $n$  = high, with significant branching initiated and terminated by other sugar residues) (product X-3875, lot 125C-00582) were purchased from Sigma (St. Louis, MO). Proteins and carbohydrate ligands were prepared in 50 mM HEPES, pH 6.8, collected from the ultrafiltration filtrate during concentration to minimize heat of dilution effects, and degassed before use. The His-tags were not removed prior to data collection, as previous ITC experiments with these constructs showed the tags had no significant impact on the model fits (data not shown). The ITC sample cell, 1.4288 mL, was loaded with the CD (0.05 mM), CBM (0.03–0.161 mM), or CD/CBM (0.07 mM), and stirred at 310 rpm and 25 °C. A 300- $\mu$ L syringe was loaded with ligands (2.5–10 mM cellooligosaccharides, 2–5 mM xylooligosaccharides, 0.333 mM CMC, or 0.45 mM xylan) and a series of 31 injections, 1  $\mu$ L for the first injection and 10  $\mu$ L for the rest, was made at 200-s intervals. Controls were run using the sample experimental parameters as above but with the ligands injected into a sample cell containing 50 mM HEPES, pH 6.8, of ultrafiltration filtrate. Their results were then used to subtract heat of dilution effects from the experimental data. Origin software (7.0 OriginLab Corp., Northampton, MA) was used to fit the data to either a one-site binding model or a two-site sequential model, depending on the protein–ligand pair. The models were converged using 100 Simplex iterations. The molecular weight of larchwood xylan, 17.5 kDa, was calculated based upon its typical degree of polymerization of 133,<sup>23</sup> and the molecular weight of low-viscosity CMC, 90 kDa, was provided by the manufacturer.



**Figure 1.** ITC binding plot for the addition at 25 °C of (A) 10 mM cellotetraose to 0.3 mM CBM, (B) 10 mM cellopentaose to 0.15 mM CBM, (C) 5 mM cellohexaose to 0.161 mM CBM, (D) 0.33 mM CMC to 0.1 mM CBM, and (E) 0.45 mM xylan to 0.1 mM CBM. The upper plots show the raw heats of binding and the lower plots show the integrated heats of binding excluding dilution effects.

## RESULTS AND DISCUSSION

**Gene Sequencing and Amino Acid Alignment.** Gene *celB* codes for an 800-residue protein that is 74% identical to *R. flavefaciens* 17 EndB.<sup>8,15</sup> The CelB protein has a domain structure similar to that of *R. flavefaciens* 17 EndB, as it contains regions that align well with the latter's CD, CBM, and dockerin domains. Thus *R. flavefaciens* FD-1 CelB contains the second known occurrence of this CBM.

