Interaction between cellohexaose and cellulose binding domains from *Trichoderma reesei* cellulases

Maija-Liisa Mattinen[a,*], Markus Linder[b], Anita Teleman[a], Arto Annila[a]

[a] *VTT, Chemical Technology, Box 1400, FIN-02044 VTT, Finland*
[b] *VTT, Biotechnology and Food Research, Box 1500, FIN-02044 VTT, Finland*

Received 9 December 1996; revised version received 20 March 1997

Abstract Most *Trichoderma reesei* cellulases consist of a catalytic and a cellulose binding domain (CBD) joined by a linker. We have used cellohexaose as a model compound for the glucose chain to investigate the interaction between the soluble enzyme and cellulose. The binding of cellohexaose to family I CBDs was studied by NMR spectroscopy. CBDs cause line broadening effects and decreasing T_2 relaxation times for certain cellohexaose resonances, whereas there are no effects in the presence of a mutant which binds weakly to cellulose. Yet it remains uncertain how well the soluble cellooligosaccharide mimics the binding of CBD to the cellulose.

Key words: NMR spectroscopy; *Trichoderma reesei* cellulase; Cellulose binding domain; Protein-ligand interaction

1. Introduction

Cellulose has an environmental impact in the preservation of the global carbon cycle and a significant commercial potential as a raw material for many industrial processes. For these reasons it is important to understand the biodegradation of crystalline cellulose on the molecular level, i.e. to relate 3D structures of cellulolytic enzymes and cellulose with activity.

In nature *Trichoderma reesei* secretes cellulases which efficiently hydrolyse crystalline cellulose. These enzymes have a large catalytic domain and a small cellulose binding domain (CBD), which is responsible for most of the affinity of the enzyme towards cellulose. The removal of the CBD has no effect on the capacity of the catalytic domain to hydrolyse small soluble cellooligosaccharides [1–6].

The CBDs of *T. reesei* cellulases share over 70% of amino acid sequence similarity and their 3D structures are very similar too. The structures of CBD_{CBHI} and three of its analogous peptides (Y5A, Y31A, Y32A) have been determined by NMR spectroscopy [7,8]. The structures of CBD_{CBHII} and CBD_{EGI} have been modelled based on the structure of CBD_{CBHI} [9]. The preliminary NMR structure of CBD_{EGI} (our unpublished data) confirms that the model of CBD_{EGI} is correct. In protein-carbohydrate interactions the presence of tyrosine or other aromatic residues at the binding face is common [10]. According to the 3D structures family I CBDs have exposed aromatic side-chains on one face of the domain. The spacing and alignment of the aromatic rings and the adjacent glutamines and asparagines with amide groups are such that multiple interactions with the glucose units on the cellulose crystal are conceivable [5,6,8,11].

In the present work the binding of cellohexaose to CBDs from *T. reesei* cellulases (CBHI, CBHII, EGI) and to one mutant of CBD_{CBHI} (Y31A) has been investigated by ^{1}H-NMR spectroscopy. Cellohexaose is the longest readily available cellooligosaccharide which can be used to mimic the glucose chain of cellulose. Cellooligosaccharides longer than cellohexaose have limited solubility in aqueous buffers.

2. Materials and methods

2.1. Cellohexaose and peptides

The concentrated cellohexaose (Seikagaku Corp., USA) solution was prepared at pH 5 in D_2O. The adsorption of CBD to cellulose is strongest at this pH [6]. The concentration of the soluble cellohexaose was about 0.35 mM. NaN_3 (0.02%) was added as an antimicrobial agent.

CBD_{CBHI} was cloned and produced in *E. coli* [12]. Y31A was synthesised by automated solid-phase synthesis using Fmoc chemistry and purified as described earlier [13]. The primary structures of the peptides were confirmed by amino acid analysis and mass spectrometry. The peptide concentration in the cellohexaose solution at the end of titration was about 250 μM for CBD_{CBHI} and Y31A.

2.2. NMR experiments

The cellohexaose solution (300 μl) was pipetted into a Shigemi NMR tube and the 1D ^{1}H-NMR spectrum of the pure cellooligosaccharide was recorded. Then the peptide was added in 5 μl and 10 μl volumes to obtain titration plots. For the binding constant (K_d) determination the titration was also performed in two constant CBD_{CBHI} concentrations by adding known amounts of cellohexaose. The NMR spectra were acquired with a Varian UNITY 600 Spectrometer. The high field provided with a sufficient resolution for the resonances of the cellooligosaccharide. The digitally filtered spectral width was 8000 Hz, the acquisition time 1.0 s, the 45° pulse 5 μs, and the repetition time 4.0 s. For each spectrum 128 transients were collected. The chemical shifts at 5°C are relative to residual H_2O at 4.85 ppm.

