Protein Science (1995), *4*:1056-1064. Cambridge University Press. Printed in the USA.

Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I

MARKUS LINDER,[1,2] MAIJA-LIISA MATTINEN,[3] MAARIT KONTTELI,[3] GUNNAR LINDEBERG,[4] JERRY STÅHLBERG,[5] TORBJÖRN DRAKENBERG,[3] TAPANI REINIKAINEN,[2] GÖRAN PETTERSSON,[1] AND ARTO ANNILA[3]

[1] Department of Biochemistry, University of Uppsala, Uppsala, Sweden
[2] VTT Biotechnology and Food Research, Espoo, Finland
[3] VTT Chemical Technology, Espoo, Finland
[4] Department of Medical and Physiological Chemistry, University of Uppsala, Uppsala, Sweden
[5] Department of Molecular Biology, University of Uppsala, Uppsala, Sweden

(RECEIVED February 10, 1995; ACCEPTED April 3, 1995)

Abstract

Cellobiohydrolase I (CBHI) of *Trichoderma reesei* has two functional domains, a catalytic core domain and a cellulose binding domain (CBD). The structure of the CBD reveals two distinct faces, one of which is flat and the other rough. Several other fungal cellulolytic enzymes have similar two-domain structures, in which the CBDs show a conserved primary structure. Here we have evaluated the contributions of conserved amino acids in CBHI CBD to its binding to cellulose. Binding isotherms were determined for a set of six synthetic analogues in which conserved amino acids were substituted. Two-dimensional NMR spectroscopy was used to assess the structural effects of the substitutions by comparing chemical shifts, coupling constants, and NOEs of the backbone protons between the wild-type CBD and the analogues. In general, the structural effects of the substitutions were minor, although in some cases decreased binding could clearly be ascribed to conformational perturbations. We found that at least two tyrosine residues and a glutamine residue on the flat face were essential for tight binding of the CBD to cellulose. A change on the rough face had only a small effect on the binding and it is unlikely that this face interacts with cellulose directly.

Keywords: cellulase; cellulose-binding domain; mutation; NMR spectroscopy; protein–carbohydrate interaction; synthetic peptide

Cellulose is degraded in nature by the concerted action of several synergistically functioning enzymes. Depending on their mode of action, cellulolytic enzymes fall into one of two main groups, endoglucanases or cellobiohydrolases (Béguin & Aubert, 1994). The filamentous fungus *Trichoderma reesei* secrets a mixture of cellulases consisting of two CBHs and several EGs (Wood, 1992; Saloheimo et al., 1994). Sequence determination and functional studies have shown that most fungal cellulases have a bifunctional organization in which a catalytic core domain and a cellulose-binding domain are connected by a flexible glycosylated linker peptide (Van Tilbeurgh et al., 1986; Tomme et al., 1988). Both domains bind to cellulose, but the affinity of the core domain is much lower than that of the CBD (Ståhlberg et al., 1991). The intact enzyme and the isolated core domain have equal activities toward soluble substrates, but the activity of the core domain toward crystalline cellulose is severely impaired (Tomme et al., 1988). The tight binding mediated by the CBD is thus necessary for an efficient attack on crystalline cellulose. It is currently not known whether the CBD merely anchors the enzyme to the cellulose substrate, or whether it also actively facilitates the breakdown of the crystalline structure (Teeri et al., 1992).

The amino acid sequences of several fungal cellulases have been reported. All the fungal CBDs show strong sequence similarity, some residues being completely conserved and some dis-

Reprint requests to: Markus Linder, Biotechnology and Food Research, Box 1500, FIN-02044 VTT, Finland; e-mail: markus.linder@vtt.fi.

Abbreviations: CBD, cellulose-binding domain; CBH, cellobiohydrolase; CBHI, cellobiohydrolase I; EG, endoglucanase; RELAY, relayed coherence transfer spectroscopy; FID, free induction decay; NOESY, two-dimensional NOE spectroscopy; COSY, correlated spectroscopy.

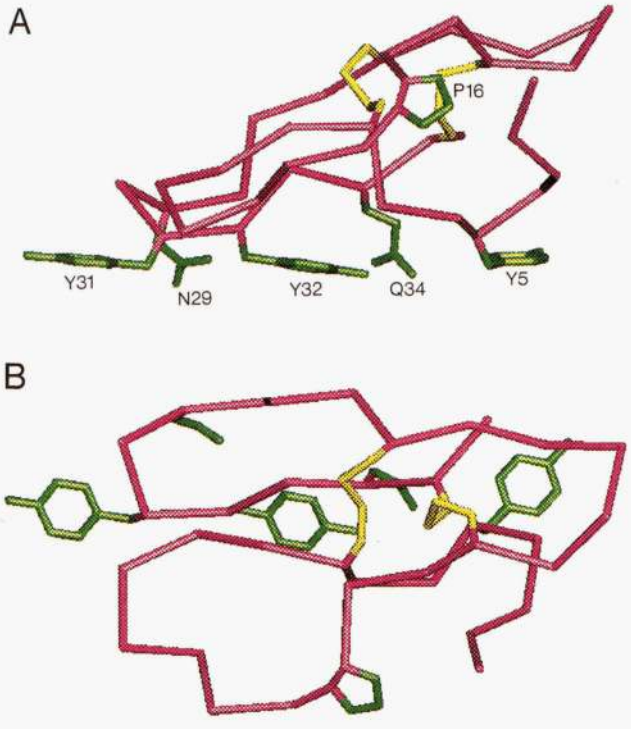

Fig. 1. α-Carbon trace (red) of the cellulose-binding domain of *T. reesei* CBHI (Kraulis et al., 1989) with the side chains of the substituted amino acids shown (green). Disulfide bridges are also shown (yellow). The view in Figure 1A is from the side and that in Figure 1B is from above, looking down at the rough face. Positions of the side chains on the flat face formed by the three tyrosines suggest functional roles in the binding to cellulose.

playing conservative substitutions. These CBDs belong to family B of CBDs as classified by Gilkes et al. (1991) or to family I as classified by Coutinho et al. (1992).

In this study we have assessed the contribution of several highly conserved residues to the binding of the *T. reesei* CBHI CBD to cellulose by using a set of synthetic peptides wherein these residues have been altered (see Kinemage 1). The advantage of using synthetic CBDs, in contrast to mutants of the intact enzyme, is that the properties of the CBD can be directly studied. Studies using intact enzymes will always be affected by the core domain, which also binds. Also, it has been shown before that a 36-residue-long synthetic analogue of the CBHI CBD and the corresponding CBD proteolytically cleaved from intact CBHI have the same affinities (Johansson et al., 1989).