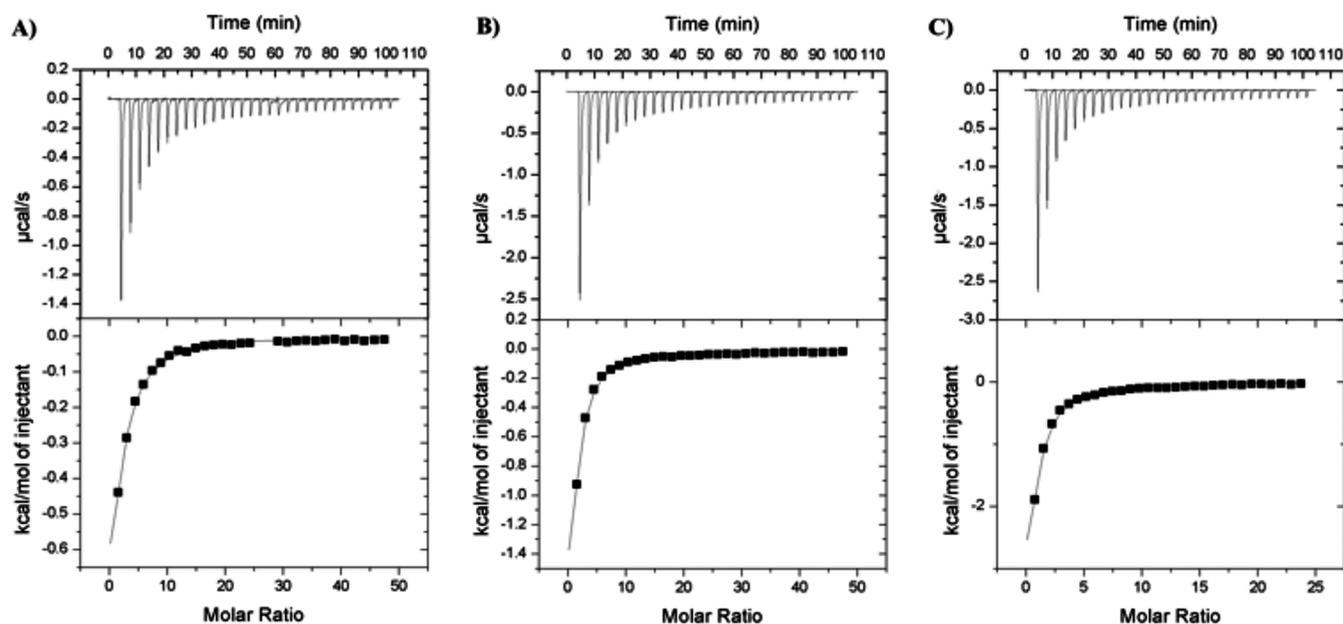
**Ligand Binding by the CBM.** The CBM binds cellotetraose ( $K_a = 635 \pm 175 \text{ M}^{-1}$ , the second value being the standard error), cellopentaose ( $K_a = 972 \pm 97 \text{ M}^{-1}$ ) and cellohexaose ( $K_a = 1620 \pm 80 \text{ M}^{-1}$ ), the values calculated using

a one-site binding model (Table 1). No significant binding of xylotetraose, xylopentaose, or xylohexaose occurs. Binding of CMC and xylan ( $K_a = 63\,600 \pm 9200 \text{ M}^{-1}$  and  $K_a = 16\,800 \pm 2000 \text{ M}^{-1}$ , respectively) is much stronger than that of the cellooligosaccharides. Stronger binding of longer ligands is likely due to multiple CBMs binding the same molecule of the longer ligands, or to their being bound by more subsites, each holding one monosaccharide residue, of a single CBM. Figure 1 shows the raw and integrated heats of binding excluding dilution effects for the CBM with the five bound ligands.

For each of these titrations  $\Delta G$ , the Gibbs free energy, is negative, showing that binding is energetically favorable. The entropy of binding,  $\Delta S$ , is strongly positive for cellotetraose and

Table 2. Thermodynamic Parameters for CD–Ligand Binding

	$K_a$ , $M^{-1}$	$\Delta H$ , kJ/mol	$\Delta S$ , J/mol-K	$\Delta G$ , kJ/mol	$n$
cellotetraose <sup>a</sup>	6960 ± 810	-6.15 ± 1.30	53.1	-22.0	1.86
cellopentaose <sup>a</sup>	8240 ± 2880	-12.8 ± 8.2	32.2	-22.3	1.05
cellohexaose <sup>a</sup>	12,800 ± 2100	-47.3 ± 29.3	-79.5	-23.5	0.408

<sup>a</sup>Single-site model.

**Figure 2.** ITC binding plot for the addition at 25 °C of (A) 10 mM cellotetraose to 0.049 mM CD, (B) 10 mM cellopentaose to 0.049 mM CD, and (C) 2.5 mM cellohexaose to 0.05 mM CD. The upper plots show the raw heats of binding and the lower plots show the integrated heats of binding excluding dilution effects.

