Trp22, Trp24, and Tyr8 Play a Pivotal Role in the Binding of the Family 10 Cellulose-Binding Module from *Pseudomonas* Xylanase A to Insoluble Ligands[+]

Tamás Ponyi,[‡,§] Lóránd Szabó,[‡,§] Tibor Nagy,[§,||] László Orosz,[‡,||] Peter J. Simpson,[⊥] Michael P. Williamson,[⊥] and Harry J. Gilbert*[,§]

Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K., Department of Molecular Biology and Biotechnology, Krebs Institute, University of Sheffield, Sheffield S10 2TN, U.K., Institute for Molecular Genetics, Agricultural Biotechnology Centre, P.O. Box 411, Gödöllô H-2101, Hungar, and Department of Biotechnology and Microbial Genetics, Gödöllô University of Agricultural Sciences, Páter Károly u. l., Gödöllô H-2101, Hungary

ABSTRACT: Aromatic amino acids are believed to play a pivotal role in carbohydrate-binding proteins, by forming hydrophobic stacking interactions with the sugar rings of their target ligands. Family 10 cellulose-binding modules (CBM10s), present in a number of cellulases and xylanases expressed by *Pseudomonas fluorescens* subsp. *cellulosa*, contain two tyrosine and three tryptophan residues which are highly conserved. To investigate whether these amino acids play an important role in the interaction of CBM10 from *P. fluorescens* subsp. *cellulosa* xylanase A (Pf Xyn10A) with cellulose, each of these residues was changed to alanine in CBM10 expressed as a discrete module or fused to the catalytic domain of Pf Xyn10A (CBM10-CD), and the capacity of the mutant proteins of CBM10-CD to bind the polysaccharide was evaluated. The data showed that W22A, W24A, and Y8A bound very weakly to cellulose compared to the wild-type protein, while Y12A retained its capacity to interact with the glucose polymer. When the W7A mutation was introduced into CBM10 the protein domain did not accumulate in *Escherichia coli*. In contrast, the W7A mutant of CBM10-CD was efficiently expressed in *E. coli*, although the protein bound very weakly to cellulose. NMR spectra of wild-type CBM10, W22A, and W24A were very similar, suggesting that the mutations did not significantly affect the protein fold. Titration of wild-type CBM10, W22A, and W24A with *N*-bromosuccinimide indicated that Trp22 and Trp24 were on the surface of the protein, while Trp7 was buried. Collectively, these data indicate that Trp22, Trp24, and Tyr8 play a direct role in the binding of Pf Xyn10A CBM10 to cellulose. The results are discussed in the light of the three-dimensional structure of CBM10 [Raghothama, S., Simpson, P. J., Szabó, L., Nagy, T., Gilbert, H. J., and Williamson, M. P. (2000) *Biochemistry* 39, 978−984].

Plant cell wall hydrolases expressed by aerobic microorganisms in general have a modular structure consisting of a catalytic domain linked to one or more noncatalytic carbohydrate-binding modules (CBMs)[1] (*1*). The majority of these CBMs interact with cellulose and are thus often referred to as cellulose-binding domains or CBDs (*1*). Currently, there are in excess of 17 families of CBMs (*2*). The three-dimensional structures of five CBMs that bind cellulose and one that interacts with xylan, from five different families, have been solved (*3*−*8*). Each protein contains several surface aromatic residues, which site-directed mutagenesis and chemical modification studies have implicated in ligand binding. In family 1, 2a, 3a, and 5 CBMs, which all bind to crystalline cellulose, the surface aromatic residues form a planar hydrophobic strip that appears to bind to the polysaccharide. In the family 2b CBM (*8*) and family 4 CBM (*6*), the aromatic residues are located in shallow and deep clefts, respectively, hence their capacity to bind single polysaccharide chains.

The molecular architecture of plant cell wall hydrolases from aerobic prokaryotes is generally complex. For example, *P. fluorescens* expresses three cellulases, a cellodextrinase, a xylanase Pf Xyn10A, and a mannanase with a common molecular architecture, which consists of a CBM2, a CBM10, and a catalytic domain (*9*). CBM10s have also been identified in a second xylanase and two other mannanases (*10*). CBM10s are small modules, comprising only ca. 45 residues, that bind to insoluble forms of cellulose with an affinity 6-fold lower than CBM2a and, in common with CBM2a and CBM3a, exhibit extremely low affinity for cellooligosaccharides and soluble forms of cellulose (*11, 12*). CBM10s therefore appear to have the same substrate specificity as CBM2as, while binding more weakly. Deletion of the CBM10 from full-length Pf Xyn10A, leaving an enzyme consisting just of the CBM2a and catalytic domain, has no measurable effect on enzyme activity *(12*). It is therefore

[+] Supported by BBSRC grant 13/P09341and a CASE award, and by equipment grants from the BBSRC and the Wellcome Trust. The Krebs Institute is a Centre for Molecular Recognition of the BBSRC/EPSRC.

* To whom correspondence should be addressed. Corresponding author: H. J. Gilbert Fax. No. 0191-2228684; Tel. No. 0191-2226962; E.mail; H. J.Gilbert@Newcastle.ac.uk.

