Design of a pH-dependent cellulose-binding domain

Markus Linder[a,*], Tarja Nevanen[a], Tuula T. Teeri[b]

[a] *VTT Biotechnology and Food Research, P.O. Box 1500, FIN-02044 VTT, Finland*
[b] *Royal Institute of Technology, Department of Biotechnology, SE-10044 Stockholm, Sweden*

Received 1 February 1999

Abstract Protein-carbohydrate interactions typically rely on aromatic stacking interactions of tyrosine, phenylalanine and tryptophan side chains with the sugar rings whereas histidine residues are rarely involved. The small cellulose-binding domain of the Cel7A cellobiohydrolase (formerly CBHI) from *Trichoderma reesei* binds to crystalline cellulose primarily using a planar strip of three tyrosine side chains. Binding of the wild-type Cel7A CBD is practically insensitive to pH. Here we have investigated how histidine residues mediate the binding interaction and whether the protonation of a histidine side chain makes the binding sensitive to pH. Protein engineering of the Cel7A CBD was thus used to replace the tyrosine residues in two different positions with histidine residues. All of the mutants exhibited a clear pH-dependency of the binding, in clear contrast to the wild-type. Although the binding of the mutants at optimal pH was less than for the wild-type, in one case, Y31H, this binding almost reached the wild-type level.

Key words: Cellulose-binding domain; Histidine titration; Protein-carbohydrate interaction; Protein engineering; Cellulase; *Trichoderma reesei*

1. Introduction

Cellulose-binding domains (CBDs) are auxiliary domains found in cellulose- and hemicellulose-degrading enzymes. One established function of the CBD is to bring and maintain the cellulase catalytic domain near to the cellulose surface during the course of the hydrolysis. If the CBD is removed from the catalytic domain, a significant decrease in its ability to hydrolyse solid substrates is observed. In addition to its important role for the enzymatic hydrolysis of cellulose, the binding of the CBD itself presents an interesting example of carbohydrate-protein interaction. In our studies we have focused on the cellulolytic system of the filamentous fungus *Trichoderma reesei*. In its very efficient mixture of hydrolytic enzymes, two cellobiohydrolases, Cel7A and Cel6A (formerly CBHI and CBHII [1]), are prevalent. The CBDs found in these and all other fungal enzymes are classified as family 1. The family 1 CBDs consist of less than 40 amino acids, their structure is stabilised by two or three disulphide bridges and they have practically no hydrophobic core [2–4].

Our efforts to understand the nature of the CBD-cellulose interactions have focused on extensive structural and functional characterisation of the Cel7A CBD [5–14]. These studies have revealed that its binding onto the crystalline cellulose surface is governed by a planar strip of three tyrosine residues which stack onto every second glucose ring in a cellulose chain [4,5,11]. Owing to the high density of binding sites, a clear-cut elution of bound CBD is often difficult even though a relatively rapid exchange rate can be experimentally demonstrated [8]. Fungal CBDs are not eluted from cellulose by small soluble oligosaccharides because they do not bind them tightly enough. We have also shown that the binding of the wild-type Cel7A CBD is not sensitive to easily manipulated experimental conditions such as pH [9]. The insight into the binding event gained by our earlier studies led us to attempt to redesign the Cel7A CBD so that the binding event could be controlled by changing the pH.

2. Materials and methods

2.1. Recombinant DNA

The vector initially used to produce the 'double CBD' [7] was used as the starting point for the DNA manipulations. The double CBD consists of the N-terminal CBD of Cel6A with part of its connecting linker region fused to part of the linker and the C-terminal CBD of Cel7A. In order to facilitate the detection and purification of the recombinant proteins the nucleotide sequences coding for the Flag (DYKDDDDK) [15] and histidine (HHHHHH) tags separated by a short linker (GS) were added immediately after the initiator methionine by PCR. The resulting plasmid was called pML1. For clarity, the amino acid numbering of the resulting double CBD corresponds to the one used previously [7], and does not include the N-terminal tags added to this particular construct. The individual numbering here used for the Cel7A CBD follows that published previously [4,5]. Two single mutations, Y5H and Y31H, and one double mutation, H4V:Y5H, were then introduced into the Cel7A CBD using PCR initiated by primers containing the mutated sequences. The PCR products were inserted as *Nco*I-*Bam*HI fragments into the T7 expression vector pET81F1+ [16]. The final constructions were transformed into *Escherichia coli* DH5α, and the corresponding plasmids were called pDT5 (Y5H), pDT6 (H4V:Y5H) and pDT7 (Y31H).