Whenever amino acid substitutions are introduced in proteins, the observed functional alterations can be caused either by the altered side-chain functionality or by a structural change induced by the mutation. To evaluate possible structural changes in the engineered peptides we examined their backbone folds by two-dimensional NMR spectroscopy and evaluated chemical shift differences between backbone protons in the wild type and the engineered peptides. Changes in dihedral angles and short interproton distances were examined to obtain information about the nature of the structural changes.

Results

Choice of the engineered peptides

The flat face of the CBHI CBD has tyrosines at positions 5, 31, 32 (Fig. 1; Kinemage 1). The aromatic nature of these positions is strictly conserved in all fungal CBDs studied so far (Fig. 2). We studied the contributions of the three tyrosines by replacing the aromatic side chain by functionally neutral alanines in the peptides Y5A, Y31A, and Y32A. Two other residues on the flat face, asparagine 29 and glutamine 34, are strictly conserved in the fungal CBDs. These residues could participate in hydrogen bonding interactions with cellulose, and their roles were studied by exchanging them for alanines in the peptides N29A and Q34A. There are no obvious interaction sites to cellulose on the other rough face. To see whether this face has any interactions with cellulose, the surface was broken by introducing the bulky, positively charged side chain of arginine at position 16 in the peptide P16R.

SOURCE		Position	1	5	10	15	20	25	30	35
Trichoderma reesei	CBHI	461	T Q S H	Y G Q C G G I G Y S G	P	T V C A S G T T C Q V L	N	P	Y	Y	S	Q	C L
Trichoderma reesei	CBHII		Q A C S S V	W G Q C G G Q N W S G	P	T C C A S G S T C V Y S	N	D	Y	Y	S	Q	C L
Trichoderma reesei	EGI	421	S C T Q T H	W G Q C G G I G Y S G	C	K T C T S G T T C Q Y S	N	D	Y	Y	S	Q	C L
Trichoderma reesei	EGII		Q Q T V	W G Q C G G I G W S G	P	T N C A P G S A C S T L	N	P	Y	Y	A	Q	C I
Trichoderma viridae	CBHI	460	T Q T H	Y G Q C G G I G Y I G	P	T V C A S G S T C Q V L	N	P	Y	Y	S	Q	C L
Agaricus bisporus	CEL 1	284	T I P Q	Y G Q C G G I G W T G	G	T G C V A P Y Q C K V I	N	D	Y	Y	S	Q	C L
Agaricus bisporus	CEL 3		Q S P V	W G Q C G G N G W T G	P	T T C A S G S T C V K Q	N	D	F	Y	S	Q	C L
Penicillium janthinellum	CBHI	502	G A R D	W A Q C G G N G W T G	P	T T C V S P Y T C T K Q	N	D	W	Y	S	Q	C L
Phanerochaete crysosporium	CBHI	480	T V P Q	W G Q C G G I G Y T G	S	T T C A S P Y T C H V L	N	P	Y	Y	S	Q	C Y
Humicola grisea	CBHI	490	K A G R	W Q Q C G G I G F T G	P	T Q C E E P Y I C T K L	N	D	W	Y	S	Q	C L
Humicola insolens	EGV	246	G C T A G R	W A Q C G G N G W S G	C	T T C V A G S T C T K I	N	D	W	Y	H	Q	C L
Trichoderma reesei	EGV	205	Q Q T L	Y G Q C G G A G W T G	P	T T C Q A P G T C K V Q	N	Q	W	Y	S	Q	C L
				f f r r r r r r f f f f f									

Fig. 2. Alignment of CBDs from fungal cellulases showing sequence similarity. Substituted amino acids are indicated by boxes. At the bottom line an **f** or an **r** indicates if the residue side chain is exposed on the flat or the rough face, respectively, in the CBHI CBD. CBDs are from: *T. reesei* CBHI (Shoemaker et al., 1983), *T. reesei* CBHII (Teeri et al., 1987), *T. reesei* EGI (Penttilä et al., 1986), *T. reesei* EGII (Saloheimo et al., 1988), *T. viridae* CBH (Cheng et al., 1990), *Agaricus bisporus* CEL1 (Raguz et al., 1992), *A. bisporus* CEL3 (Chow et al., 1994), *Penicillium janthinellum* CBHI (Koch et al., 1993), *Phanerochaete chrysosporium* CBHI (Sims et al., 1988), *Humicola grisea* CBHI (Azevedo et al., 1990), *H. insolens* EGV (Rasmussen et al., 1991), and *T. reesei* EGV (Saloheimo et al., 1994). Residue numbering used here is the same as in Kraulis et al. (1989). To get the corresponding position in the intact CBHI enzyme, add 461 to the residue number.

Adsorption of the peptides on crystalline cellulose

The binding of the peptides to crystalline cellulose was determined by measuring the equilibrium concentration after adsorption at different starting concentrations. By an iterative method, minimizing the least-square residuals, one-binding-site Langmuir adsorption isotherms were fitted on the data points (Fig. 3). The efficiency of adsorption was characterized by calculating partition coefficients (Klyosov et al., 1986; Beldman et al., 1987) from the initial slopes of the adsorption isotherms by linear regression (inset in Fig. 3). The following values were obtained: wild-type 1.7 L/g, P16R 1.0 L/g, N29A 0.6 L/g, Q34A 0.2 L/g, and Y321A 0.07 L/g. The largest effect of the amino acid substitutions was seen for peptides Y5A and Y32A, which totally lost their affinity to cellulose.

NMR data

The differences in backbone NH and CαH chemical shifts between the wild-type peptide and the analogues for each residue are shown in Figure 4. The $^3J_{HN\alpha}$ coupling constants were measured from COSY spectra and converted to dihedral Φ-angles using the Karplus relation (Karplus, 1963; Pardi et al., 1984). The obtained coupling constants given as three intervals, $^3J_{HN\alpha} \leq 5.5$ Hz, $5.5 \text{ Hz} < {}^3J_{HN\alpha} < 8$ Hz, and $^3J_{HN\alpha} \geq 8$ Hz, are shown in Figure 5A. In Figure 5B sequential NH-NH and C$^\alpha$H-NH NOEs for loop and β-strand regions are shown. Changes in cross peaks between backbone protons (NH-NH, C$^\alpha$H-NH, and C$^\alpha$H-C$^\alpha$H) not present in NOESY spectra of the wild type and the analogues are shown in Figure 6.