[‡] Gödöllô University of Agricultural Sciences.

[§] University of Newcastle upon Tyne.

[||] Institute for Molecular Genetics.

[⊥] University of Sheffield.

[1] Abbreviations: CBM, carbohydrate binding module; CBM10, family 10 carbohydrate-binding module from *Pseudomonas fluorescens* subspecies *cellulosa*; CD, catalytic domain.

10.1021/bi9921642 CCC: $19.00
Published on Web 01/14/2000

not clear why a CBM2a-CBM10 combination should have been selected in so many different enzymes. The current study is part of a continuing investigation into this problem.

CBM10s contain three tryptophan and two tyrosine residues which are completely conserved *(10)*. In view of the fact that aromatic residues (Tyr and Trp and to a small extent His) have been shown to play an important role in the interaction between various polysaccharide-binding domains and their respective ligands *(13, 14)*, it is likely that some of the aromatic residues in CBM10s mediate the binding of these modules to cellulose. To test this hypothesis we have used a site-directed mutagenesis and chemical modification approach to investigate the importance of the aromatic residues of CBM10 from Pf Xyn10A. The data indicated that Trp22, Trp24, and Tyr8 play a pivotal role in the interaction of CBM10 with its insoluble ligand.

MATERIALS AND METHODS

Bacterial Strains, Plasmids and Culture Conditions. The *Escherichia coli* strains used in this study were BL21(DE3): pLysS (Novagen) and XL1-Blue (Stratagene). The plasmids used, which consisted of regions of the gene (*xynA*) encoding Pf Xyn10A cloned into pET expression vectors, were as follows: pJG2, encoding CBM10 from Pf Xyn10A (CBM10); and pJG3, encoding a truncated version of Pf Xyn10A (CBM10-CD) comprising the catalytic domain and CBM10. Details of how these plasmids were constructed were described previously (*11*). All of the *E. coli* strains were cultured in Luria broth (LB) supplemented with 50 μg/mL ampicillin. To generate the recombinant proteins, *E. coli* BL21(DE3) harboring pLysS and the appropriate recombinant plasmids was cultured to mid-exponential phase (A_{550} = 0.5) in LB containing ampicillin at 30 °C, after which time isopropylthio-β-D-galactopyranoside was added to a final concentration of 1 mM and the cultures were incubated for a further 4 h at 30 °C.

Purification of Recombinant Proteins. CBM10, CBM10-CD, and their corresponding mutants all contained N-terminal His_{10} tags and were therefore purified by nickel ion affinity chromatography, using a Talon metal-affinity column. The purification protocols were as described previously *(12)*. Approximately 2 mg of native and mutant forms of CBM10 (except W7A which did not accumulate) and 30 mg of the corresponding CBM10-CD proteins were produced per liter of culture.

Site-Directed Mutagenesis. Mutants of CBM10 and CBM10-CD were generated by site-directed mutagenesis of the appropriate regions of *xynA* using the Transformer kit supplied by CloneTec. The selection primer used in the Transformer method was 5′ GACTGGTGAGTATTCAACCAAGTC 3′, and the mutagenic primers were as follows: W7A, 5′GTGCCGTAC**GC**ATTACATTG 3′; W22A, 5′ CTTCCCAGCC**GGC**TCCATTGGTGG 3′; W24A, 5′ CTGGTCTTC**GGC**GCCCCATCCATTGG 3′; Y8A, 5′ GTGTGCCG**GC**CCAATTAC 3′; Y12A, 5′ GCACAGTGGA**GC**GAGTGTG 3′.

The nucleotides in bold are the mutations incorporated into *xynA*.

Ligand Binding Assays. A qualitative analysis of the capacity of CBM10, CBM10-CD, and their respective mutants to bind to Avicel was determined as follows: an equal volume of purified protein (0.5 mg/mL) and Avicel (5% w/v) was mixed and incubated at 4 °C for 1 h. The cellulose was pelleted, the supernatant removed, and the polysaccharide washed three times in 1 mL of 50 mM sodium phosphate buffer, pH 7.0. The final cellulose pellet was resuspended in 100 μL of 10% (w/v) SDS, boiled for 5 min, and subjected to SDS−PAGE *(15)*. Quantitative determination of cellulose binding was performed essentially as described by Bolam et al. *(12)*. Briefly, 0.5−10 μM of each protein, in 50 mm Tris/HCl buffer, pH 8.0, was mixed with 500 μg of Avicel in a final volume of 450 μL at 4 °C for 1 h. The cellulose was then pelleted by centrifugation at 13000*g* for 1 min, and the concentration of soluble protein was measured at 280 nm.