The spectra were deconvoluted into individual Lorentzian lines to determine the full linewidth at half-height. The complete calculated spectrum and each of its individual components were compared with the measured spectrum. Spectra for T_2 analysis were acquired with the CPMG pulse sequence. The 2D NOESY spectrum of the pure cellohexaose and the solution corresponding the highest CBD_{CBHI} concentration in cellohexaose were recorded with standard pulse sequences and phase cycling [14] with a 500 ms mixing time. All experiments were performed at 5 and 20°C.

3. Results

The resonances of the cellohexaose broadened when CBD_{CBHI} was added into the solution (Fig. 1). Increasing amounts of the peptide caused progressively larger line broadening (Fig. 2), but the chemical shifts of the cellooligosaccharide protons were not affected by the additions of CBD. There-

*Corresponding author. Fax: (358) (9) 456 7026.
E-mail: maija.mattinen@vtt.fi

Abbreviations: CBD, cellulose binding domain; CBHI, cellobiohydrolase I; CBHII, cellobiohydrolase II; EGI, endoglucanase I; Y31A, CBD_{CBHI} where tyrosine 31 is replaced by alanine

PII S0014-5793(97)00356-6

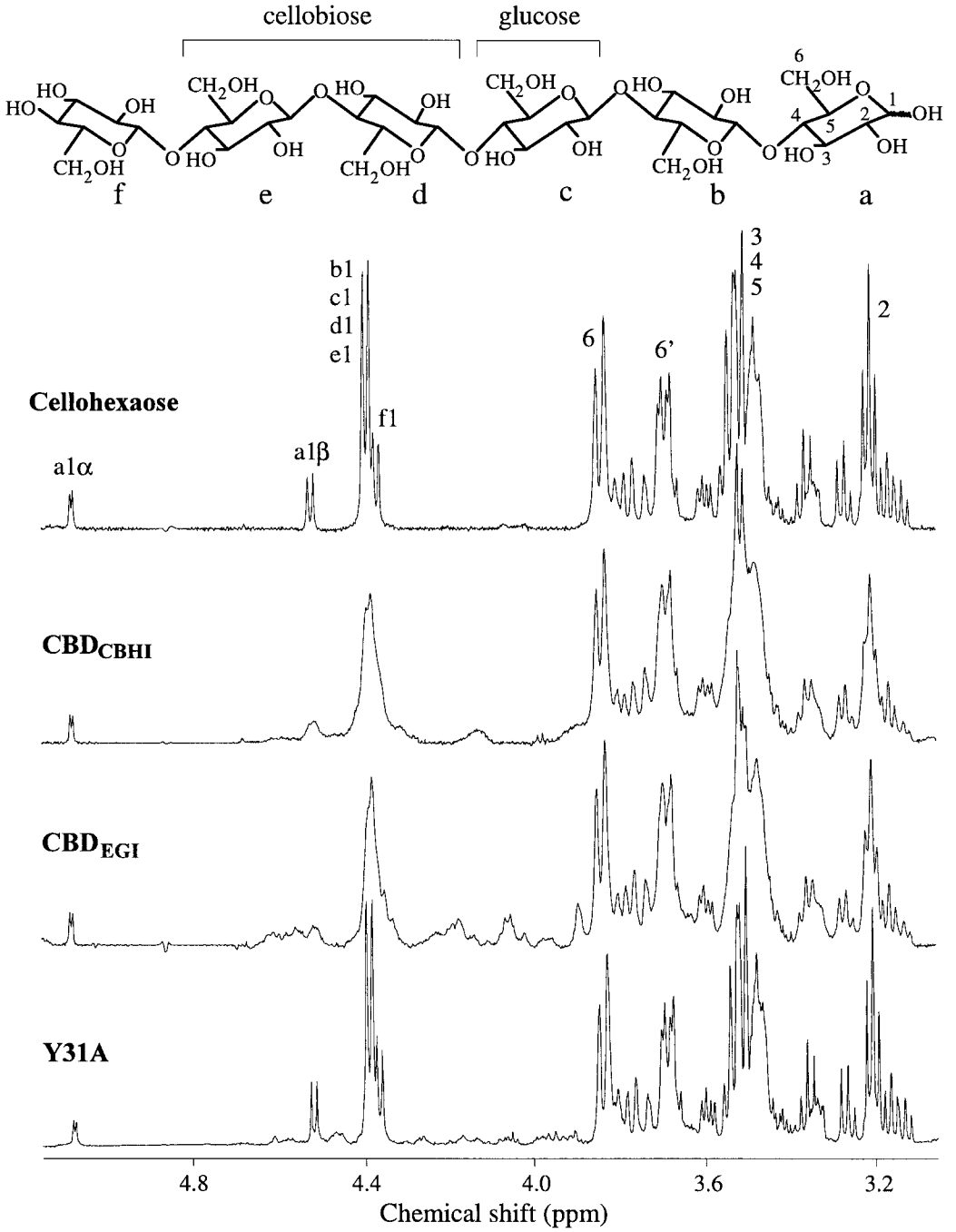

Fig. 1. 1D ^{1}H-NMR spectra of cellohexaose, pure and in the presence of CBDs (CBHI, EGI, and Y31A) at 5°C. Resonance assignments are according to Ikura and Hikichi [30]. The ratio of cellohexaose to CBD is 2:1 for CBD$_{CBHI}$ and Y31A. For CBD$_{EGI}$ the ratio is 3:1.