Structural characterization of the engineered peptides

Y5A

In this peptide, the chemical shifts of residues in the N-terminus and on the flat face, especially Y32, have changed greatly compared to the changes observed in most of the other analogues. The coupling constants also show more dissimilarities to the wild type than do the other analogues. The $^3J_{HN\alpha}$ coupling constants corresponding to the turns formed by residues 3–6 and 10–13 in the wild type are no longer characteristic for this type of turn, indicating changes in the N-terminus and in the loop between strands β1 and β2. The NOEs shown in Figure 5B support these conclusions (Wagner et al., 1986). In the diagonal plot (Fig. 6), new short-range NOEs are found between protons in the N-terminal part and no new long-range NOEs. In the NOE spectra many long-range NOEs typical for the wild-type structure are missing. An overall loss of structural compactness seems thus to result from the amino acid substitution.

Y32A

The changes in chemical shifts for Y32A are small, less than 0.4 ppm and 0.2 ppm outside the area of substitution. Differences are mainly found adjacent to the substitution and in the N-terminal region. Some new long-range NOEs between strands β1 and β3 are found, indicating changes in the β-sheet. The coupling constants show good agreement with the wild type, also in the turn containing the exchanged residue. Further evidence for the intactness of the turn is seen from the NOEs in Figure 5B. The changes in chemical shifts thus seem to arise from the re-

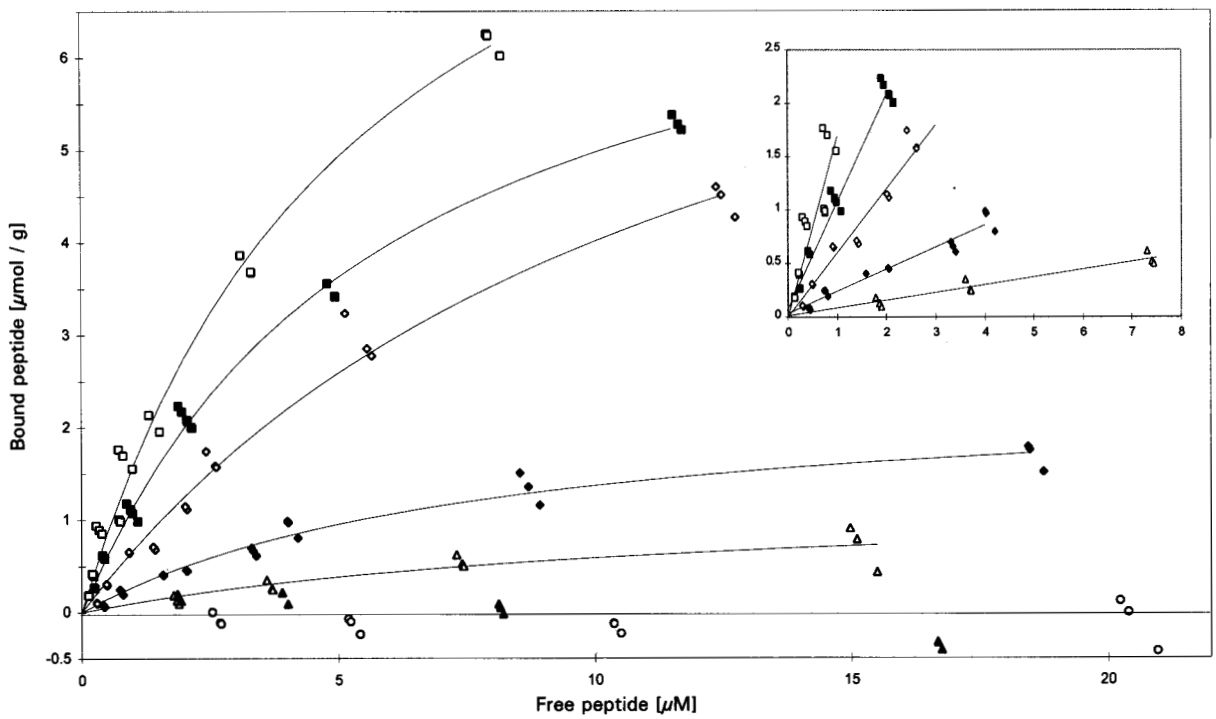

Fig. 3. Adsorption isotherms of the wild-type and engineered peptides. The largest decrease in binding occurred when the tyrosines on the flat face were replaced. □, Wild type; ■, P16R; ◇, N29A; ◆, Q34A; △, Y31A; ▲, Y5A; ○, Y32A. **Inset** shows the initial slopes from which the partition coefficients were calculated.