Titration of CBM10 with N-Bromosuccinamide. A 2 μL aliquot of 10 mM *N*-bromosuccinimide (NBS) was added to wild-type and mutant forms of CBM10, at a concentration of approximately 50 μM for wild-type protein (70 μM for W22A and W24A), in 50 mM sodium acetate buffer, pH 4.5, in a volume of 625 μL. The mixture was incubated for 3 min at 25 °C and the A_{280} value recorded from a 250−320 nm scan of the protein using a Pharmacia Ultrospec 4000 spectrophotometer. Further NBS aliquots were added to the protein until the A_{280} value plateaued, or started to rise. When denatured protein was titrated with NBS, CBM10 and its mutants were dialyzed into 50 mM sodium acetate buffer, pH 4.5, containing 8 M urea. The concentration of wild-type and mutant forms of CBM10 were determined from their molar extinction coefficients at 280 nm, which were 19 700 M^{-1} cm^{-1} for the wild-type protein and 14 000 M^{-1} cm^{-1} for the mutants W22A and W24A. The equation used to determine the number of tryptophans oxidized was as follows:

$$(\Delta A_{280}/4330) \times \text{reaction volume } (\mu\text{L}) = \mu\text{mol of tryptophan oxidized} \quad (1)$$

Equation 1 is based on the molar extinction coefficients for tryptophan and the amino acid oxidized by NBS as 5700 M^{-1} cm^{-1} and 1370 M^{-1} cm^{-1}, respectively *(16)*, thus the ΔA_{280} for the oxidation of a 1 M solution of tryptophan is 4330.

NMR and Circular Dichroism Spectroscopy. For NMR spectroscopy, wild-type and mutant forms of CBM10 were concentrated to 700 μM in 500 μL of 50 mM sodium phosphate buffer containing 100 mM sodium chloride and 10 mM sodium azide, pH 4.5, 10% (v/v) ^2H_2O. NMR spectra were recorded with a Bruker DRX 500 spectrometer, and the ^1H chemical shifts were referenced to an internal standard of 3-trimethylsilyl-2,2,3,3-d_4-propionate at 0.00 ppm. Experiments were acquired at 303 K with a resolution of 8192 complex points over a sweep width of 12 500 Hz. All data were processed with a Gaussian convolution using FELIX 97.0 (Molecular Simulations Inc.). Circular dichroism spectroscopy was performed as described previously using a Jobin−Yvon CD6 spectropolarimeter (*17*). Briefly, the spectra were obtained at a residue concentration of 5.0 to 5.5 mM in 10 mM Tris/HCl buffer, pH 8.0, at 25 °C using a 0.1 mm path-length quartz cuvette. Each spectrum was accumulated from 20 to 30 scans between 190 and 250 nm, at a scan rate of 60 nm/min.

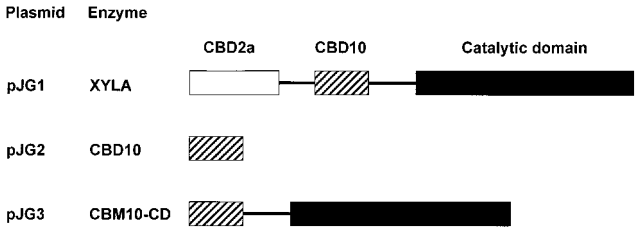

FIGURE 1: Schematic of Pf Xyn10A and its derivatives. The different modules in Pf Xyn10A are depicted as follows: catalytic domain (black bar); CBM2 (white bar); CBM10 (striped bar); linker sequences (—).

RESULTS

Generation of Mutants of CBM10. The family 10 CBM of Pf Xyn10A was expressed either as a discrete domain (CBM10) or joined to the enzyme's C-terminal catalytic domain by a serine-rich linker sequence (CBM10-CD; Figure 1). All the tyrosine and tryptophan residues were individually mutated to alanine to generate the following mutations in both CBM10 and CBM10-CD: W7A, W22A, W24A, Y8A, and Y12A. The number of the residue mutated refers to its position in the CBM10 domain: residue 1 in CBM10 corresponds to residue 120 in full-length Pf Xyn10A (*18*). All the mutant proteins could be expressed in *E. coli*, except CBM10 containing the W7A mutation, which did not accumulate in *E. coli*. The enzymes were purified by nickel ion affinity chromatography to apparent homogeneity as judged by SDS−PAGE (Figure 2). The observation that the W7A mutation prevented the accumulation of CBM10 in *E. coli* suggests that the mutant polypeptide did not fold correctly.

Affinity of CBM10 Mutants for Cellulose. To quantify the binding of native and mutant forms of CBM10 to cellulose, a series of binding isotherm experiments were performed using native and mutant forms of CBM10-CD. The data (Figure 3 and Table 1), which fitted a one-site binding model, were used to calculate the affinity constant for binding to cellulose. The results showed that the affinity for cellulose of the mutant proteins in which any of the three Trp residues or Tyr8 had been modified, was greatly reduced compared to native CBM10-CD. Suprisingly, the number of available binding sites on the cellulose was variable for the different forms of CBM10-CD. Similar amounts of wild-type and Y12A were required to saturate the cellulose, while considerably less binding sites on the polysaccharide were available to the other four proteins, particularly for the mutants Y8A and W22A. The explanation for the variable number of cellulose-binding sites for the different forms of CBM10-CD is currently not obvious. Mutation of Tyr12 had very little effect on either ligand binding or bound protein at saturation.