cellopentaose, but less so for cellohexaose. The enthalpy of binding,  $\Delta H$ , is negative for each celooligosaccharide and becomes more negative with increasing chain length. These two observations suggest ligand binding is enthalpy-driven and undergoes enthalpy–entropy compensation. Binding of larger ligands is mainly enthalpy-driven, as  $\Delta S$  is negative for CMC and xylan, while their  $\Delta H$  values are strongly negative. The stoichiometry of the binding reactions,  $n$ , is close to 1 mol ligand:1 mol CBM for cellohexaose and CMC, as expected. The binding of xylan by the CBM occurs at almost a 2:1 ratio, and the binding of cellotetraose and cellopentaose by the CBM occurs at a nearly a 4:1 ratio. Interestingly,  $n$ -values were close to or above unity for binding of the polysaccharides CMC and xylan to the CBM. The CBM likely binds internally to the polysaccharide chain, which would result in  $n$ -values less than unity, as the CBM would have multiple binding sites per polysaccharide molecule. This may suggest that nonspecific binding occurs in the case of xylan, or that the CBM binding pocket has many subsites in the case of cellotetraose and cellopentaose.

Although the CBM binds celooligosaccharides, its binding constants are low when compared to those of other CBMs that bind celooligosaccharides.<sup>1,24</sup> Binding constants of the order of  $5 \times 10^4 M^{-1}$  were measured in those works rather than the constants of  $\sim 10^3 M^{-1}$  found here. Constants of the same order of magnitude as the CBM were measured by Boraston et al. for the individual CBMs of a tandem of three CBMs.<sup>25</sup>

**Ligand Binding by Inactive Mutants of the CD.** Titrations of the CD and CD/CBM were performed with

cellotetraose, cellopentaose, and cellohexaose to determine if a synergistic effect occurs when the CD and CBM are linked. The catalytic acid was mutated to a glutamine residue to prevent any ligand hydrolysis, which would result in a heat of reaction obscuring the binding data as well as a change in the character of the ligand.