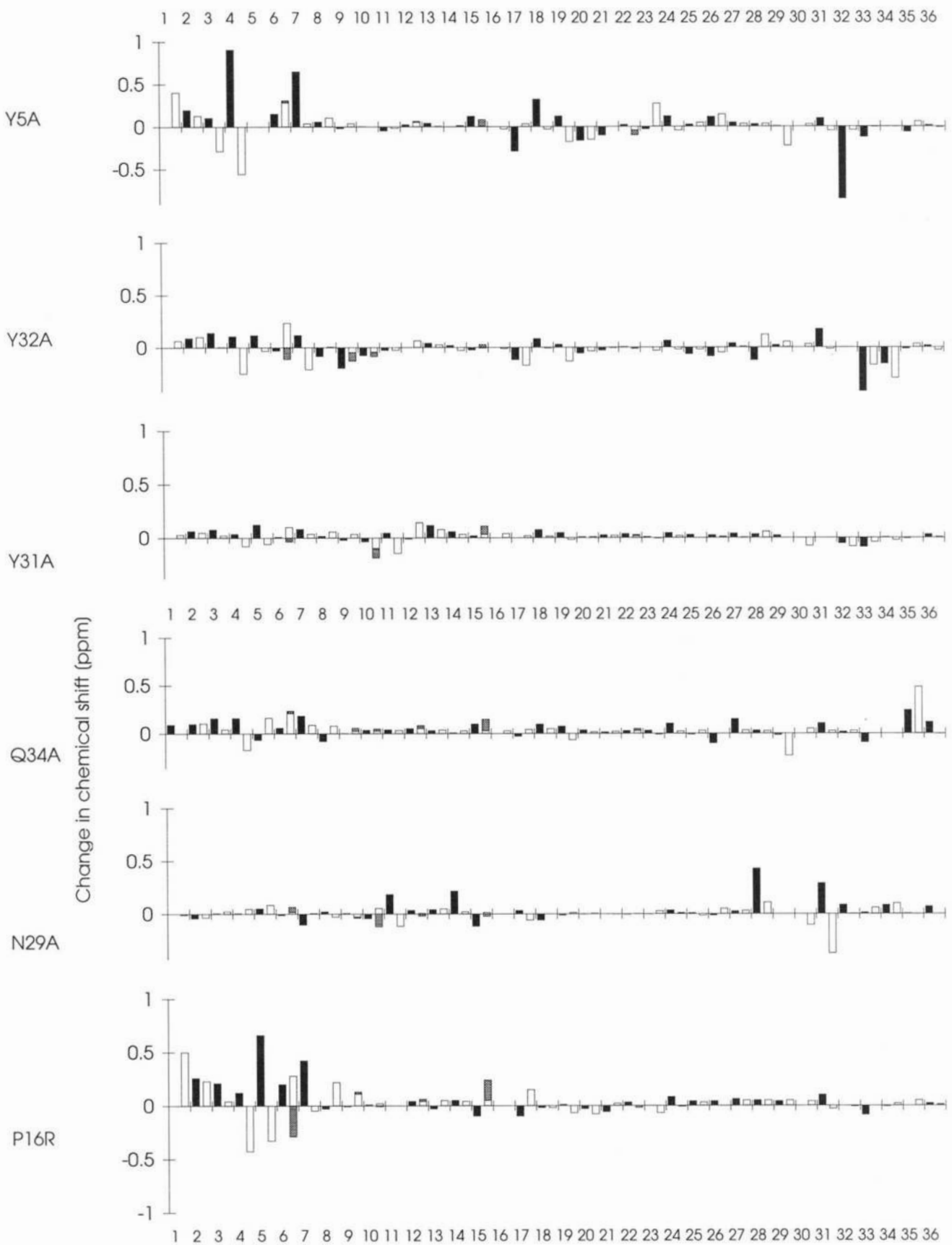

Fig. 4. Differences in chemical shifts of backbone protons between wild-type and engineered peptides. Black bars are NH shifts and white bars are $C^{\alpha}H$ shifts. For glycine residues, white bars are $C^{\alpha 1}H$ shifts and gray bars $C^{\alpha 2}H$ shifts.

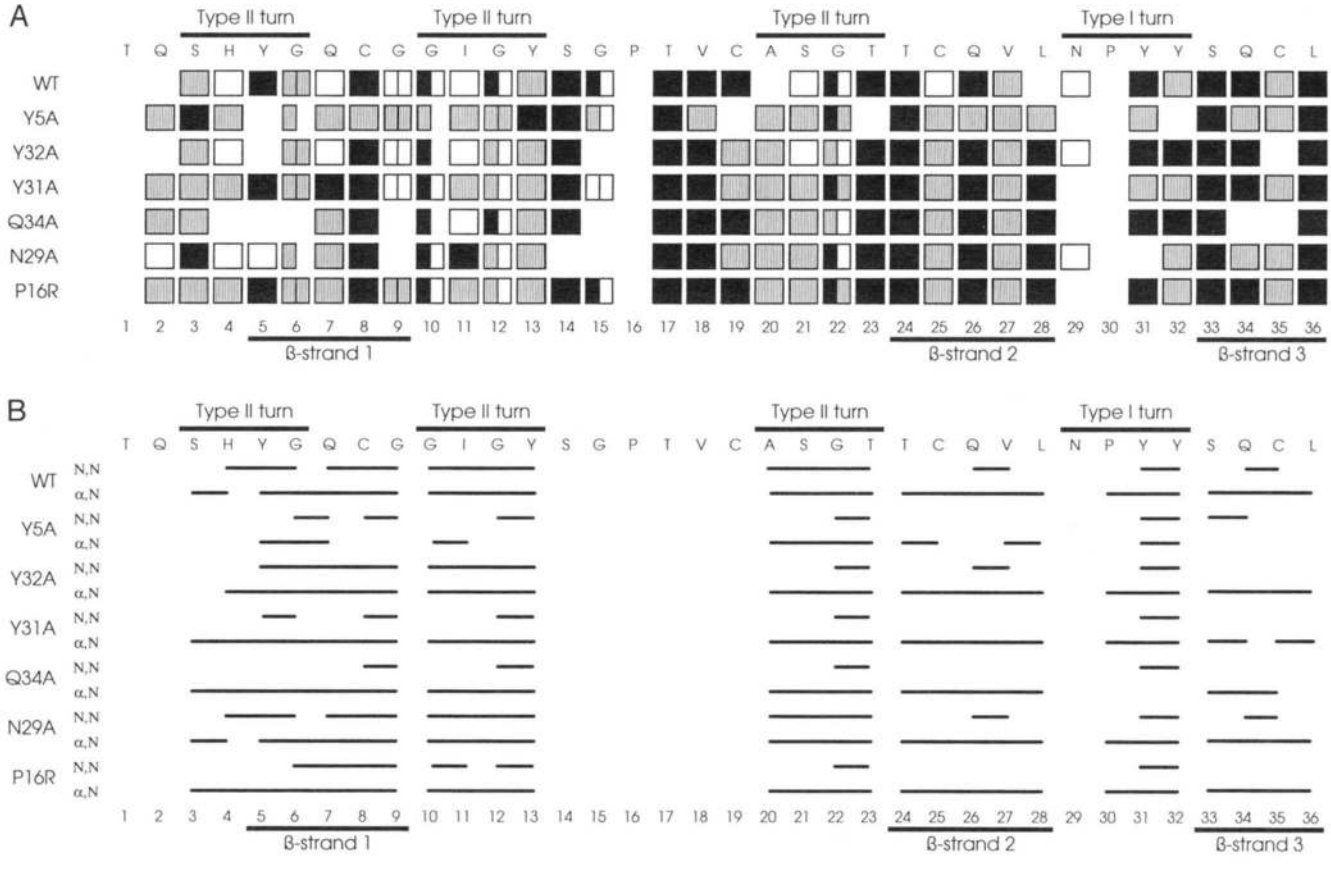

Fig. 5. A: Coupling constants ($^3J_{HN\alpha}$) for wild-type and engineered peptides. Changes in typical patterns for coupling constants in structural elements such as turns and β-strands can be seen in this figure. White boxes, $^3J_{HN\alpha} \leq 5.5$ Hz (corresponding to $\Phi > -70°$); gray boxes, 5.5 Hz $< {}^3J_{HN\alpha} <$ 8 Hz ($-90° < \Phi < -70°$); black boxes, $^3J_{HN\alpha} \geq 8$ Hz ($-150° < \Phi < -100°$). The coupling constant for Y5 in Y31A and P16R is 9 Hz, which corresponds to $\Phi \sim 60°$. **B:** Summary of short-range NOEs involving $C^{\alpha}H$ and NH protons (d_{NN} and $d_{\alpha N}$) in turn and β-strand regions. In the β-strand regions, typical patterns are observed in all cases except for Y5A. The largest changes in turn regions are observed in the N-terminus.

moval of the aromatic ring rather than from the disruption of the structure.

Y31A

The largest changes in chemical shifts, which were less than 0.12 ppm, are observed for the residues that are spatially close to the substituted residue. One new long-range NOE between the β1- and β3-strands indicates a local change near the changed residue, but there is no substantial change in the overall fold of the peptide. The coupling constants show only a few dissimilarities, one of them being at the substituted residue. The turn containing this residue does not seem to be disrupted, because the NOE pattern in this turn is the same as in the wild type. The NMR spectra thus indicate a well-conserved fold and compactness.