Table 1: Affinity of Wild-Type and Mutants Forms of CBM10 for Cellulose

protein	mutation	affinity (K_r^+)[a] for Avicel
CBM10-CD	none	1.65
CBM10-CD	W7A	0.13
CBM10-CD	W22A	0.04
CBM10-CD	W24A	0.10
CBM10-CD	Y8A	0.12
CBM10-CD	Y12A	1.22

[a] K_r^+, relative affinity (l/g).

NMR and CD Spectra of Native and Mutants of CBM10. To investigate whether the mutations introduced into CBM10 altered the gross structure of the protein, native CBM10 and the W22A and W24A mutants of the domain were subjected to NMR analysis. Figure 4 shows that signals from the side chain of Ile31 and from the aromatic ring of Tyr8, which are strongly ring current shifted, are in similar positions in the two mutant proteins W22A and W24A. Calculations of ring current shifts using the solution structure *(19)* show that Ile31 and Tyr8 are shifted almost entirely by the aromatic rings of Trp7 and Phe43. Therefore, the similar chemical shifts of these signals in the mutant proteins demonstrate that the relative positions of Ile31, Trp7, and Phe43, which together form most of the hydrophobic core, remain almost unchanged in the mutants (the Ile31 γ-methyl signal at 0.35 ppm is also partially ring-current shifted by Trp22, hence the small change in position in the W22A mutant spectrum). These data suggest that changing either Trp22 or Trp24 to Ala did not cause a substantial change in the conformation of the protein. This view is supported by CD spectra of the

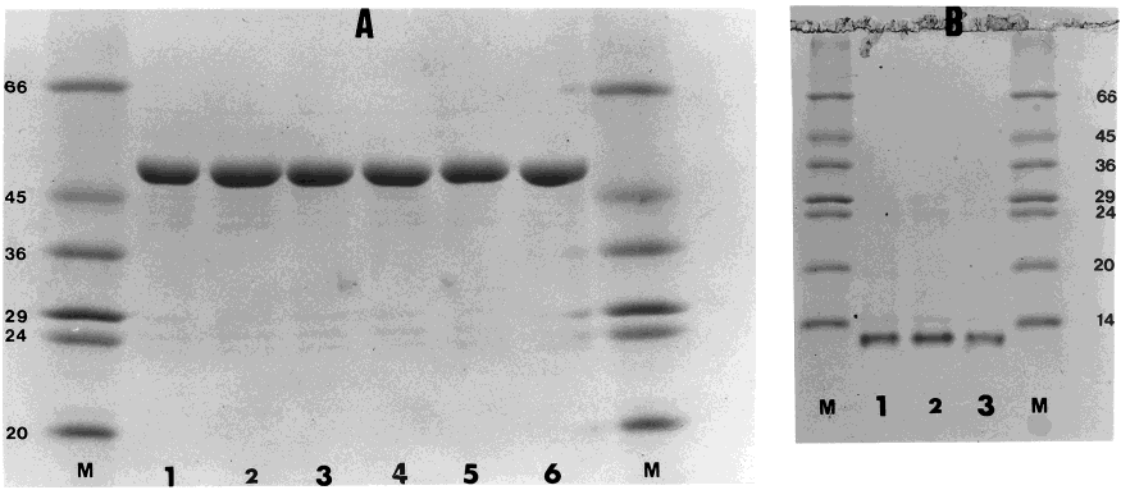

FIGURE 2: SDS−PAGE of mutants of CBM10 and CBM10-CD. Panel A contains wild-type (1) and mutants W7A (2), W22A (3), W24A (4), Y8A (5), and Y12A (6) of CBM10-CD. Panel B contains wild-type (1) and mutants W22A (2) and W24A (3) of CBM10. The proteins were subjected to SDS−PAGE using a 10% (w/v) polyacrylamide gel. Low molecular weight size markers from Sigma are in lane M, and the size of proteins ($M_r \times 10^3$) are shown.

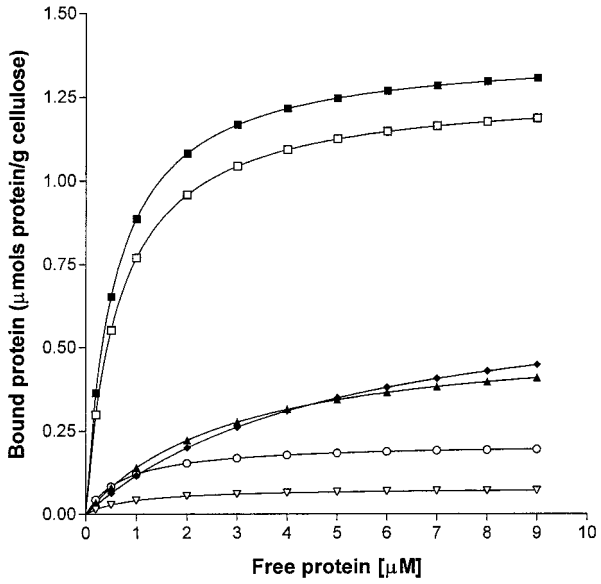

FIGURE 3: Binding isotherms of wild-type and mutants of CBM10-CD using Avicel as the ligand. Binding isotherm measurements of wild-type (■) CBM10-CD and the following mutants of the protein; W7A (▲), Y8A (○), Y12A (□), W22A (▽), and W24A (◆) were performed as described in Materials and Methods, using Avicel as the ligand.