fore, the free and bound cellohexaose were in a fast exchange. The line broadening phenomenon (Fig. 1) was observed also for CBDs of CBHII (not shown) and EGI, but not for a mutant of CBD$_{CBHI}$ (Y31A) which was previously shown to bind only very weakly to cellulose [11]. The pure amino acids (tyrosine and tryptophan) which were employed as controls did not cause any line broadening effects (not shown). A detailed comparison of the cellohexaose spectra acquired in the presence of different CBDs with the spectrum of the pure cellohexaose revealed, that the line broadening effects were most significant for the a1β protons in the reducing end of the cellohexaose and for the internal and terminal 1 protons (b1, c1, d1, e1, f1). The line broadening effects were also seen in the spectral region of the 3, 4, and 5 protons. Because the lines of these protons overlap in the spectrum of the pure cellohexaose, it was impossible to identify unambiguously which protons caused the line broadening effects in the presence of CBD. For CBD$_{CBHI}$ and Y31A these effects were studied as a function of peptide concentration at 5 and 20°C.

In the case of CBD$_{CBHI}$ the largest line broadening effects were seen for the terminal 1 protons (Fig. 2D). The difference between the linewidths corresponding to the highest peptide concentration and the starting point, where no CBD was added, was 9 Hz at 5°C and 6 Hz at 20°C. The broadening