Q34A

The differences in chemical shifts in Q34A are moderate along the entire backbone. Some coupling constants in the N-terminus could not be measured, but those measured show large similarities to the wild type. A new long-range NOE between A34 and G6 may be a consequence of the reduced size of the side chain. In all, the NMR data indicate only small conformational perturbations.

N29A

The NMR measurements show the largest differences in chemical shifts adjacent to the mutated residue; namely in L28 and Y31, and some changes in residues 10–15, which are close to Y31 in space, but very small differences in the N-terminus. The $^3J_{HN\alpha}$ coupling constants show moderate differences along the chain, but larger differences in the N-terminus. Several new NOEs were found for N29A. These are all near short distances found in the wild-type structure, with the exception of the NOE between T24 and C19. The loop containing these residues, on one side of the rough face, may thus be somewhat flattened. The presence of additional, not unambiguously assigned, cross peaks suggests that the A29 to P30 peptide bond may also adapt the *cis* conformation leading to increased flexibility in the loop formed by residues 29–32. This could not be confirmed by measuring the exchange peak between conformations at 70 °C due to precipitation at higher temperatures (data not shown). Apart from these findings, the fold of the peptide seems to be correct.

P16R

The changes in the chemical shifts near the N-terminus are greater for this peptide than for the others. The coupling constants also show large differences near the N-terminus. The new

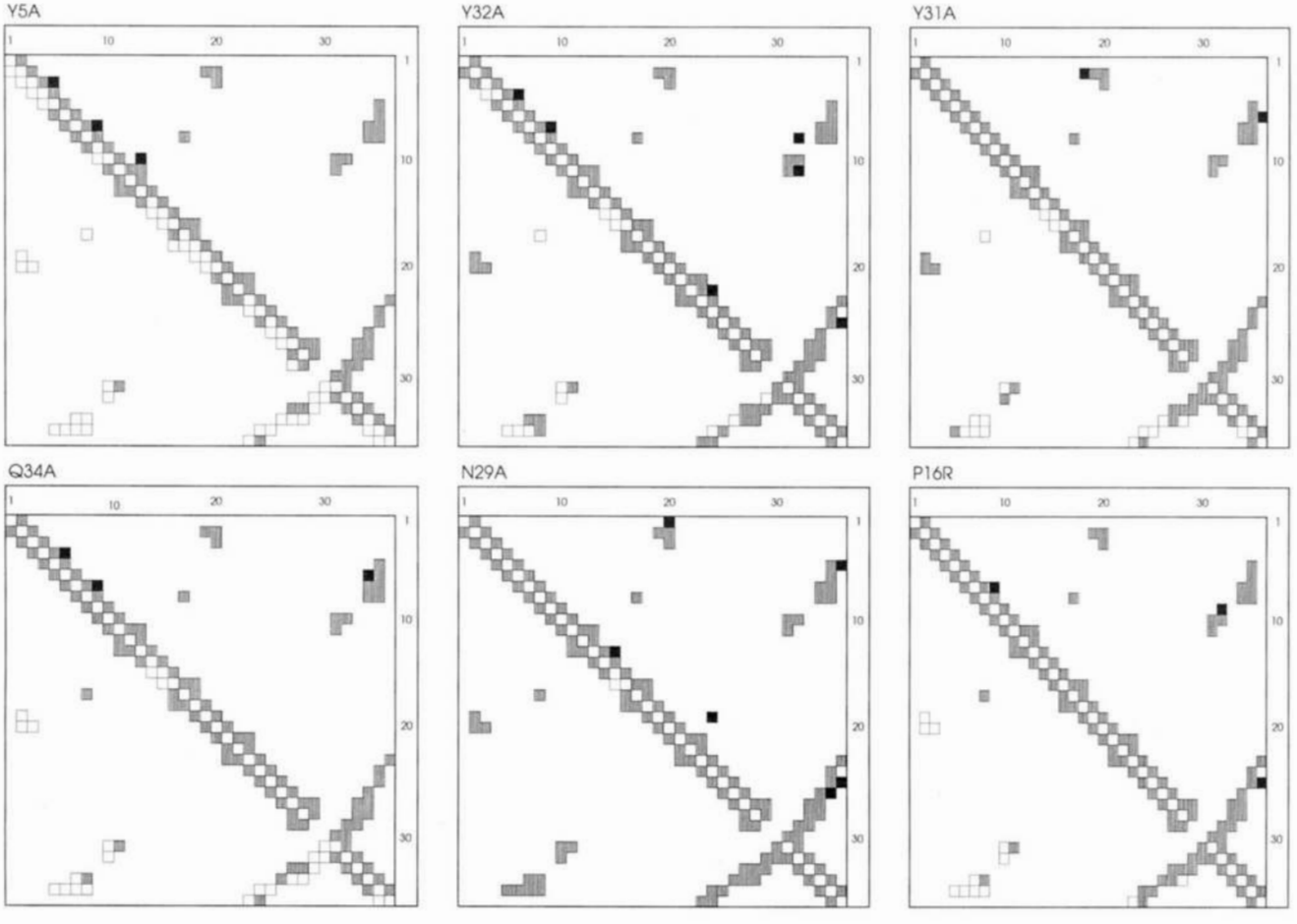

Fig. 6. Diagonal plots of interresidue NOEs between backbone protons (HN, $C^{\alpha}H$; for glycine HN, $C^{\alpha 1}H$, $C^{\alpha 2}H$). Wild-type NOEs are shown as gray boxes. In the upper half of each diagram, new NOEs for the engineered peptides are shown as black boxes. Most new long-range NOEs were observed between strands $\beta 1$ and $\beta 3$ (upper right part of the figure). These indicate changes in the β-sheet, but also verify that the β-sheet structure still exits. The same phenomenon is also observed in Figure 4, where the largest changes in shift are seen in the $\beta 1$- and $\beta 3$-strand residues. In the lower half of each diagram, wild-type NOEs not found in the analogues are shown as white boxes.

NOEs are in sequence close to NOEs observed in the wild type, although the new NOEs to G9 indicate a change in the indention of the rough face. Apart from the changes in the N-terminus, the data indicate a conserved structure. In addition to the backbone NOEs, NOEs to the arginine side chain of P16R were also identified. The β and γ protons showed NOEs to the amide protons of both T17 and V18, but no NOEs were detected for the other side chain protons. This indicates that the side chain points outward from the molecule in the direction of the N-terminus.