W22A and W24A mutants of CBM10-CD, which were indistinguishable from the native protein (data not shown).

N-Bromosuccinimide Titrations (NBS). To evaluate whether Trp7, Trp22, and Trp24 are on the suface of CBM10 or in the internal hydrophobic core, the aromatic side chains were oxidized by titration of the protein with NBS. The oxidation of the tryptophan side chains was monitored by measuring a decrease in the A_{280} of the protein. An example of the titrations is displayed in Figure 5. The A_{280} of native CBM10 decreased up to a molar ratio of NBS:protein of 8:1, and the absorbance then plateaued. These data are consistent with the oxidation of two tryptophan residues. Further additions of the oxidizing agent in a molar excess greater than 13 (over the protein) caused an increase in the A_{280} value of the protein. When native CBM10, which had been denatured with urea, was titrated against NBS, the decrease in the A_{280} value up to a molar ration of NBS:protein of 10:1 was consistent with the modification of three tryptophan residues; there was no increase in absorbance on subsequent additions of the oxidizing agent. The increase in A_{280} on the addition of excess NBS to native CBM10 is likely to be caused by oxidation of tyrosine residues as the product has an ϵ_{280} of 5400×10^3 M^{-1} cm^{-1}, which is approximately three times larger than the equivalent value for tyrosine at 280 nm (*20*, *21*). This view is supported by the observation that NBS does not react with tyrosine in urea *(16)*, hence, the A_{280} value did not increase when urea-denatured CBM10 was incubated with a large excess of the oxidizing agent. These data indicate that only two of the three tryptophan residues of native CBM10 are accessible to the oxidizing agent, and the third tryptophan is only oxidized by NBS when the protein is unfolded. The results therefore suggest that two of the tryptophan amino acids are on the surface of the CBM, while the third residue is buried in the hydrophobic core of the protein.

To investigate which of the tryptophan residues have been modified by NBS, W22A and W24A were also titrated with the oxidizing agent. The data, presented in Figure 5, show that in both mutant proteins only one surface Trp has been modified by NBS, indicating that Trp22 and Trp24 are on the surface of the CBM. By implication, Trp7 is therefore buried in the interior.

DISCUSSION

The binding of carbohydrate-binding modules to their respective ligands is mediated primarily by hydrophobic stacking interactions between sugar pyranose rings and aromatic amino acids (*22*). CBM10 contains six aromatic residues; three tryptophans, two tyrosines, and a phenylalanine. Phenylalanine has not been observed to act as a surface-exposed ligand for saccharides in any CBMs examined to date and is always buried in the interior. Site-directed mutagenesis data, presented in this paper, concentrated on the Tyr and Trp residues and showed that Y12A did not affect the affinity of the protein for cellulose, suggesting that Tyr12 does not play a direct role in the binding of the polysaccharide to the CBM. This is consistent with the

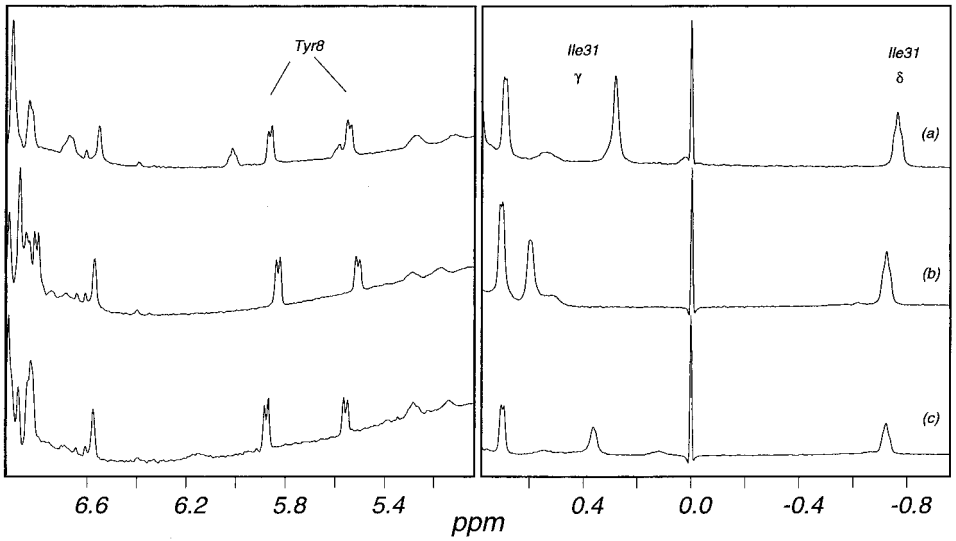

FIGURE 4: NMR spectra of (a) wild-type CBM10, (b) W22A, (c) W24A. The positions in the wild-type spectrum of signals from Tyr8 aromatic protons, and Ile31 side-chain protons are indicated. The sharp signal at 0.0 ppm is the chemical shift reference.