2.2. Protein expression and purification

For protein production, the plasmids containing the wild-type Cel7A CBD and the three different mutated CBDs were transformed to the *E. coli* strain BL21 DE3 pLysS strain of Novagen, USA. The cells were grown in 300 ml LB medium in 2.5 l flasks at 220 rpm to a cell density of 0.4 to 0.6 at 600 nm, after which the protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.5 mM. Following further cultivation of 4–6 h, the cells were harvested by centrifugation (5 min, 6000×*g*, 4°C). The cell pellet was suspended in 20 ml 50 mM Tris pH 7.9 and frozen. Because of the T7 lysozyme present, the cells underwent lysis when thawed. The DNA was broken by sonicating with a sonication tip (MSE instruments) and the sample centrifuged at 20 000×*g* for 15 min. The pellet was repeatedly washed with the same buffer and then dissolved in 2 ml 6 M guanidine HCl in 50 mM Tris pH 7.9. The dissolved pellet was then diluted with 50 ml refolding buffer, 50 mM Tris, 0.2 M NaCl pH 7.9 containing 77 mg reduced glutathione and 31 mg oxidised glutathione. The sample was left over night in a 150 ml open Erlenmeyer bottle at room temperature and stirring slowly with a magnetic bar. Protein precipitates and other solid materials were removed by centrifugation at 47 000×*g* for 30 min. The protein was then run on a nickel-loaded chelating Sepharose FF

*Corresponding author. Fax: +358 (9) 455 2103.
E-mail: markus.linder@vtt.fi

Abbreviations: CBD, cellulose-binding domain

PII: S0014-5793(99)00253-7

(Pharmacia Biotech) column and eluted with 50 mM Tris pH 7.9 containing 0.2 M NaCl and 0.5 M imidazole. The protein was further purified and desalted using reversed-phase chromatography with a 1 cm×25 cm Vydac C4 column and a gradient from 0.1% trifluoroacetic acid to acetonitrile/0.1% trifluoroacetic acid in 300 ml. After evaporation of the solvent from the peak fractions 3 ml 0.2 M HEPES pH 8.5 was added to their residues. To remove the N-terminal tags and obtain the isolated Cel7A CBD, 10 units of trypsin (Sigma T-4019) were added and an incubation at 37°C performed over night. The cleavage products were isolated by repeating the reversed-phase chromatography. The production and purification was followed using SDS-PAGE and Western blotting with the anti-FLAG M3 monoclonal antibody (Eastman-Kodak). As a final step mass spectroscopy (MALDI) was used to verify the identity of each polypeptide.

2.3. Protein labelling

For binding studies the peptides were tritium labelled by reductive methylation essentially as described by [17]. About 1 mg of each lyophilised peptide was dissolved in 1 ml 0.2 M HEPES pH 8.5. The dissolved peptide and all other reagents were kept on ice. Fifty μl of 3.7% formaldehyde was added. One ampoule of tritium-enriched $NaBH_4$ (100 mCi, 12 Ci/mmol; Amersham) was dissolved in 1 ml 0.2 M HEPES pH 8.5 buffer and added to each peptide sample. The tubes were incubated 30 min on ice and subsequently purified by reversed-phase chromatography as described above.