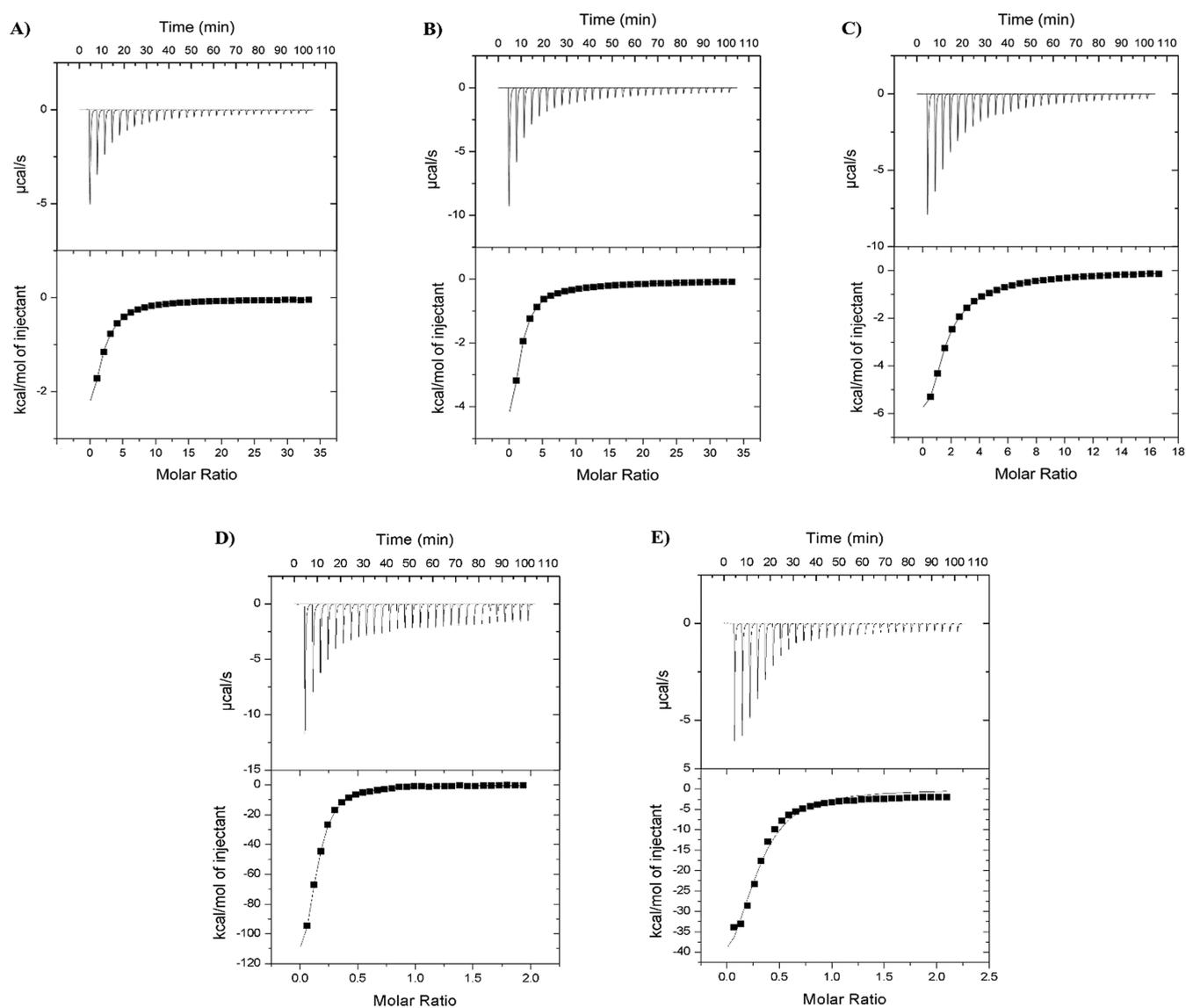
The CD binds cellotetraose ( $K_a = 6960 \pm 810 M^{-1}$ ), cellopentaose ( $K_a = 8240 \pm 2880 M^{-1}$ ), and cellohexaose ( $K_a = 12\,800 \pm 2100 M^{-1}$ ) using a one-site binding model (Table 2). Figure 2 shows the raw and integrated heats of binding excluding dilution effects for the CD with the three celooligosaccharides tested. Based upon the binding stoichiometry, it appears that the CD binding pocket can accommodate two molecules of cellotetraose per molecule of enzyme, but only one molecule of cellopentaose or cellohexaose. This suggests that the glycon and aglycon binding sites are each at least four subsites long. This is consistent with the subsite structure found in the crystal structures of three GH44 family members, all of which have more than eight subsites.<sup>17–19</sup>

Titrations of the CD with CMC and xylan were inconclusive because the data showed little curvature up to molar ratios of two molecules of substrate to one molecule of CD, where the solubility of the substrates reached saturation. This led to binding constants nearly equivalent to their standard deviations, suggesting that binding was weak. Although the calorimetric data proved inconclusive,  $k_{cat}/K_M$  data for both substrates with the CD alone were positive, showing that they could be hydrolyzed in the absence of the CBM.<sup>21</sup>

Table 3. Thermodynamic Parameters for CD/CBM–Ligand Binding

	$K_d, M^{-1}$	$\Delta H, \text{kJ/mol}$	$\Delta S, \text{J/mol-K}$	$\Delta G, \text{kJ/mol}$	$n$
cellotetraose <sup>a</sup>	$7140 \pm 220$	$-28.1 \pm 0.5$	-20.6	-22.0	n/a
	$173 \pm 55$	$-24.8 \pm 5.1$	-40.6	-12.8	n/a
cellopentaose <sup>a</sup>	$14,600 \pm 900$	$-35.4 \pm 1.0$	-39.0	-23.8	n/a
	$638 \pm 76$	$-43.9 \pm 1.7$	-93.7	-16.0	n/a
cellohexaose <sup>a</sup>	$22,900 \pm 3800$	$-41.6 \pm 2.8$	-56.0	-24.9	n/a
	$2320 \pm 290$	$-38.4 \pm 2.7$	-64.4	-19.2	n/a
CMC <sup>b</sup>	$452,000 \pm 15,000$	$-657 \pm 13$	-2100	-32.3	0.125
xylan <sup>b</sup>	$184,000 \pm 36,000$	$-221 \pm 21$	-640	-30.2	0.255

<sup>a</sup>Two-site sequential model. First and second rows are first and second sites, respectively. <sup>b</sup>Single-site model.