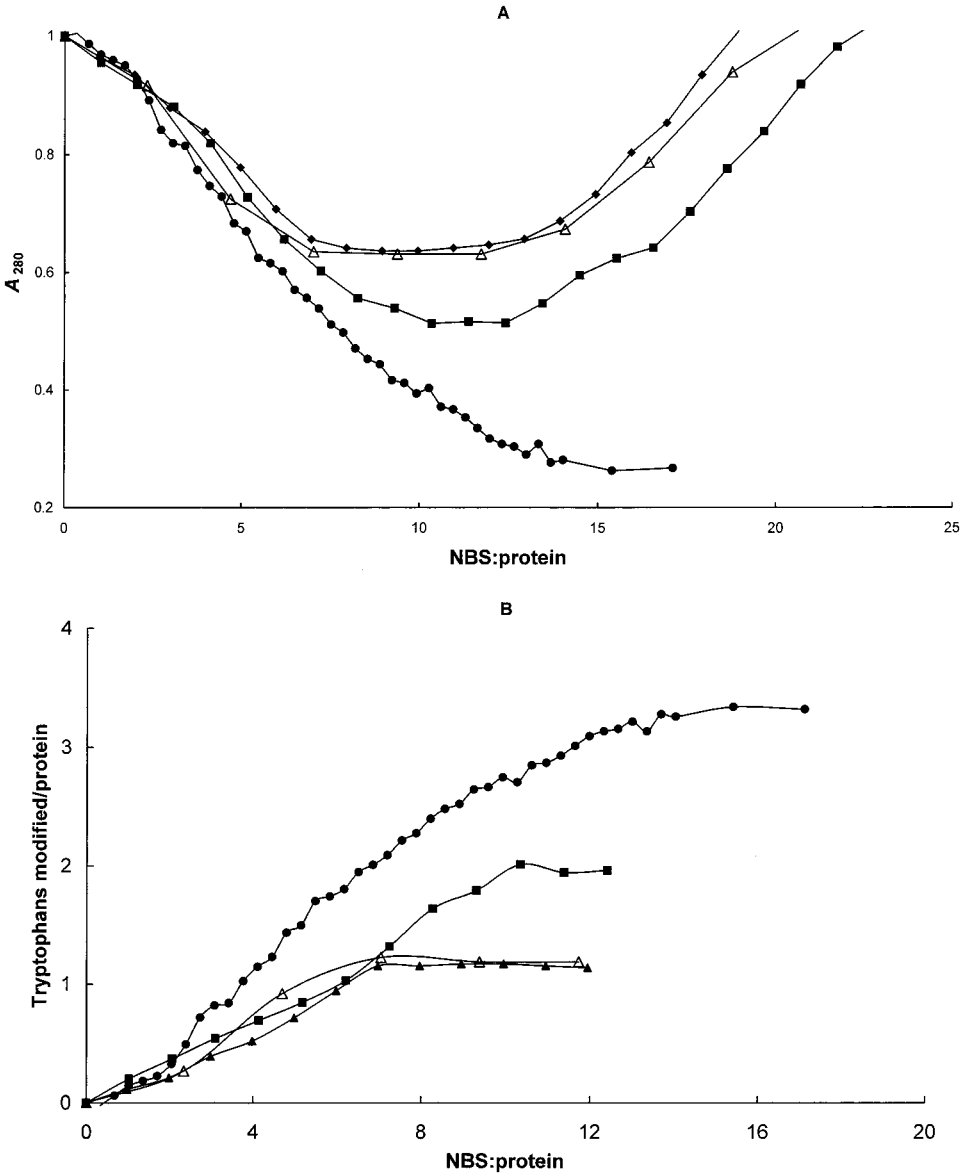

FIGURE 5: NBS titrations of wild-type and mutant forms of CBM10. Native (■) and denatured (●) wild-type CBM10 and W22A (▽) and W24A (◆) mutants of the protein (folded) were titrated with NBS as described in Materials and Methods. Panel A shows the change in A_{280} of the proteins titrated with the oxidizing agent, while Panel B shows the number of tryptophan residues modified by NBS.

solution structure of CBM10 which is presented in the preceding paper in this issue (*19*) and shows that Tyr12 is in a semiburied location within the protein, precluding a direct interaction between this amino acid and the ligand.

