Studies of cellulose binding by cellobiose dehydrogenase and a comparison with cellobiohydrolase 1

Gunnar HENRIKSSON*, Andres SALUMETS*‡, Christina DIVNE† and Göran PETTERSSON*§

*Department of Biochemistry, University of Uppsala, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden, and †Department of Molecular Biology, University of Uppsala, Uppsala, Sweden

The binding isotherm to cellulose of cellobiose dehydrogenase (CDH) from *Phanerochaete chrysosporium* has been compared with that of cellobiohydrolase 1 (CBH 1) from *Trichoderma reesei*. CDH binds more strongly but more sparsely to cellulose than does CBH 1. In a classical Scatchard analysis, a better fit to a one-site binding model was obtained for CDH than for CBH 1. The binding of both enzymes decreased in the presence of ethylene glycol, increased in the presence of ammonium sulphate and was unaffected by sodium chloride. Attempts to localize the cellulose-binding site on CDH have also been made by exposing enzymically digested CDH to cellulose and isolating the cellulose-bound peptides. The results suggest that the cellulose-binding site is located internally in the amino acid sequence of CDH.

INTRODUCTION

Specific binding to cellulose has been reported for several cellulases, hemicellulases and cellulosome-forming proteins [1–6] as well as for one β-glycosidase [7]. Many bacterial and fungal cellulases have a characteristic molecular organization, with a catalytic domain and a separate non-catalytic cellulose-binding domain (CBD) interconnected via a linker peptide [2,6]. This two-domain structure provides an elegant solution to the problems associated with catalytic action on a surface rather than in solution [3]. A similar organization has been reported for glucoamylases from *Aspergillus niger* [8] and might also be valid for chitinases [9]. Instead of having an independently folded binding domain, cellulases from anaerobic fungi such as *Neocallimastix* and *Piromyces* and the bacterium *Clostridium thermocellum* form aggregates in which cellulose binding is often mediated by separate proteins [10–13].

Fungal CDBs all belong to the same family (type 2 in the classification of Béguin and Aubert [6], group B in the classification of Gilkes et al. [2], and family I in that of Coutinho et al. [14]). They are approx. 35 residues long and are located either at the N-terminus or at the C-terminus of the mature proteins [6]. The three-dimensional solution structure of a synthetic CBD, corresponding to residues 462–497 of the fungal cellulase cellobiohydrolase 1 (CBH 1) from *Trichoderma reesei*, has been determined by means of homonuclear two-dimensional NMR spectroscopy [15]. From the structure and studies of synthetic CBD mutants [16] it was concluded that tyrosine residues are likely to be important for binding to the cellulose surface. Two conserved disulphide bridges are also characteristic but these are probably needed for structural integrity of the domains. Bacterial CBDs can be classified into at least four different families [2,6,14]. They are generally larger than the fungal CBDs, varying in length from 63 to 240 residues, and are located either at one of the termini or internally [6]. A disruptive

effect on cellulose fibres has been observed for the CBD of endoglucanase A from *Cellulomonas fimi* [17], and in this type of CBD (type 1 [6] or group A [2]), tryptophan residues are important for binding [18].

In some cases it has been possible to study isolated CBDs obtained by proteolysis [17,19], by genetic construction [1,18] or by synthesis [20]. The binding of *T. reesei* CBH 1 to cellulose is probably the most extensively studied [3,16,21,22]. Removal of the CBD from CBH 1 greatly diminishes both the cellulose binding and the cellulolytic activity of the enzyme [3,23]. Results of binding experiments with the complete enzyme and its isolated catalytic domain and CBD all fit better to a multiple-binding-site model [24] than to a one-site model in a classical Scatchard analysis. This does not necessarily indicate that the cellulose molecules have different classes of binding site for the enzyme, because overlapping binding sites might be an explanation for this phenomenon [24]. In the intact enzyme, however, the two-domain structure of CBH 1 might explain the results [22]. The three-dimensional structure determination of *T. reesei* CBH 1 [25] has revealed that the catalytic domain contains a cellulose-binding tunnel 50 Å (5 nm) long designed to bind cellulose chains of up to 10 glucose units in length (C. Divne, unpublished work). Furthermore the choice of cellulose is important because the degree of crystallinity varies. In this study we have used cellulose from *Acetobacter xylinum*, a cellulose considered to be relatively homogeneous.

Cellulose binding has also been reported for the non-hydrolytic haemoflavoenzyme cellobiose dehydrogenase (CDH), formerly called cellobiose oxidase, from the white-rot fungus *Phanerochaete chrysosporium* [26,27]. CDH is an extracellular enzyme (molecular mass approx. 89 kDa [28], 754 amino acid residues [29,30]) that is produced under cellulolytic conditions by a number of wood-degrading fungi [31–33]. It carries one FAD cofactor and one cytochrome *b*-type haem as prosthetic groups [34], each in a separate domain. The two domains can be

separated by proteolytic cleavage; the individual domains have been shown to retain their respective activities [26]. CDH oxidizes cellobiose, cellodextrins and other soluble saccharides to their corresponding lactones [26,35,36], and possible electron acceptors for the enzyme include quinones [26,35,36], O_2 [35], aromatic cation radicals [37], I_3^- [38] and ferricyanide [39]. The FAD-containing domain of CDH has been shown to contain the active site and the cellulose-binding site [26]. The fact that the sites do not coincide [26] has led to the suggestion that binding of cellulose by CDH might be due to a CBD.