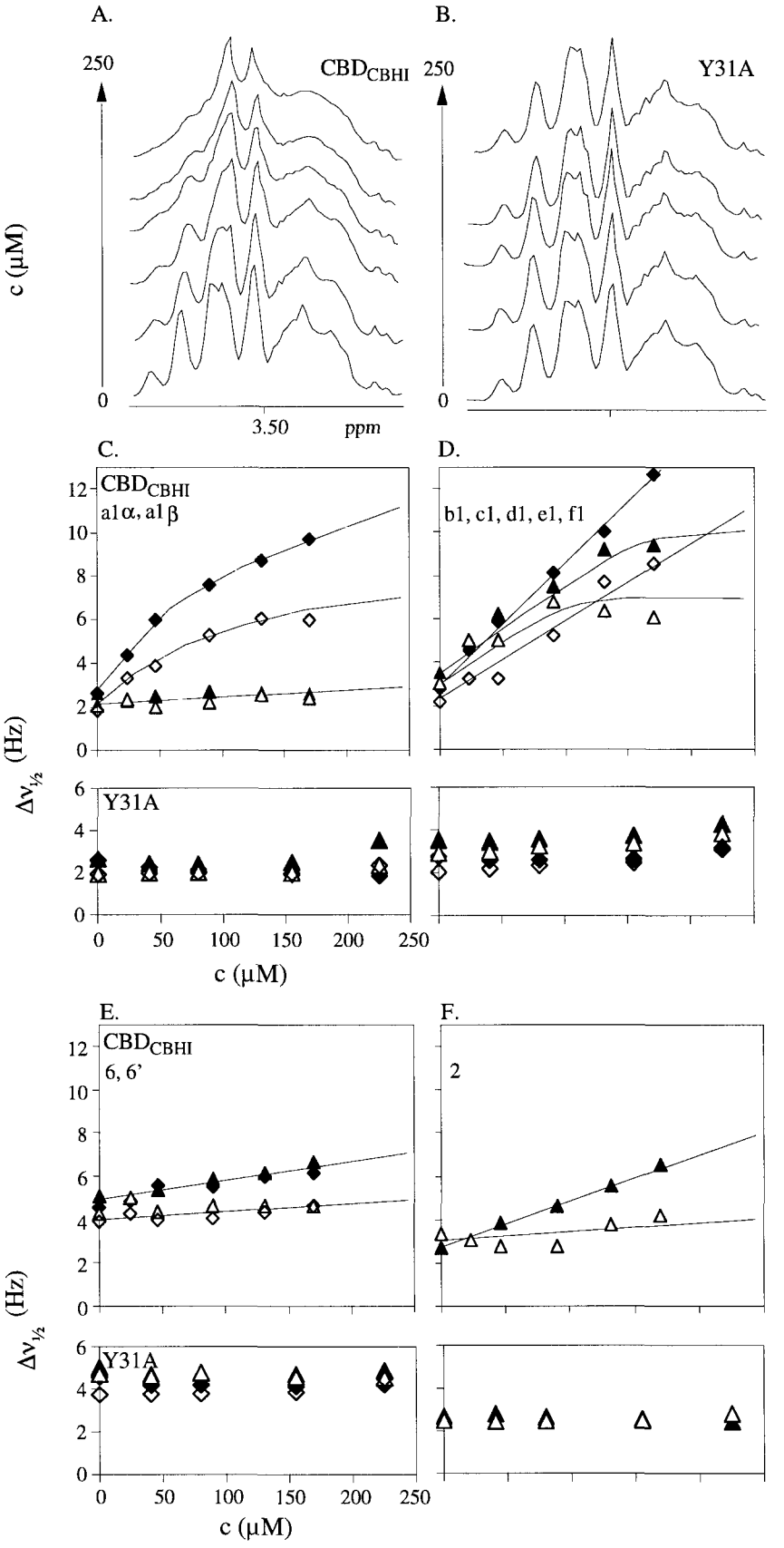

Fig. 2. A series of spectral regions of protons 3, 4, and 5 for (A) CBD_{CBHI} and (B) Y31A acquired at 5°C. Linewidths ($\Delta\nu_{1/2}$) of certain cellohexaose protons as a function of peptide (CBD_{CBHI}, Y31A) concentration: (C) a1α (△) and a1β (◇); (D) b1, c1, d1, e1 (△) and f1 (◇); (E) 6 (△) and 6′ (◇); (F) 2 (△). Solid symbols correspond to experiments at 5°C and open symbols at 20°C. Lines are guides to the eye.