Discussion

The structure of the wild-type CBD of *T. reesei* CBHI has been determined by NMR spectroscopy (Kraulis et al., 1989). It is a wedge-shaped molecule with overall dimensions of 30 × 10 × 18 Å (Fig. 1; Kinemage 1). The main secondary structure is an irregular antiparallel triple-stranded β-sheet, where strand $\beta 1$ (residues 5–9) is hydrogen bonded to strand $\beta 3$ (residues 33–36), and this in turn is hydrogen bonded to strand $\beta 2$ (residues 24–28). The structure reveals two distinct faces, one of which is remarkably flat and the other rougher. The molecule is slightly amphiphilic, the flat face being more hydrophilic and the rough more hydrophobic. The amino acid substitutions made in this work did not essentially alter the conformation of the main chain, with the exception of the peptide Y5A. The N-terminal region seems to be the region most sensitive to changes. In the wild-type structure, Y5 is in a type II turn at a position usually occupied by a glycine (Richardson, 1981). Its torsion angles place it far outside of the allowed regions in a Ramachandran plot (Kraulis et al., 1989). This relatively unfavorable conformation is stabilized by a hydrogen bond between the amide hydrogen of Y5 and the CO of C35, and possibly by a histidine–tyrosine interaction (Loewenthal et al., 1992) between H4 and Y5. This special secondary structure seems to be the reason for the sensitivity of this region to amino acid substitutions.

In the binding experiments, the wild-type peptide and the analogues followed Langmuir adsorption isotherms at the measured concentrations (Fig. 3). Extrapolating the curves for wild type, P16R, and N29A gives a saturation value of 9 ± 1 μmol/g, suggesting that the available number of sites is the same and that

it is only the affinity that has been reduced as a result of these substitutions. Unfortunately, extrapolations do not necessarily give a correct value (Klotz, 1989). The difficulties in achieving saturating levels of CBD on cellulose seem to be inherent to this type of studies (Gilkes et al., 1992; Ong et al., 1993). In the case of Q34A and Y31A, estimation of saturating levels is more uncertain because only a weak curvature can be noted in the isotherms. Using an equal number of available sites, we can calculate the difference in free energy of binding from Equation 1:

$$\Delta\Delta G = -RT\ln(K_{mut}/K_{wt}), \quad (1)$$

where K_{mut} and K_{wt} are the partition coefficients of the mutant and wild type, respectively. This gives values of 1.2, 2.4, 4.9, and 7.3 kJ/mol for P16R, N29A, Q34A, and Y31A, respectively. The values for both N29A and Q34A are in the range expected for the loss of a hydrogen bond donor or acceptor (Fersht et al., 1985). The value obtained for P16R is less than the expected contribution of a hydrogen bond.

A striking feature of the flat face is the three exposed tyrosine residues (Kinemage 1). Several findings indicate that these residues are involved in the binding to cellulose. The aromatic function of the residues is strictly conserved in all fungal CBD sequences so far studied (Fig. 2). Stacking of aromatic residues against the faces of sugar rings has been observed in most protein-carbohydrate interactions described (Vyas, 1991; Quiocho, 1993). It has also been shown earlier, by specific nitration, that tyrosines are important for the binding of CBHI to cellulose (Claeyssens & Tomme, 1989), and the importance of Y31 has been demonstrated by site-directed mutagenesis (Reinikainen et al., 1992). Interestingly, the spacing of the tyrosines equals that between every second pyranose ring in crystalline cellulose. Two other strictly conserved residues, N29 and Q34, on the same face, could interact with cellulose through hydrogen bonding. Both asparagine and glutamine are commonly involved in protein-carbohydrate interactions (Vyas, 1991).

The binding properties of the peptides support a model for the binding in which the simultaneous participation of at least Y31, Y32, and Q34 is required. Because of the evident structural changes in the peptide Y5A, unambiguous conclusions cannot be drawn about the function of Y5. Nevertheless, it is likely to be involved in the binding because of its position on the flat face. In addition, the results do not in any way contradict this. The role of N29 remains more difficult to assess. The loss of stability of the loop containing Y31 and Y32 may be sufficient to explain the reduced affinity in this peptide.

In the peptides Y32A and Y31A, the same type of substitution was made, but the adsorption showed a clear difference. Examining the structure of the wild-type CBD reveals that a substitution at position 32 is more likely to cause conformational changes in the neighboring side chains. Thus, we expect that the substitution at position 32 would cause a larger disruption of the flat face, which is in accordance with the experimental results.

Although there is no obvious way in which the rough face of the CBD could interact with cellulose, it does contain a few conserved residues that could participate in hydrogen bonding or hydrophobic interactions (Figs. 1, 2). Interactions of this face with crystalline cellulose have been studied earlier by breaking the surface by an arginine introduced in place of proline-16 by site-directed mutagenesis of the intact CBHI enzyme (Reinikainen et al., 1992). When expressed in yeast, the mutated protein had significantly lowered binding and direct interactions of the rough surface with cellulose could not be ruled out. Here the corresponding mutation was introduced in the peptide P16R. The reduction of binding relative to the wild-type peptide was not as large as would have been expected if the rough face were involved in binding. Structural perturbations of the N-terminal loop, which affect the region around Y5, may well be sufficient to explain the decrease observed. It is therefore very unlikely that the rough face directly interacts with cellulose. If both the flat and the rough faces were involved in the binding we would expect that changes on the flat face causing only small structural perturbations would not decrease the rough face affinity. Thus, a residual affinity due to the rough face adsorption would be seen for the peptides where the flat face was changed. No such affinity can be noted in the isotherms.

As can be seen in Figure 2, there are some additional amino acids that are conserved in homologous sequences but not investigated by substitutions in this study. Glutamine-7, which is partly exposed on the flat face, could well be involved in the interaction with cellulose. The amide nitrogen of this residue forms an intramolecular hydrogen bond with the carbonyl oxygen of glycine-9 and has thus a clear structural role (Kraulis et al., 1989). A substitution of this amino acid to any other would mean a disruption of this hydrogen bond and would make assessment of a possible functional role extremely difficult. The conservation of the adjacent segment 7–10 also indicates structural importance, especially because these residues are mostly buried inside the peptide. The conservation of the aromatic functionality in position 13 suggests an important role for this residue. It is, however, not exposed at the surface but has seemingly an important structural role indicated by a hydrogen bond from the side chain to the amide hydrogen of glycine-15 (Kraulis et al., 1989).

Our data indicate that specific interactions are involved in the binding of the CBD to cellulose and that a correct positioning of the participating side chains on the flat face of the CBD is important. It is not known whether the binding of the CBD can occur at several different types of sites at the cellulose surface. Crystalline cellulose presents at least two surfaces with different characteristics, giving at least two potential classes of sites (Henrissat et al., 1988). We note, however, that both surfaces on cellulose present a very similar geometry and that our results are compatible with a binding to either one or both surfaces.