2.4. Binding studies

For binding studies the lyophilised peptide was dissolved in water to a concentration of about 20 μmol/l. The concentrations were determined by UV adsorption at 280 nm (pH 6), using the molar adsorption coefficient of 4080 for the mutants and 5360 for the wild-type. Bacterial microcrystalline cellulose was prepared as in [18] and used as substrate. The cellulose was kept as a suspension (2.71 g/l) in water. The buffers (0.2 M) used were glycine-HCl (pH 2.5, 3.5), acetate (pH 3.6, 4.4, 5.0, 5.1, 5.2, 5.3, 5.4), citrate/phosphate (pH 5.0, 5.4, 5.6, 6.0, 6.4, 6.6, 7.0), HEPES (pH 6.8, 7.2, 7.6, 8.0), glycine-NaOH (pH 8.5, 9.5, 10.5). Bovine serum albumin (1%) was added to all buffers to prevent non-specific binding.

In a typical binding experiment 50 μl cellulose, 37.5 μl buffer and 12.5 μl CBD were mixed in 1.5 ml Eppendorf tubes for 30 min, centrifuged at 20 000×*g* for 5 min, and 25 μl from the supernatant taken for scintillation counting. Binding isotherms were made at pH 8.5 using glycine-NaOH.

3. Results

3.1. Design of the mutations

The binding interaction of the Cel7A CBD is dominated by aromatic stacking of three tyrosine residues with the glucose rings on a cellulose chain (see Fig. 1). Tyrosine residues have a high intrinsic pK_a value and the binding event has been found to be relatively insensitive to changes in pH [9]. In other fungal CBDs, some of the tyrosines are sometimes replaced by tryptophan or phenylalanine residues, but neither is sensitive to pH. Although histidine residues have not been found as part of the binding site of any fungal CBDs, our earlier data indicated that a histidine at position 31 of the Cel7A CBD can at least partly take the binding role of a tyrosine [12]. Since the imidazole side chain of histidine has a pK_a near neutral, we decided to explore the utilisation of histidine protonation in order to make the binding interaction sensitive to pH. The first mutation (Y31H) introduced a histidine in place of the tyrosine at the exposed tip of the CBD (Fig. 1). In many fungal CBDs the tyrosine at position 5 has been replaced by a tryptophan, and here we replaced it by a histidine (Y5H). This single mutation allowed us to explore possible pH effects of a pair of histidines stacking with each other (H5 and the wild-type H4, see Fig. 1). In the wild-type Cel7A CBD, H4 has been assumed to stabilise the conformation of Y5 [5,10], but the corresponding residue has been replaced by a valine in e.g. the endoglucanases Cel6A and Cel12A. The double mutation (H4V:Y5H) was thus made to probe the role of the H4 in combination with the Y5H mutation.

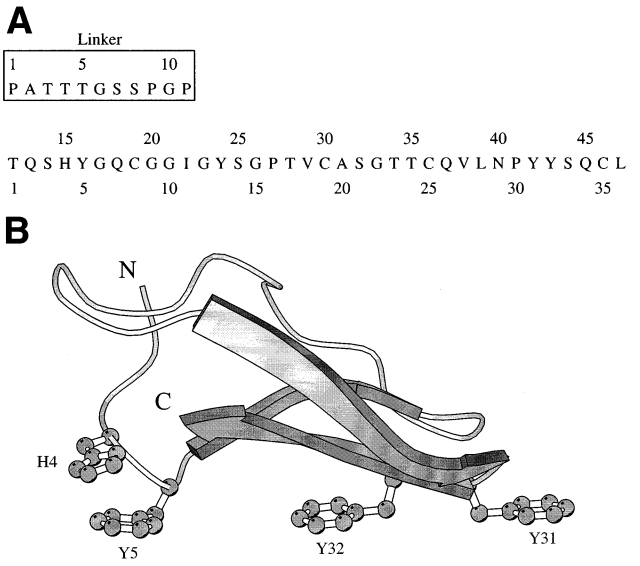

Fig. 1. A: Amino acid sequence of the Cel7A CBD. The protein used in the current study has been obtained by proteolytic cleavage from a double CBD. After the cleavage, 11 residues of the interdomain linker remain attached to the N-terminus of the Cel7A CBD (numbering above the sequence 1–47). For clarity the numbering used to describe the key binding residues in the Cel7A CBD follows that used earlier (numbers below the sequence 1–36). B: The α carbon backbone structure of the Cel7A CBD as determined by NMR [4] (PDB accession: 1cbh). Side chains are drawn for the three tyrosines Y5, Y31 and Y32, which form the cellulose binding face of the CBD. Tyrosine side chains in the positions 5 and 31 were replaced by histidine side chains in the single mutants Y31H and Y5H. In order to avoid stacking of two histidine side chains, the side chain H4 was replaced by a valine residue in the double mutant H4V:Y5H. The structure was drawn using the program Molscript [26].