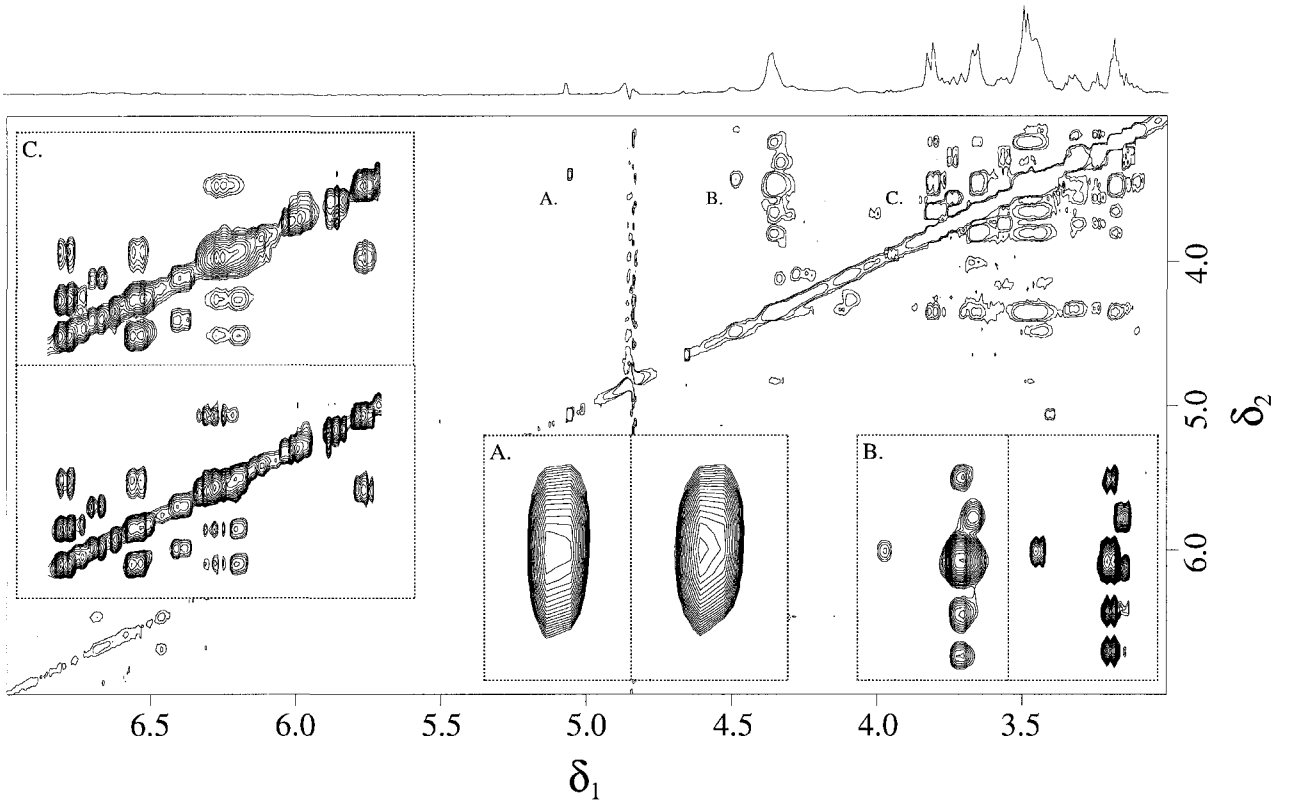

Fig. 3. NOESY spectrum of cellohexaose in the presence of CBD_{CBHI} at 5°C. In the insets certain cross peaks of the cellooligosaccharide are compared with the NOEs obtained from the pure cellohexaose solution using the same experimental conditions. Pure cellohexaose is shown on the right (A and B) and below (C) the insets.

was obvious also for a1β protons (Fig. 2C). These linewidths increased more than 6 Hz at 5°C and about half of that at 20°C. The increase in the linewidth was linear up to 75 μM peptide concentration at both temperatures. For internal 1 protons (Fig. 2D) the effect was of the same order of magnitude as for a1β protons. Also the linewidths of protons 3, 4, and 5 increased, but this spectral region (3.45–3.55 ppm) was too crowded to be interpreted specifically. Therefore, only a series of spectral expansions in this region is shown as a function of peptide concentration (Fig. 2A). For proton 2 (Fig. 2F) the increase in the linewidths was about 3 Hz at 5°C and less than 2 Hz at 20°C. For protons 6 and 6′ (Fig. 2E) the linewidths broadened less than 2 Hz at both temperatures. No line broadening effects were observed for a1α protons (Fig. 2C) in the reducing end of the cellohexaose as a function of peptide concentration at either temperatures. From the cellohexaose titration experiments at constant CBD concentrations the equilibrium binding constant (K_d) was calculated to be 350 ± 90 μM for CBD_{CBHI} at 5°C [15]. The accuracy was limited by the resolution of spectra. For CBD_{CBHII} and CBD_{EGI} the binding constants were comparable with K_d of CBD_{CBHI}.

The same experiments were repeated for the Y31A mutant of CBD_{CBHI} (Fig. 2). No line broadening effects were observed for any of the well resolved cellohexaose resonances. At both temperatures the error of linewidths was at most 1 Hz. For the spectral region of protons 3, 4, and 5 no line broadening effects were observed at high concentrations of Y31A. However, small changes in the spectral appearance were caused by the lines of Y31A.

T_2 relaxation measurements confirmed the linewidth measurements. The effects were observed for all native CBDs but not for Y31A. T_2 rates decreased for a1β, internal and terminal 1 protons on an average to a third of the values of pure cellohexaose. T_2 analysis could not be performed for the resonances of 3, 4, and 5 protons due to the overlap of these lines.

Transferred NOE experiments carried out at 5°C for the solution corresponding to the highest CBD_{CBHI} concentration in cellohexaose (Fig. 3) revealed broadened cross peaks for cellohexaose protons, as expected based on the 1D measurements. Also for CBD_{CBHI} some cross peaks, for example between ring protons of tyrosines, were observed. However, no unambiguous cross peaks between the cellohexaose protons and the ring protons of tyrosines were seen.

We demonstrated that the line broadening effects disappear by adding CBHII enzyme into mixtures of CBD and cellohexaose to produce these degradation products of cellohexaose (data not shown). As a consequence all resonances of the cellooligosaccharides narrowed and were equally narrow as recorded from a reference sample. Simultaneously, the intensity of the lines of the terminal (f1) and reducing end (a1) protons increased and correspondingly the intensity of the lines of the internal protons decreased. Consequently, we conclude that CBD_{CBHI} and CBD_{EGI} do not bind to short soluble cellooligosaccharides like cellotriose and cellobiose.