Materials and methods

Peptide structure

Coordinates for the NMR structure of the CBHI CBD were obtained from the Protein Data Bank (Bernstein et al., 1977), entry p1CBH.

Peptide synthesis and purification

Peptides were synthesized by automated solid-phase synthesis using Fmoc chemistry and purified as described earlier (Lindeberg et al., 1991). The identity of the purified peptides was confirmed by amino acid analysis and time-of-flight plasma desorption mass spectrometry.

Adsorption experiments

Lyophilized peptide was dissolved in 50 mM Na-acetate buffer, pH 5.0, containing 50 mM NaCl to yield a concentration of about 30–40 μM. The concentration of the peptide in the stock solution was determined by amino acid analysis.

Serial dilutions of the peptide solutions were mixed with an equal volume of an aqueous suspension of cellulose from *Acetobacter xylinum* to a final concentration of 1.1 mg/mL. All adsorption experiments were made in triplicate. The suspensions were mixed at 4 °C for 20 h with a magnetic stirrer and centrifuged (3,000 rpm, 10 min). The concentration of the peptide remaining in the supernatant was determined using analytical HPLC. Concentrations were calculated from the integrated peak areas using a calibration curve. The cellulose was grown and prepared as described elsewhere (Herstin, 1963; Gilkes et al., 1992).

Analytical methods

Analytical HPLC was run on a Vydac PRO/PEP C-18 column, with a 0.1% TFA/water:0.09% TFA/acetonitrile gradient from 20 to 30% acetonitrile in 10 min, with detection by UV adsorption at 210 nm.

The amino acid analyses were done on an LKB model 4151 Alpha Plus amino acid analyzer using ninhydrin detection. The samples (20 μg) were hydrolyzed for 24 h at 110 °C with 6 N HCl containing 2 mg/mL phenol and norleucine as an internal standard.

NMR spectroscopy

The peptides were dissolved in ^1H_2O, containing about 7% ^2H_2O for the deuterium lock, to a concentration of 5–10 mg/mL. NaN_3 (0.02%) was added and the pH was adjusted to 3.9. NMR experiments were performed with a Varian Unity 600-MHz spectrometer. All spectra were recorded at 15 °C. Standard pulse sequences and phase cycling were used to obtain COSY (Marion & Wüthrich, 1983), RELAY-COSY (Wagner, 1983) ($\tau = 35$ ms), NOESY (mixing time 200 ms) (Macura & Ernst, 1980), and sensitivity-enhanced TOCSY spectra (Braunschweiler & Ernst, 1983; Bax & Davis, 1985; Cavanagh & Rance, 1990) (mixing times 50, 80, and 120 ms). The spectral width in both dimensions was always 6,400 Hz with 2,048 complex points in each FID. The number of increments varied somewhat but was typically 512 for COSY, 400 for RELAY-COSY, 256 for TOCSY, and 400 for NOESY spectra. Window functions were sinebell for COSY and RELAY-COSY spectra and shifted sinebell square for NOESY and TOCSY spectra in both dimensions. Solvent suppression was achieved by continuous wave presaturation (1.2 s) and residual signal was further reduced by a time-domain data convolution method with a sinebell filter with 20–40 points (Sodano & Delepierre, 1993). Fourier transformation was performed with the program Felix (version 2.3, Biosym Technologies) and resulted in a final data size of 2K × 1K real data points. The sequential assignment was based on both direct (COSY) and relayed through-bond (RELAY-COSY, TOCSY) and through-space (NOESY) connectivities. Spectral assignments were aided by the computer graphics program Felix. $^3J_{HN\alpha}$ coupling constants were determined from COSY spectra by the J-doubling technique (Le Parco et al., 1992; McIntyre & Freeman, 1992).

Acknowledgments

This work was in part supported by the Finnish Academy, the Foundation of Neste Inc., the E. and G. Ehrnrooth Fund, the Swedish National Board for Industrial and Technical Development, the Swedish Research Council for Engineering Sciences, and Nordisk Forskerutdanningsakademi. We thank Dr. David Eaker, Dr. Tuula Teeri, and Professor Olle Teleman for useful discussions and comments on the manuscript. Janne Kerovuo is thanked for his assistance in interpreting NMR spectra.

References

Azevedo M, Felipe MSS, Astolfi-Filho S, Radford A. 1990. Cloning, sequencing and homologies of the *cbh-1* (exoglucanase) gene of *Humicola grisea* var. *thermoidea*. *J Gen Microbiol 136*:2569–2579.

Bax A, Davis DG. 1985. MLEV-17 based two-dimensional homonuclear magnetisation transfer spectroscopy. *J Magn Reson 65*:355–360.

Béguin P, Aubert JP. 1994. The biological degradation of cellulose. *FEMS Microbiol Rev 13*:25–58.

Beldman G, Vorlagen AGJ, Rombouts FM, Searle-van Leeuwen MF, Pilnik W. 1987. Adsorption and kinetic behavior of purified endoglucanases and exoglucanases from *Trichoderma viride*. *Biotechnol Bioeng 30*:251–257.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson 53*:521–528.

Cavanagh J, Rance M. 1990. Sensitivity improvement in isotropic mixing (TOCSY) experiments. *J Magn Reson 88*:72–85.

Cheng C, Tsukagoshi N, Udaka S. 1990. Nucleotide sequence of the cellobiohydrolase gene from *Trichoderma viride*. *Nucleic Acids Res 18*:5559.

Chow CM, Yagüe E, Raguz S, Wood DA, Thurston CF. 1994. The *cel3* gene of *Agaricus bisporus* codes for a modular cellulase and is transcriptually regulated by the carbon source. *Appl Environ Microbiol 60*:2779–2785.

Claeyssens M, Tomme P. 1989. Structure–function relationships of proteins from *Trichoderma reesei*. In: Kubicek CP, ed. *Trichoderma reesei cellulases: Biochemistry, genetics, physiology and application*. Cambridge, UK: Royal Society of Chemistry. pp 1–11.

Coutinho JB, Gilkes NR, Warren RAJ, Kilburn DG, Miller RC. 1992. The binding of *Cellulomonas fimi* endoglucanase C (Cen C) to cellulose and Sephadex is mediated by the N-terminal repeats. *Mol Microbiol 6*:1243–1252.

Fersht AR, Shi JP, Knill-Jones J, Lowe DM, Wilkinson AJ, Blow DM, Brick P, Carter P, Waye MMY, Winter G. 1985. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature 314*:235–238.

Gilkes NR, Henrissat B, Kilburn GD, Miller RC, Warren RAJ. 1991. Domains in microbial β-1,4-glycanases: Sequence conservation, function, and enzyme families. *Microbiol Rev 55*:303–315.