**Figure 3.** ITC binding plot for the addition at 25 °C of (A) 10 mM cellotetraose to 0.07 mM CD/CBM, (B) 10 mM cellopentaose to 0.07 mM CD/CBM, (C) 5 mM cellohexaose to 0.07 mM CD/CBM, (D) 0.33 mM CMC to 0.05 mM CD/CBM, and (E) 0.45 mM xylan to 0.05 mM CD/CBM. The upper plots show the raw heats of binding and the lower plots show the integrated heats of binding excluding dilution effects.

#### Ligand Binding by Inactive Mutants of the CD/CBM.

Binding curves from titrations of the CD/CBM with cellotetraose, cellopentaose, and cellohexaose were fitted better with a two-site sequential binding model than with either a one-site or two-site binding model, giving binding constants for cellotetraose of  $7140 \pm 220 M^{-1}$  and  $173 \pm 55 M^{-1}$ , for cellopentaose of  $14\,600 \pm 900 M^{-1}$  and  $638 \pm 76 M^{-1}$ , and for

cellohexaose of  $22\,900 \pm 3800 M^{-1}$  and  $2320 \pm 290 M^{-1}$  for the first and second sites, respectively (Table 3). The first binding constants are consistent with, although larger than, the binding constants for the same cellooligosaccharides bound by the CD (first binding site), while the much smaller second binding constants are roughly consistent with those of the CBM binding the same ligands (second binding site).

The CD/CBM was also titrated with CMC and xylan. Binding isotherms were best fitted with a one-site binding model, suggesting that these polysaccharides span both the CD and CBM binding sites. Binding constants for CMC and xylan are  $452\,000 \pm 15\,000\text{ M}^{-1}$  and  $109\,000 \pm 24\,000\text{ M}^{-1}$ , respectively (Table 3). The binding stoichiometry indicates that each CD/CBM molecule binds much less than one molecule of CMC or xylan. This is of particular note in light of the  $n$ -value fit from the binding data of xylan to the CBM, which was less than unity. Figure 3 shows the raw and integrated heats of binding excluding dilution effects for the CD/CBM with the five ligands tested.

Interestingly, the CD/CBM shows two different binding motifs depending on the length of the substrate. The implication of the two-site sequential binding mechanism for the CD/CBM binding cellooligosaccharides is that the distance between the binding site on the CBM and the active site on the CD is larger than the length of a cellohexaose molecule, whereas a single molecule with a high degree of polymerization, CMC or xylan, is so long that it binds to both the CBM and the CD active site as a single binding site. The stoichiometry of the binding models for CMC and xylan binding to the CD/CBM is less than unity. This indicates that it is improbable for a single molecule of CMC or xylan to bind the CBM and the CD active site as if they were two distinct binding sites because the substrates are so long that multiple CD/CBM molecules bind a molecule of CMC or xylan.

Analysis of the  $K_a$  values of the three constructs for cellooligosaccharide binding shows a synergistic effect of binding on cellopentaose and cellohexaose when the CD and CBM are both present, in that the sums of  $K_a$  values for binding by separate CBM and CD domains are much less than those for the combined CD/CBM. This is the first demonstration of a synergistic effect on  $K_a$  when a CBM is bound to a GH. No synergism was observed when the CD and CBM bound cellotetraose. The synergistic effect of combining of CD and CBM on binding of CMC and xylan is left to speculation, as their CD binding data were inconclusive. The increases in catalytic activity ( $k_{cat}/K_M$ ) on CMC and xylan for the CD/CBM compared to the CD (but lower catalytic activity on xyloglucan and lower rates on cellopentaose and cellohexaose)<sup>21</sup> may result from a synergistic effect of binding, as CBMs can have such an effect on activity of GHs.<sup>26,27</sup> Furthermore, synergistic effects on  $K_a$  have been observed when CBMs are present in tandem.<sup>25,28,29</sup>