4. Discussion

So far there is no unambiguously proven model at the res-

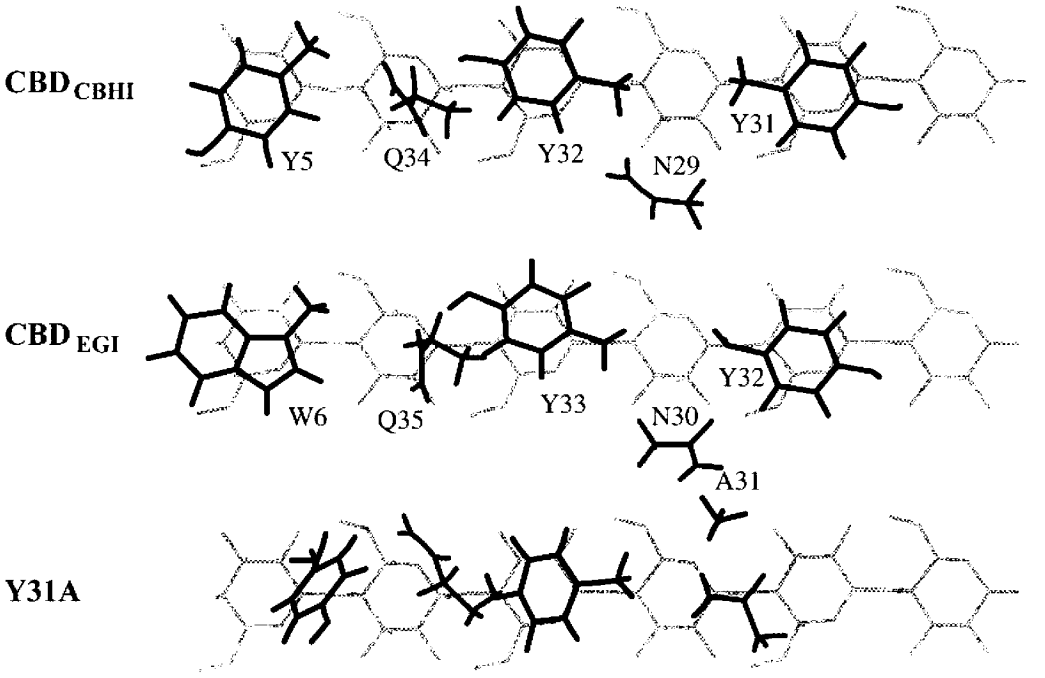

Fig. 4. Binding model. Wild type CBDs (CBHI, EGI) and an engineered peptide of CBD_{CBHI} (Y31A) are positioned on the glucose chain (only six glucose units are drawn) with aromatic rings stacked on top of rings. Only the aromatic residues (Y5, Y32, Y31, Y33, W6), glutamines (Q34, Q35), asparagines (N29, N30) and mutated residue (A31) are shown.

idue level of how CBD interacts with crystalline cellulose and if it can act as a molecular unzipper of the cellulose chains [6]. Although the 3D structures of certain fungal CBDs are known, studies of the interaction are at least partly complicated by the complex structure of cellulose.

The 3D structure of crystalline cellulose has been studied intensively with different methods during several years [16–18]. Recently molecular modelling in combination with atomic force microscopy (AMF) techniques has been used to analyse cellulose surfaces [19]. Cellulose is composed of parallel linear polymeric chains of glucose, which pack together by regular hydrogen bonding networks. The glucose monomers are joined by β-1,4 linkages which create a two-fold screw axis along the polymer. Therefore, the repeating unit is a cellobiose. In addition to the crystalline regions there are crystal defects, dislocations, chain ends, twists, etc., and amorphous regions, which may be initial target sites for cellulases [20].

According to several investigations family I CBDs (CBHI, CBHII and EGI) bind to cellulose via their flat face [5,6,11]. Adsorption measurements in combination with structural analyses of several mutated CBDs of CBHI have lead to the conclusion that the residues Y5, Y31, Y32 and Q34 at the flat face are functionally important. Particularly the number of aromatic residues and their precise spatial arrangement are critical for the specific binding [8,21]. In CBD_{CBHI} the rough face is not directly involved in the function of the CBD [6,11].