3.2. Protein production, purification and labelling

The mutations were generated and produced in *E. coli* as described in Section 2. The final yields of cleaved Cel7A CBD and the three mutants were about 3–4 mg from 1 l of culture medium. The experimentally determined mass of every fragment was found to be within 1 Da of that calculated for a fully oxidised fragment (experimental/calculated masses, respectively: Y5H 4669.00/4669.12 Da; Y31H 4670.09/4669.12 Da; H4V:Y5H 4631.18/4631.11 Da). The mutant fragments were different from the wild-type fragment in their retention times in reversed-phase HPLC. The wild-type eluted at 35.4%, H4V:Y5H and Y5H at 34.1% and Y31H at 32.4% acetonitrile. The tritium labelling gave the following specific activities: Y5H 0.45 Ci/mmol, Y31H 4.7 Ci/mmol, H4V:Y5H 1.5 Ci/mmol. The wild-type control used was from a batch of labelled Cel7A CBD (1.38 Ci/mmol) used in an earlier study [8]. The protein had been stored frozen at −70°C and was re-purified by reversed-phase chromatography prior to use. The tritium labelling provides a very convenient method of following the binding and, as shown earlier, does not affect the binding [8,9].

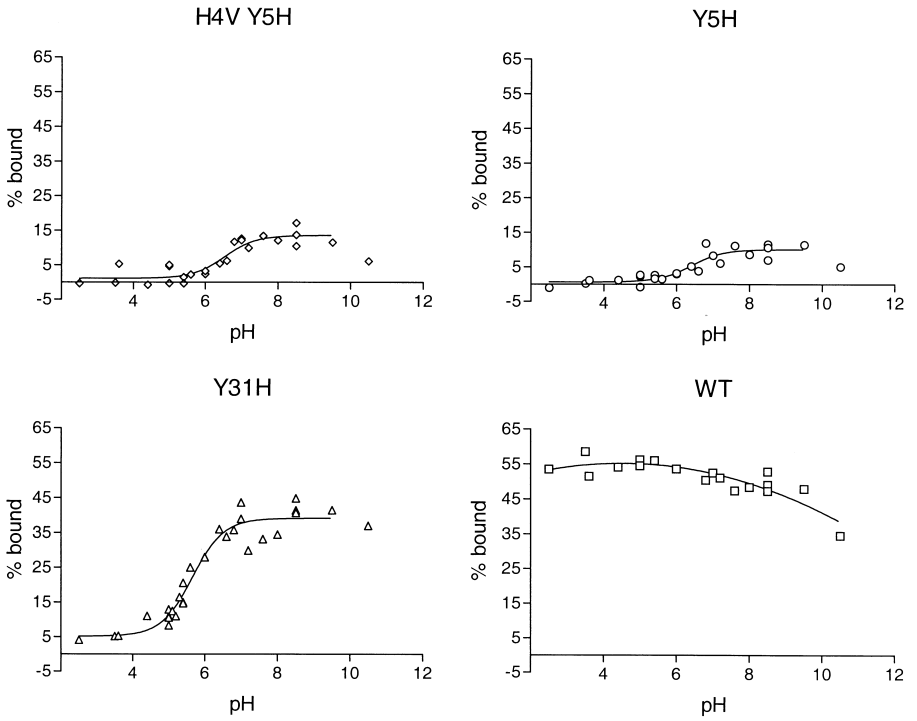

Fig. 2. Titration curves showing the extent of binding at different pHs. For the three mutants the fitted curve is a sigmoidal response curve and for the wild-type a polynomial curve.