Gilkes NR, Jervist E, Henrissat B, Tekant B, Miller RC, Warren RAJ, Kilburn DG. 1992. The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. *J Biol Chem 267*:6743–6749.

Henrissat B, Vigny B, Buleon A, Perez S. 1988. Possible adsorption sites of cellulases on crystalline cellulose. *FEBS Lett 231*:177–182.

Herstin S. 1963. Bacterial cellulose. *Methods Carbohydr Chem 3*:4–9.

Johansson G, Ståhlberg J, Lindeberg G, Engström Å, Pettersson G. 1989. Isolated fungal cellulase terminal domains and a synthetic minimum analogue bind to cellulose. *FEBS Lett 243*:389–393.

Karplus M. 1963. Vicinal proton coupling in nuclear magnetic resonance. *J Am Chem Soc 85*:2870–2871.

Klotz IM. 1989. Ligand-protein binding affinities. In: Creighton TE, ed. *Protein function*. Cambridge, UK: IRL Press. pp 25–54.

Klyosov AA, Mitkevich OV, Sinitsyn AP. 1986. Role of the activity and adsorption of cellulases in the efficiency of the enzymatic hydrolysis of amorphous and crystalline cellulose. *Biochemistry 25*:540–542.

Koch A, Weigel CTO, Schulz G. 1993. Cloning, sequencing, and heterologous expression of a cellulase encoding cDNA (*cbh1*) from *Penicillium janitellum*. *Gene 124*:57–65.

Kraulis PJ, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J,

Gronenborn AM. 1989. Determination of the three-dimensional structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. *Biochemistry 28*:7241–7257.

Le Parco JM, McIntyre L, Freeman R. 1992. Accurate coupling constants from two-dimensional correlation spectra by "I deconvulsion." *J Magn Reson 97*:553–567.

Lindeberg G, Bennich H, Engström Å. 1991. Purification of synthetic peptides: Immobilized metal ion affinity chromatography. *Int J Pept Protein Res 38*:253–259.

Loewenthal R, Sancho J, Fersht AR. 1992. Histidine-aromatic interactions in barnase. *J Mol Biol 224*:759–770.

Macura A, Ernst RR. 1980. Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy. *Mol Phys 41*:95–117.

Marion D, Wüthrich K. 1983. Two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants. *Biochem Biophys Res Commun 113*:967–974.

McIntyre L, Freeman L. 1992. Accurate measurement of coupling constants by J doubling. *J Magn Reson 96*:425–431.

Ong E, Gilkes NR, Miller RC, Warren RAJ, Kilburn DG. 1993. The cellulose-binding domain (CBD_{Cex}) of an exoglucanase from *Cellulomonas fimi*: Production in *Escherichia coli* and characterization of the polypeptide. *Biotechnol Bioeng 42*:401–409.

Pardi A, Billeter M, Wüthrich K. 1984. Calibration of the angular dependence of the amide proton-C-alpha proton coupling constant, $^3J_{HN\alpha}$, in a globular protein. *J Mol Biol 180*:741–751.

Penttilä M, Lehtovaara P, Nevalainen H, Bhikhabhai R, Knowles J. 1986. Homology between cellulase genes of *Trichoderma reesei*: Complete nucleotide sequence of the endoglucanase I gene. *Gene 45*:253–263.

Quiocho FA. 1993. Probing the interactions between proteins and carbohydrates. *Biochem Soc Trans 21*:442–448.

Raguz S, Yague E, Wood DA, Thurston CF. 1992. Isolation and characterization of a cellulose-growth-specific gene from *Agaricus bisporus*. *Gene 119*:183–190.

Rasmussen G, Mikkelsen J, Schulein M, Hagen F, Hjort C, Hastrup S. 1991. A cellulase preparation comprising an endoglucanase enzyme. *World Pat WO 91/17243*. pp 50–51.

Reinikainen T, Ruohonen L, Nevanen T, Laaksonen L, Kraulis P, Jones TA, Knowles JKC, Teeri TT. 1992. Investigation of the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. *Proteins Struct Funct Genet 14*:475–482.

Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv Protein Chem 34*:167–339.

Saloheimo A, Henrissat B, Hoffrén AM, Teleman O, Penttilä M. 1994. A novel, small endoglucanase gene, *egl5*, from *Trichoderma reesei* isolated by expression in yeast. *Mol Microbiol 12*:219–228.

Saloheimo M, Lehtovaara P, Penttilä M, Teeri TT, Ståhlberg J, Johansson G, Pettersson G, Claeyssens M, Tomme P, Knowles J. 1988. EGIII, a new endoglucanase from *Trichoderma reesei*: The characterization of both gene and enzyme. *Gene 63*:11–21.

Shoemaker S, Schweickart V, Lander M, Gelfand D, Kwok S, Myambo K, Innis M. 1983. Molecular cloning of exo-cellobiohydrolase I derived from *Trichoderma reesei* strain L27. *Bio/Technology 1*:691–696.

Sims P, James C, Broda P. 1988. The identification, molecular cloning and characterization of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exo-cellobiohydrolase I gene from *Trichoderma reesei*. *Gene 74*:411–422.

Sodano P, Delepierre M. 1993. Clean and efficient suppression of the water signal in multidimensional NMR spectra. *J Magn Reson 104*:88–92.

Ståhlberg J, Johansson G, Pettersson G. 1991. A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. *Bio/Technology 9*:286–290.

Teeri TT, Lehtovaara P, Kauppinen S, Salovuori I, Knowles J. 1987. Homologous domains in *Trichoderma reesei* cellulolytic enzymes: Gene sequence and expression of cellobiohydrolase II. *Gene 51*:43–52.

Teeri TT, Reinikainen T, Ruohonen L, Jones TA, Knowles JKC. 1992. Domain function in *Trichoderma reesei* cellobiohydrolases. *J Biotechnol 24*:169–176.

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

Van Tilbeurgh H, Tomme P, Claeyssens M, Bikhabai R, Pettersson G. 1986. Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. *FEBS Lett 204*:223–227.

Vyas NK. 1991. Atomic features of protein-carbohydrate interactions. *Curr Opin Struct Biol 1*:732–740.

Wagner G. 1983. Two-dimensional relayed coherence transfer spectroscopy of a protein. *J Magn Res 55*:151–156.

Wagner G, Neuhaus D, Wörgötter E, Vašák M, Kägi JRH, Wüthrich K. 1986. Nuclear magnetic resonance identification of "half turn" and 3(10)-helix secondary structure in rabbit liver metallothionein-2. *J Mol Biol 187*: 131–135.

Wood TM. 1992. Fungal cellulases. *Biochem Soc Trans 20*:46–53.