The four side-chains at the flat face of CBD_{CBHI} (Y5, Y31, Y32, Q34) and CBD_{EGI} (W6, Y32, Y33, Q35) can be aligned along the cellulose chain (Fig. 4). For CBD_{CBHI} the alignment is within the precision of the structures [7]. A similar alignment is obtained for CBD_{EGI}. For the Y31A mutant of CBD_{CBHI} residues Y5, N29, Y32, and Q34 are along the same line at the flat face. A31 is slightly above and to the side of this alignment, where it should not interfere with the weak but observed binding to the crystalline cellulose. These observations suggest a model of the binding in which the three aligned aromatic rings stack on every other glucose unit of the cellulose polymer. In the cellulose crystal form I, which occurs in nature, this surface is at the obtuse corner [22].

The planar polar amides of the glutamines and asparagines and hydroxyls of the tyrosines are potential for hydrogen bonding with oxygens of glucose. The amides that are off the alignment may assist in the inter polymer hydrogen bond breaking. Such a face-to-face stacking of aromatic residues with glucose and extensive hydrogen bonding of polar groups with cellooligosaccharide hydroxyls have been observed in several protein-carbohydrate interactions [23]. Hydrophobic effects have been found to primarily mediate interactions between cellulose and family IV CBD [24].

The line broadening effects and decreasing T_2 relaxation times of cellohexaose resonances in the presence of native CBDs (CBHI, CBHII and EGI) suggest that cellohexaose may serve at least to some extent as a model compound to mimic the CBD-cellulose interaction. The binding is specific because no effects were observed for the Y31A mutant and for the reference compounds (tyrosine and tryptophan). Furthermore, the short cellooligosaccharides cellobiose and cellotriose do not qualify as model compounds. Consistent with the cellulose structure where there is no α-glycosidic bond the broadening of the a1α line is negligible. Curiously, the line broadening effects for 1, 3, and 5 protons of cellohexaose are substantially larger than that of the 2, 6 and 6′ protons. The line broadening effects of proton 4 cannot be distinguished from the overlapping lines of protons 3 and 5. Concerning the binding model, presented above, it is conceivable that the aromatic rings could preferably align with the hydropho-

bic patches of C1, C3, and C5 of glucose which are larger than the patches of C2 and C4. However, the specific proof for the binding model remains to be obtained, because no unambiguous transferred NOEs were observed between the CBD and the cellohexaose. Since the line broadening effects decreased with increasing temperature it was not possible to choose a temperature high enough, so the cross peaks of the free cellohexaose vanish [15].

We find the observations for cellohexaose-CBD binding somewhat contradict the cellulose-CBD binding. The affinity of the native CBDs towards cellulose follows the order $CBD_{CBHI} \approx CBD_{CBHII} < CBD_{EGI}$ [11,12], but the results with cellohexaose indicate no significant difference between the CBDs. The discrepancies between the results of the soluble model compound and native solid substrate may arise from several sources. The flat face of CBD_{CBHI} differs from that of CBD_{CBHII} and CBD_{EGI} which have tryptophan in the position 5 instead of tyrosine. Therefore, the periodicity of the aromatic rings of CBD_{CBHII} and CBD_{EGI} do not match the hydrophobic patches of the cellooligosaccharide as well as in the case of CBD_{CBHI}.

Recently published data with family II CBD, which adsorbs to both crystalline and amorphous cellulose [25,26], support the role of the tryptophan residues in cellooligosaccharide binding, but the interactions are weak. The results with family IV CBD, which binds solely to amorphous cellulose, reveal that tyrosines rather than tryptophans are involved in the cellooligosaccharide binding [27]. The binding face of one family IV CBD is concave forming a cleft into which the cellooligosaccharide chains, but not crystalline cellulose, can penetrate [28]. For the family IV CBD the structure of the binding site is clearly different from that of CBDs belonging to families I and II, which have three aligned aromatic sidechains on the binding face [7,25]. For the family III CBDs the residues which interact with cellulose are unknown and there is no proof for the specific binding to short cellooligomers [29].

The flexible soluble oligosaccharide fails to provide a distinct rigid surface, where the family I CBD could attach. The strong and specific binding of CBD to the flat and rigid surface of crystalline cellulose rather than to the soluble and flexible cellooligosaccharide polymer to be hydrolysed by the core domain is, of course, precisely how the CBD may assist in the breakdown of crystalline cellulose.

Acknowledgements: This work has been supported by the Academy of Finland. The authors also thank Professor Olle Teleman and Dr. Anna-Marja Hoffrén for providing the model of CBD_{EGI}.