3.3. Binding studies

All of the binding experiments in this study were done at low protein concentrations to exclude the possibility for lateral protein-protein interactions affecting the binding [5,12]. To study whether pH influenced the binding of the different CBD variants, the same initial concentration of each peptide was used in a series of binding experiments carried out in buffers at different pHs. As seen in Fig. 2, every mutant was clearly inferior to the wild-type CBD in low pH. However, for all three mutants a clear shift to an improved level of binding was observed at around pH 6. In the case of the mutant Y31H binding approaching that of the wild-type CBD was observed when full binding isotherms were determined at pH 8.5 (Fig. 3). The linearity of these isotherms also shows that lateral protein-protein interactions did not significantly influence binding of the CBDs in these experiments. In agreement with our earlier data [9], binding of the wild-type was not very sensitive to pH and only a small decline in binding was observed in alkaline pH beyond 9–10 (Fig. 2).

4. Discussion

We show here that a simple trick of replacing tyrosine residues with histidine residues at the CBD-binding face can be used to generate peptides that clearly exhibit pH-dependent binding. This is in sharp contrast to the wild-type Cel7A CBD which binds equally well over a wide range of pH values. The titration data show that the mid-points of the titration are about 5.6 (Y31H), 6.3 (Y5H) and 6.4 (H4V:Y5H). The value 5.6 here obtained for the Y31H is somewhat lower than the value of about 6.4 given in literature for model compounds and the pK_a value of 7.1 of imidazole [19,20]. However, the titration value of a residue in a peptide is unavoidably affected by other amino acid residues. As seen in Fig. 1, the histidines at positions 5 and 31 are indeed in very different environments which can influence their apparent pK_a values.

In agreement with our previous results [5,10] we found here that the maximum affinity of the Cel7A CBD was more sensitive to mutations at position 5 than those at position 31. We have shown before that introduction of an alanine at position 5 results in a distorted CBD structure, whereas a corresponding mutation at position 31 hardly affects the structure [10]. We therefore assume that the reduced binding here observed for the position 5 mutants can also be largely explained by structural effects. Stacking of the two histidines in the Y5H mutant seems to reduce the binding more than the introduction of a single histidine in H4V:Y5H although the effect of the pH was very similar for these two mutants. All variants, including the wild-type CBD had decreased binding at pH over 10, presumably owing to the ionisation of the tyrosine side chains at this pH [20].

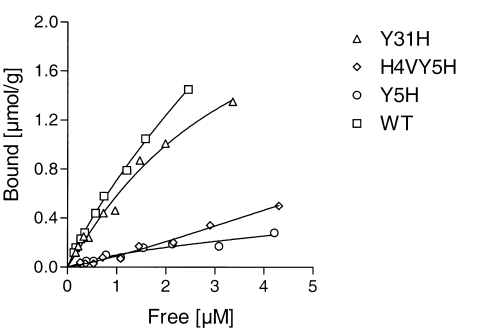

Fig. 3. Binding isotherms determined at pH 8.5, showing the relative decrease in affinity due to the different mutations. The fitted curves are hyperbolas (simple one site binding).

The best combination of pH-dependent but nevertheless efficient binding was achieved in the mutant Y31H. Studies of many different protein-carbohydrate interactions indicate a stacking interaction of the sugar rings and the aromatic side chains of tyrosines, phenylalanines or tryptophans [21–23]. Histidine residues are almost never found in this kind of stacking, and only in one CBD family a histidine side chain seems to be able to replace one tyrosine or tryptophan residue on the cellulose-binding face [24,25]. This is probably because of differences in the nature of its π system, resulting from its heterocyclic composition. On the other hand, our finding that the histidine side chain in Y31H can function in binding nearly as well as the tyrosine it replaced indicates that the involvement of π-electrons in the wild-type protein-carbohydrate interaction can also occur with the imidazole ring. It is possible that in nature, histidine residues are not chosen for protein-sugar interactions simply because they fail to provide a general solution for binding over a range of conditions. However, pH-dependent binding/elution of an engineered CBD can provide an easy elution system in chromatographic and other practical applications relying on CBDs.