As with binding by the CBM,  $\Delta G$  values for binding of all ligands by the CD and CD/CBM are strongly negative. Binding of cellooligosaccharides by the CD gives progressively more negative  $\Delta H$  values with increasing chain length, with positive  $\Delta S$  values for cellotetraose and cellopentaose and negative values for cellohexaose. All  $\Delta H$  and  $\Delta S$  values for ligand binding by the CD/CBM are negative, becoming much more so with CMC and xylan. This again leads to the conclusion that all binding with all constructs is enthalpy-driven with enthalpy–entropy compensation.

## CONCLUSIONS

The binding and thermodynamic properties of CBM, CD, and CD/CBM constructs have been quantified for a variety of ligands. All three constructs bind cellotetraose, cellopentaose, and cellohexaose with increasing affinity as chain length increases. The CBM and CD/CBM have higher association constants for CMC than for xylan, with those for CMC being

greater by more than an order of magnitude than those for cellohexaose. Although the CBM binds xylan, no binding of xylohexaose, xylopentaose, or xylohexaose occurs. The presence of the CBM on the C-terminus of the CD yields a synergistic effect on the association constants of cellopentaose and cellohexaose. Binding of all ligands with the three constructs results in energetically favorable negative values of  $\Delta G$ , caused by a mixture of enthalpic and entropic effects.

The mechanistic cause of the observed synergism is unknown. It may arise from an increase in the local concentration of substrate near both binding sites due to the presence of a second binding site (CBM) in near proximity to the first (CD). There could be an orientation effect resulting from slippage of a cellohexaose molecule from the nearby CBM into the active site in a more preferable orientation or confirmation for binding. It is also possible the synergistic mechanism results from allosteric effects of substrate binding to the CBM, impacting active site structure.

The synergistic effects, preference for cellooligosaccharides, and increase of the association constants with increased degree of polymerization provide clues to the role of this xyloglucanase/endoglucanase within the larger *Ruminococcus flavefaciens* FD-1 cellulosome complex. The unusual lack of CBMs in the scaffoldin architecture makes the presence of the member of a novel CBM family attached to the CD relatively important. The ability of the CD/CBM construct to bind and digest different types of polysaccharides allows the cellulosomes to digest polysaccharides better than shorter oligosaccharides, providing a variety of shorter substrates for the other cellulosome components to digest. Yet the ability of the CD/CBM to digest cellotetraose through cellohexaose likely gives it a more defined role within the cellulosome complex as the longer diverse set of polysaccharides dwindles in favor of shorter substrates. The synergistic nature of the CD/CBM complex is likely to cause an increase in  $k_{cat}$  against cellooligosaccharides and presumably against longer polysaccharides.

## ASSOCIATED CONTENT

### Supporting Information

SI Figure 1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [reilly@iastate.edu](mailto:reilly@iastate.edu). Tel: +1-515-294-5968. Fax: +1-515-294-2689.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This project was funded by the National Research Initiative of the U.S. Department of Agriculture Cooperative State Research, Education and Extension Service through contract 2007-35504-18252. We are very grateful for this assistance. We also thank Professor Bryan White for his donation of the pBAW101 plasmid.

## REFERENCES

- (1) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochem. J.* **2004**, *382*, 769–781.