References

[1] H. van Tilbeurgh, P. Tomme, M. Clayssens, R. Bhikhabhai, G. Pettersson, FEBS Lett 204 (1986) 223–227.
[2] T. Teeri, P. Lehtovaara, S. Kauppinen, I. Salovuori, J. Knowles, Gene 51 (1987) 43–52.
[3] P. Tomme, H. van Tilbeurgh, O. Pettersson, J. van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, M. Claeyssens, Eur J Biochem 170 (1988) 575–581.
[4] J. Ståhlberg, G. Johansson, G. Pettesson, BioTechnology 9 (1991) 286–290.
[5] T. Reinikainen, L. Ruohonen, T. Nevanen, L. Laaksonen, P. Kraulis, T.A. Jones, J. Knowles, T. Teeri, Proteins 14 (1992) 475–482.
[6] T. Reinikainen, O. Teleman, T. Teeri, Proteins 22 (1995) 392–403.
[7] P.J. Kraulis, M.G. Clore, M. Nilges, A.T. Jones, G. Petterson, J. Knowles, A.M. Gronenborn, Biochemistry 28 (1989) 7241–7257.
[8] M. Mattinen, M. Kontteli, J. Kerovuo, M. Linder, A. Annila, G. Lindeberg, T. Reinikainen, T. Drakenberg, Prot Sci 6 (1997) 294–303.
[9] A.-M. Hoffrén, T. Teeri, O. Teleman, Prot Eng 8 (1995) 443–450.
[10] N.K. Vyas, Curr Opin Struct Biol 1 (1991) 723–740.
[11] M. Linder, M. Mattinen, M. Kontteli, G. Lindeberg, J. Ståhlberg, T. Drakenberg, T. Reinikainen, G. Pettersson, A. Annila, Prot Sci 4 (1995) 1056–1064.
[12] M. Linder, I. Salovuori, L. Ruohonen, T. Teeri, J Biol Chem 271 (1996) 21268–21272.
[13] G. Lindeberg, H. Bennich, Å. Engström, Int J Peptide Protein Res 38 (1991) 253–259.
[14] A. Macura, R.R. Ernst, Mol Phys 41 (1980) 95–117.
[15] F. Ni, Prog Nuclear Magnet Reson Spectrosc 26 (1994) 517–606.
[16] A.P. Heiner, J. Sugiyama, O. Teleman, Carbohydr Res 273 (1995) 207–223.
[17] J. Sugiyama, R. Vuong, H. Chanzy, Macromolecules 24 (1991) 4168–4175.
[18] R.H. Atalla, D.L. Vanderhart, Science 223 (1984) 238–285.
[19] L. Kuutti, J. Peltonen, J. Pere, O. Teleman, J Microsc 178 (1995) 1–6.
[20] D. Hon, Cellulose 1 (1994) 1–25.
[21] M. Linder, G. Lindeberg, T. Reinikainen, T. Teeri, G. Pettersson, FEBS Lett 372 (1995) 96–98.
[22] K.H. Gardner, J. Blackwell, Biopolymers 13 (1974) 1975–2001.
[23] F.A. Quiocho, Annu Rev Biochem 55 (1986) 287–315.
[24] A.L. Creagh, E. Ong, E. Jervis, D.G. Kilburn, C.A. Haynes, Proc Natl Acad Sci USA (1996) in press.
[25] G.-Y. Xu, E. Ong, N.R. Gilkes, D.G. Kilburn, D.R. Muhandiram, M. Harris-Brandts, J.P. Carver, L.E. Kay, T.S. Havery, Biochemistry 34 (1995) 6993–7009.
[26] N. Din, J.B. Coutinho, N.R. Gilkes, E. Jervis, D.G. Kilburn, R.C. Miller Jr., E. Ong, P. Tomme, R.A.J. Warren, Prog Biotechnol 10 (1995) 261–270.
[27] P.E. Johnson, P. Tomme, M.D. Joshi, L.P. McIntosh, Biochemistry 35 (1996) 13895–13906.
[28] P.E. Johnson, M.D. Joshi, P. Tomme, D.G. Kilburn, L.P. McIntosh, Biochemistry 35 (1996) 14381–14394.
[29] J. Tormo, R. Lamed, A.J. Chirino, E. Morag, E.A. Bayer, Y. Shoham, T.A. Steitz, EMBO J 15 (1996) 5739–5751.
[30] M. Ikura, K. Hikichi, Carbohydr Res 163 (1987) 1–8.