Acknowledgements: We wish to thank Dr. Olle Teleman for fruitful discussions and the Academy of Finland for financial support. Dr. John Londesborough is thanked for linguistic revision of the manuscript.

References

[1] Henrissat, B., Teeri, T.T. and Warren, R.A. (1998) FEBS Lett. 425, 352–354.
[2] Linder, M. and Teeri, T.T. (1997) J. Biotechnol. 57, 15–28.
[3] Teeri, T.T. (1997) Trends Biotechnol. 15, 160–167.
[4] Kraulis, P.J., Clore, G.M., Nilges, M., Jones, T.A., Pettersson, G., Knowles, J. and Gronenborn, A.M. (1989) Biochemistry 28, 7241–7257.
[5] Linder, M., Mattinen, M.L., Kontteli, M., Lindeberg, G., Ståhlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G. and Annila, A. (1995) Protein Sci. 4, 1056–1064.
[6] Linder, M., Lindeberg, G., Reinikainen, T., Teeri, T.T. and Petterson, G. (1995) FEBS Lett. 372, 96–98.
[7] Linder, M., Salovuori, I., Ruohonen, L. and Teeri, T.T. (1996) J. Biol. Chem. 271, 21268–21272.
[8] Linder, M. and Teeri, T.T. (1996) Proc. Natl. Acad. Sci. USA 93, 12251–12255.
[9] Linder, M., Nevanen, T., Söderholm, L., Bengs, O. and Teeri, T.T. (1998) Biotechnol. Bioeng. 60, 642–647.
[10] Mattinen, M., Kontteli, M., Kerovuo, J., Linder, M., Annila, A., Lindeberg, G., Reinikainen, T. and Drakenberg, T. (1997) Protein Sci. 6, 294–303.
[11] Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T.A., Knowles, J.K.C. and Teeri, T.T. (1992) Proteins 14, 475–482.
[12] Reinikainen, T., Teleman, O. and Teeri, T.T. (1995) Proteins 22, 392–403.
[13] Mattinen, M., Linder, M., Drakenberg, T. and Annila, A. (1998) Eur. J. Biochem. 256, 279–286.
[14] Mattinen, M., Linder, M., Teleman, A. and Annila, A. (1997) FEBS Lett. 407, 291–296.
[15] Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L. and Conlon, P.J. (1988) Biotechnology 5, 1204–1210.
[16] Tanhauser, S.M., Jewell, D.A., Tu, C.K., Silverman, D.N. and Laipis, P.J. (1992) Gene 117, 113–117.
[17] Tack, B.F., Dean, J., Eilat, D., Lorenz, P.E. and Schechter, A.N. (1980) J. Biol. Chem. 255, 8842–8847.
[18] Gilkes, N.R., Jervis, E., Henrissat, B., Tekant, B., Miller Jr., R.C., Warren, R.A. and Kilburn, D.G. (1992) J. Biol. Chem. 267, 6743–6749.
[19] Tanford, C. (1962) Adv. Protein Chem. 17, 69–165.
[20] Kyte, J. (1995) Structure in Protein Chemistry, Garland, New York, 606 pp.
[21] Quiocho, F.A. (1993) Biochem. Soc. Trans. 21, 442–448.
[22] Vyas, N.K. (1991) Curr. Opin. Struct. Biol. 1, 732–740.
[23] Elgavish, S. and Shaanan, B. (1997) Trends Biochem. Sci. 22, 462–467.
[24] Tormo, J., Lamed, R., Chirino, A.J., Morag, E., Bayer, E., Shoham, Y. and Steitz, T. (1996) EMBO J. 15, 5739–5751.
[25] Sakon, J., Irwin, D., Wilson, D.B. and Karplus, P.A. (1997) Nature Struct. Biol. 4, 810–817.
[26] Kraulis, P. (1991) J. Appl. Cryst. 24, 946–950.