- (2) Shoseyov, O.; Shani, Z.; Levy, I. Carbohydrate binding modules: Biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 283–295.
- (3) Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238.
- (4) Doerner, K. C.; White, B. A. Assessment of the endo-1,4- $\beta$ -glucanase components of *Ruminococcus flavefaciens* FD-1. *Appl. Environ. Microbiol.* **1990**, *56*, 1844–1850.
- (5) Doerner, K. C.; Howard, G. T.; Mackie, R. I.; White, B. A.  $\beta$ -Glucanase expression by *Ruminococcus flavefaciens* FD-1. *FEMS Microbiol. Lett.* **1992**, *93*, 147–153.
- (6) Flint, H. J.; Martin, J.; McPherson, C. A.; Daniel, A. S.; Zhang, J. X. A bifunctional enzyme, with separate xylanase and  $\beta(1,3-1,4)$ -glucanase domains, encoded by the *xymD* gene of *Ruminococcus flavefaciens*. *J. Bacteriol.* **1993**, *175*, 2943–2951.
- (7) Vercoe, P. E.; Finks, J. L.; White, B. A. DNA sequence and transcriptional characterization of a  $\beta$ -glucanase gene (*celB*) from *Ruminococcus flavefaciens* FD-1. *Can. J. Microbiol.* **1995**, *41*, 869–876.
- (8) Berg Miller, M. E.; Antonopoulos, D. A.; Rincon, M. T.; Band, M.; Bari, A.; Akraiko, T.; Hernandez, A.; Thimmapuram, J.; Henrissat, B.; Coutinho, P. M.; Borovok, I.; Jindou, S.; Lamed, R.; Flint, H. J.; Bayer, E. A.; White, B. A. Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1. *PLoS One* **2009**, *4*, e6650.
- (9) Ding, S. Y.; Rincon, M. T.; Lamed, R.; Martin, J. C.; McCrae, S. I.; Aurilia, V.; Shoham, Y.; Bayer, E. A.; Flint, H. J. Cellulosomal scaffoldin-like proteins from *Ruminococcus flavefaciens*. *J. Bacteriol.* **2001**, *183*, 1945–1953.
- (10) Rincon, M. T.; Ding, S. Y.; McCrae, S. I.; Martin, J. C.; Aurilia, V.; Lamed, R.; Shoham, Y.; Bayer, E. A.; Flint, H. J. Novel organization and divergent dockerin specificities in the cellulosome system of *Ruminococcus flavefaciens*. *J. Bacteriol.* **2003**, *185*, 703–713.
- (11) Rincon, M. T.; Martin, J. C.; Aurilia, V.; McCrae, S. I.; Rucklidge, G. J.; Reid, M. D.; Bayer, E. A.; Lamed, R.; Flint, H. J. ScaC, an adaptor protein carrying a novel cohesin that expands the dockerin-binding repertoire of the *Ruminococcus flavefaciens* 17 cellulosome. *J. Bacteriol.* **2004**, *186*, 2576–2585.
- (12) Bayer, E. A.; Belaich, J. P.; Shoham, Y.; Lamed, R. The cellulosomes: Multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* **2004**, *58*, 521–554.
- (13) Jindou, S.; Borovok, I.; Rincon, M. T.; Flint, H. J.; Antonopoulos, D. A.; Berg, M. E.; White, B. A.; Bayer, E. A.; Lamed, R. Conservation and divergence in cellulosome architecture between two strains of *Ruminococcus flavefaciens*. *J. Bacteriol.* **2006**, *188*, 7971–7976.
- (14) Bayer, E. A.; Lamed, R.; White, B. A.; Flint, H. J. From cellulosomes to cellulosomes. *Chem. Rec.* **2008**, *8*, 364–377.
- (15) Rincon, M. T.; McCrae, S. I.; Kirby, J.; Scott, K. P.; Flint, H. J. EndB, a multidomain family 44 cellulase from *Ruminococcus flavefaciens* 17, binds to cellulose via a novel cellulose-binding module and to another *R. flavefaciens* protein via a dockerin domain. *Appl. Environ. Microbiol.* **2001**, *67*, 4426–4431.
- (16) Warner, C. D. Structures, kinetics, and modeling of glycoside hydrolase family 44 endoglucanases from *Clostridium acetobutylicum* and *Ruminococcus flavefaciens* FD-1. Ph. D. Dissertation, Department of Chemical and Biological Engineering, Iowa State University, Ames, IA. 2010.