In contrast, conversion of either Trp22 or Trp24 to alanine greatly reduced the affinity of the protein for its ligand. In principle, loss of binding could arise either from direct loss of a binding interaction or from disruption of the three-dimensional structure of the protein. Disruption of the structure is easily characterized by observation of changes to ^{1}H NMR proton signals that are shifted by aromatic ring currents. In NMR spectra of proteins, aromatic rings of tyrosine, tryptophan, and phenylalanine residues cause ring current shifts of adjacent protons away from their expected random-coil (unfolded protein) frequency. In the spectra shown in Figure 4, signals due to the δ and γ methyl groups of Ile31 and the ring protons of Tyr8 are significantly different from their random coil chemical shift values and are thus characteristic of the CBM10 protein fold. Any changes to the structure of the hydrophobic core, down to a few tenths of an ångström, are readily observed as moves back toward the random coil shift values. 1D NMR data indicated that the 3D structures of W22A and W24A were very similar to the wild-type CBM10, suggesting that the two tryptophan residues do not play an important structural role in CBM10 and, therefore, play a direct role in the binding of the protein to cellulose. This view is supported by the structure of the domain presented in the previous paper in this issue (*19*), which shows that both Trp22 and Trp24 are on the surface of the CBM and are thus in an appropriate location to interact with plant structural polysaccharides.

To further probe the position of Trp22 and Trp24 in CBM10, wild-type and mutant forms of the protein were titrated with NBS. This reagent preferentially oxidizes tryptophan indole rings on the surface of proteins causing a reduction in A_{280}. The reduction in the A_{280} of wild-type CBM10 by NBS was consistent with the structure of the protein (*19*), in which Trp22 and Trp24 are exposed on the surface and available for oxidation. Trp7 is only oxidized in denaturing conditions, even though it is partially exposed.

However, a detailed study of the structure of the protein (*19*) shows that although the Nϵ and adjacent Cδ of Trp7 are surface exposed, the rest of the residue is buried. NBS is generally assumed to react with tryptophan by a nucleophilic attack of the Cγ of tryptophan on the bromine atom of the reagent. This attack is blocked in the structure of the native protein, consistent with the observed lack of reaction.

The data described in this report are consistent with a previous study (*23*) which showed that only the surface tryptophan residues of CBM1 from *Trichoderma reesei* endoglucanase II reacted with NBS. In contrast, the three surface and the two internal tryptophan residues of $CBM2_{Cex}$ from *Cellulomonas fimi* Cex were modified by the oxidizing agent, although the exposed aromatic residues reacted more rapidly with the reagent than the two buried amino acids (*24*). It was suggested that structural fluctuations of $CBM2_{Cex}$ may make the internal tryptophans accessible to NBS, if only transiently, or oxidation of the surface residues destabilized the protein, making the other two tryptophans available to the reagent. The observation that the internal tryptophan of CBM10 was not modified, even by a large excess of the oxidizing agent, suggests that the protein domain has little mobility around Trp7. The rigidity of the protein is evidenced by essentially unchanged NMR spectra up to at least 50 °C (*19*). In addition, oxidation of the surface tryptophan residues does not appear to destabilize the protein as Trp7 was not subsequently modified.

Although the data presented in this report strongly indicate that Trp7 is internal, changing this amino acid to an alanine residue greatly reduced the protein's capacity to bind to cellulose. It is therefore likely that this residue does not play a direct role in ligand binding, but is important in the correct folding of the protein. From the structure (*19*), it appears that substituting Trp7 with alanine will disrupt the internal packing of the protein. This view is supported by the observation that although wild-type, W22A and W24A mutants of CBM10 could be produced in *E. coli*, W7A did not accumulate in the enteric bacterium. It is well established that single amino acid substitutions that significantly disrupt the structure of proteins prevent their accumulation in *E. coli*, presumably because these proteins are particularly susceptible to proteolysis (*25*). In contrast, W7A was produced in *E. coli* when appended to the catalytic domain of Pf Xyn10A. This is not particularly surprising as small unfolded recombinant polypeptides only accumulate in *E. coli* when fused to large globular proteins such as β-galactosidase (*26*).

The cellulose ligands identified in this work (i.e., two tryptophans and one tyrosine) are typical for domains that bind to crystalline cellulose. Cellulose-binding proteins also typically have other polar residues in the binding site that hydrogen bond to the specific substrate (*5*). From the structure (*19*), it was speculated that Thr10, Tyr12, Ser29, and Gln39 would be possible polar ligands in that they are highly conserved and close to the cellulose-binding site. In this work, we have shown that mutation of Tyr12 to alanine has very little effect on the binding. The most common polar ligands for saccharides are Glx and Asx, and Gln39 is therefore the most promising candidate for a specific hydrogen-bonding interaction.

This study, together with previous results (*13*, *19*), has revealed that the CBM2a and CBM10 of *P. fluorescens* both bind to cellulose using three exposed coplanar aromatic rings.

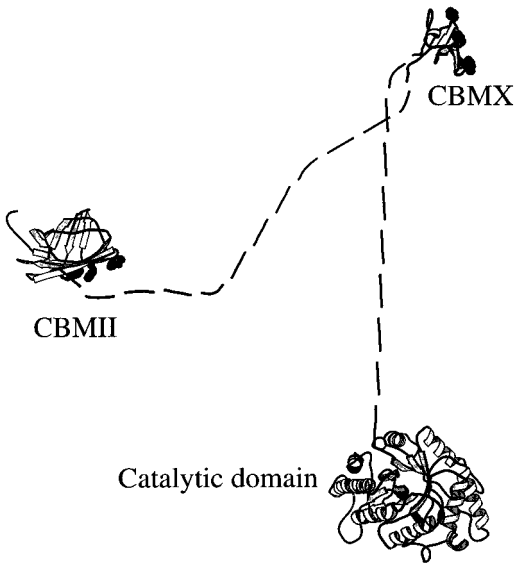

FIGURE 6: Model of the structure of full-length Pf Xyl10A. The linkers between the globular domains are flexible and are therefore indicated by dashed lines. The length of the linker is correct, but the spatial locations of the globular domains are chosen at random.