- (17) Kitago, Y.; Karita, S.; Watanabe, N.; Kamiya, M.; Aizawa, T.; Sakka, K.; Tanaka, I. Crystal structure of Cel44A, a glycoside hydrolase family 44 endoglucanase from *Clostridium thermocellum*. *J. Biol. Chem.* **2007**, *282*, 35703–35711.
- (18) Nam, K. H.; Kim, S. J.; Hwang, K. Y. Crystal structure of CelM2, a bifunctional glucanase-xylanase protein from a metagenome library. *Biochem. Biophys. Res. Commun.* **2009**, *383*, 183–186.
- (19) Warner, C. D.; Hoy, J. A.; Shilling, T. C.; Linnen, M. J.; Ginder, N. D.; Ford, C. F.; Honzatko, R. B.; Reilly, P. J. Tertiary structure and characterization of a glycoside hydrolase family 44 endoglucanase from *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* **2010**, *76*, 338–346.
- (20) Ariza, A.; Eklof, J. M.; Spadiut, O.; Offen, W. A.; Roberts, S. M.; Besenmatter, W.; Friis, E. P.; Skjot, M.; Wilson, K. S.; Brumer, H.; Davies, G. Structure of a *Paenibacillus polymyxa* xyloglucanase from GH family 44 with xyloglucan. 2011. Protein Data Bank accession 2YIH.
- (21) Warner, C. D.; Go, R. M.; Garcia-Salinas, C.; Ford, C.; Reilly, P. J. Kinetic characterization of a glycoside hydrolase family 44 xyloglucanase/endoglucanase from *Ruminococcus flavefaciens* FD-1. *Enzyme Microb. Technol.* **2011**, *48*, 27–32.
- (22) Warner, C. D.; Ford, C.; Reilly, P. J. Kinetics and modeling of disproportionation reactions catalyzed by *Clostridium acetobutylicum* and *Ruminococcus flavefaciens* endoglucanases. *Biotechnol. Bioeng.* **2012**, submitted for publication.
- (23) Panbangred, W.; Shinmyo, A.; Kinoshita, S.; Okada, H. Purification and properties of endoxylanase produced by *Bacillus pumilus*. *Agric. Biol. Chem.* **1983**, *47*, 957–963.
- (24) Najmudin, S.; Guerreiro, C.; Carvalho, A. L.; Prates, J. A. M.; Correia, M. A. S.; Alves, V. D.; Ferreira, L. M. A.; Romao, M. J.; Gilbert, H. J.; Bolam, D. N.; Fontes, C. Xyloglucan is recognized by carbohydrate-binding modules that interact with  $\beta$ -glucan chains. *J. Biol. Chem.* **2006**, *281*, 8815–8828.
- (25) Boraston, A. B.; McLean, B. W.; Chen, G.; Li, A.; Warren, R. A. J.; Kilburn, D. G. Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase. *Mol. Microbiol.* **2002**, *43*, 187–194.
- (26) Din, N.; Damude, H. G.; Gilkes, N. R.; Miller, R. C.; Warren, R. A. J.; Kilburn, D. G. C-1-C-X revisited – Intramolecular synergism in a cellulase. *Proc. Natl. Acad. Sci., U.S.A.* **1994**, *91*, 11383–11387.
- (27) Fernandes, A. C.; Fontes, C.; Gilbert, H. J.; Hazlewood, G. P.; Fernandes, T. H.; Ferreira, L. M. A. Homologous xylanases from *Clostridium thermocellum*: Evidence for bifunctional activity, synergism between xylanase catalytic modules and the presence of xylan-binding domains in enzyme complexes. *Biochem. J.* **1999**, *342*, 105–110.
- (28) Linder, M.; Salovuori, I.; Ruohonen, L.; Teeri, T. T. Characterization of a double cellulose-binding domain – Synergistic high affinity binding to crystalline cellulose. *J. Biol. Chem.* **1996**, *271*, 21268–21272.
- (29) Bolam, D. N.; Xie, H.; White, P.; Simpson, P. J.; Hancock, S. M.; Williamson, M. P.; Gilbert, H. J. Evidence for synergy between family 2b carbohydrate binding modules in *Cellulomonas fimi* xylanase 11A. *Biochemistry* **2001**, *40*, 2468–2477.