The CBM10 has an affinity ca. 6 times weaker than that of CBM2a, and there is no measurable synergy in binding between the two domains, raising the question of the function of CBM10. The results presented in this paper and the preceding paper in this issue (*19*) essentially complete the structure determination of all of the globular parts of the protein: there is already a structure of the catalytic domain (*27*) and of a homologous CBM2 (*4*). We have therefore modeled the structure of full-length Xyn10A (Figure 6) using the coordinates for these domains together with an extended linker chain. The most striking aspect of the model is that the binding domains are separated from the catalytic domain by long and flexible linkers. It is also noteworthy that the linkers from the CBM10 both protrude from the side of the domain facing away from the binding site, suggesting that the CBM10 may act as a "secondary" binding domain. The binding domains both bind preferentially to cellulose, although the catalytic domain attacks xylan. The model therefore places an upper limit (of ca. 14 nm) on the spatial separation between the cellulose- and xylan-binding sites in the plant cell wall.

In summary, the data presented in this report show that Trp22 and Trp24 are located on the surface of CBM10, and play a direct role in binding cellulose. Tyr8 also appears to be involved in ligand binding, while Trp7 and Tyr12 do not interact directly with the polysaccharide.

REFERENCES

1. Tomme, P., Warren, R. A., and Gilkes, N. R. (1995) *Adv. Microbiol. Physiol. 37*, 1−81.
2. Henrissat, B. (1999) in *Recent Advances in Carbohydrate Engineering* (Gilbert, H. J., Davies, G. J., Svensson, B., and Henrissat, B., Eds.) (in press).
3. Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Petterson, G., Knowles, J., and Gronenborn, A. M. (1989) *Biochemistry 28*, 7241−7257.
4. Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris the-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) *Biochemistry 34*, 6993−7009.

5. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J. 15*, 5739−5751.
6. Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) *Biochemistry 35*, 13895−13906.
7. Brun, E., Moriaud, F., Gans, P., Blackledge, M. J., Barras, F., and Marion, D. (1997) *Biochemistry 36*, 16074−16086.
8. Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J., and Williamson, M. P. (1999) *Structure 7*, 853−864.
9. Hazlewood, G. P., and Gilbert, H. J. (1998) *Progr. Nucleic Acid Res. Mol. Biol. 61*, 211−241.
10. Millward-Sadler, S. J., Davidson, K., Hazlewood, G. P., Black, G. W., Gilbert, H. J., and Clarke, J. H. (1995) *Biochem. J. 312*, 39−48.
11. Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998) *Biochem. J. 331*, 775−78.
12. Gill J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., and Gilbert, H. J. (1999) *Biochem. J. 342*, 473−480.
13. Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J., and Orosz, L. (1998) *FEBS Lett. 429*, 312−316.
14. Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C., Jr, Warren, R. A. J., and Kilburn, D. G. (1994) *Mol. Microbiol. 11*, 747−755.
15. Laemmli, U. K. (1970) *Nature (London) 227*, 680−685.
16. Spande, T. F., and Witkop, B. (1967) *Methods Enzymol. 11*, 498−506.
17. 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.
18. Hall, J., Hazlewood, G. P., Huskisson, N. S., Durrant, A. J., and Gilbert, H. J. (1989) *Mol. Microbiol. 3*, 1211−1219.
19. Raghothama, S., Simpson, P. J., Szabó, L., Nagy, T., Gilbert, H. J., and Williamson, M. P. (2000) *Biochemistry* 39, 978−984.
20. Bray, M. R., Carriere, A. D. and Clarke, A. J. (1994) *Anal. Biochem. 221*, 278−284.
21. Bray, M. R., and Clarke, A. J. (1995) *Biochemistry 34*, 2006−2014.
22. Quiocho, F. A. (1986) *Annu. Rev. Biochem. 55*, 287−315.
23. Macarron, R., Henrissat, B., Van Beeuman, J., Dominguez, J. M., and Claeyssens, M. (1995) in *Enzymatic degradation of insoluble substrates* (Saddler, J. N., and Penner, M. H., Eds.) pp 164−173, ACS Symposium Series No 618, American Chemical Society, Washington DC.
24. Bray, M. R., Johnson, P. E., Gilkes, N. R., McIntosh, L. P., Kilburn, D. G., and Warren, R. A. J. (1996) *Protein Sci. 5*, 2311−2318.
25. Gilbert, H. J., and Drabble, W. T. (1980). *J. Gen. Microbiol. 117*, 33−45.
26. Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F., and Boyer, H. W. (1977) *Science 198*, 1056−1063.
27. Harris, G. W., 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*, 1107−1116.

BI9921642