The biological function of CDH is not clear. However, synergy
with cellulases in the breakdown of cellulose has been demonstrated [40], and it has recently been shown that, in the presence of hydrogen peroxide and Fe^{III}, CDH can degrade carboxymethylcellulose, xylan and lignin [41]. Co-operation with the lignolytic enzyme manganese peroxidase has also been suggested ([42]; reviewed in [32,33]). Moreover, CDH has been reported to bind to the fungal cell wall [43] and thus it has been proposed that the cellulose binding is unintended rather than a natural function.

In this work we have compared the cellulose binding of CDH with that of *T. reesei* CBH 1, performed a screening of binding to a number of insoluble polysaccharides, and investigated the physical nature of cellulose binding by adding salts and ethylene glycol to the solvents. We have also made attempts to identify the cellulose-binding site.

MATERIALS AND METHODS

Materials

CDH was purified as described previously [26,44], and CBH 1 as described by Bhikhabhai et al. [45]. Birchwood xylan, crab-shell chitin and potato starch were purchased from Sigma (St. Louis MO, U.S.A.). Cellodextrin mixture II was obtained from Merck (Darmstadt, Germany) and reduced to alditols (the aldehyde group reduced to a primary alcohol) as described in [46]. *A. xylinum* cellulose was prepared as described previously [47,48]. Insoluble mannan (ivory nut) was from Megazyme (Sidney, Australia). Ether-extracted birchwood powder was a gift from Bert Pettersson (STFI, Stockholm, Sweden). Endoglycosidase H was from New England Biolabs (Beverly, MA, U.S.A.). Other chemicals were of analytical grade. The computer program Ultrafit was from Biosoft (Cambridge, U.K.). An extracellular protease from a strain of *Arthrobacter* species was prepared as described by Hofsten et al. [49].

Deglycosylation

CDH is probably mainly N-glycosylated [29]. CDH (0.4 mg) was mixed with 5 units of endoglycosidase H in 50 mM sodium citrate, pH 5.5, and incubated at 37 °C for 36 h. Deglycosylated CDH was purified by ion-exchange chromatography with a MonoQ column (Pharmacia, Uppsala, Sweden) and eluted with a sodium acetate (pH 4.0) gradient (0.040–1 M). Compared with non-deglycosylated CDH, enzyme treated with endoglycosidase H had a smaller apparent molecular mass [29] and formed a sharper band on a native PAGE gel, indicating the greater homogeneity of the material.

Binding studies

Binding experiments were performed at room temperature in 50 mM ammonium acetate, pH 5.0, under constant agitation with a magnetic stirrer for 1 h (with the exception of the kinetic experiment) for CDH, and for 30 min for CBH 1. The poly-

saccharide concentration was 1.2 mg/ml (1 mg/ml for the kinetic experiment). The samples were centrifuged and the supernatants collected. Concentrations of CBH 1 were determined spectrophotometrically at 280 nm. Except for the kinetic experiment, where the absorbance was measured at 420 nm, concentrations of CDH were determined with the standard activity assay [26,50]. The program Ultrafit was used for non-linear regression calculations.

Preparation of cellulose-binding peptides of CDH

CDH (2.5 mg) was transferred to 25 mM sodium phosphate, pH 7 with a PD-10 disposable column (Pharmacia, Sweden). A protease from *Arthrobacter* [49] was added to CDH in proportions of 1:30 by mass and the mixture was incubated for 15 min at room temperature, after which the pH was adjusted to 4 with acetic acid to terminate the proteolysis. Cellulose (2.1 mg/ml suspended in water) was added and the mixture was incubated under constant agitation for 15 min. The sample was centrifuged and the pellet was washed twice with 10 mM ammonium acetate, pH 4. Cellulose-bound material was eluted with 80 % (v/v) formic acid, freeze-dried and separated by SDS/PAGE [51]. Bands visible after staining with Coomassie Blue were cut out, subjected to trypsin and sequenced in an Applied Biosystems 470A gas–liquid phase sequencer equipped with an Applied Biosystems 120A phenylthiohydantoin analyser [52].

RESULTS

In Figure 1 the binding isotherm for cellulose of CDH is compared with that of CBH 1. When fitted to the one-binding-site model with the program Ultrafit, the relative mismatch of the CDH data was considerably lower than that of the CBH 1 data. The data for CDH followed a straight line in a Scatchard plot, whereas those for CBH 1 did not (Figure 2). Both the dissociation constant and the capacity were lower for CDH than for CBH 1 (K_d 0.64 mM, capacity 2.1 mmol/g for CDH; K_d 3.6 mM, capacity 5.1 mmol/g for CBH 1). The addition of ethylene glycol decreased the binding, whereas $(NH_4)_2SO_4$ increased the binding for both CDH and CBH 1. NaCl at concentrations up to 1 M did not influence the cellulose binding of either enzyme (Figure 3). None of the above substances influenced the CDH activity [26] to any measurable extent. The hydrolysis of *p*-nitrophenyl lactoside [53] by CBH 1 was not influenced by NaCl or $(NH_4)_2SO_4$. Preincubation in 80 % (v/v) ethylene glycol decreased the activity of CBH 1 only slightly. Equilibrium in
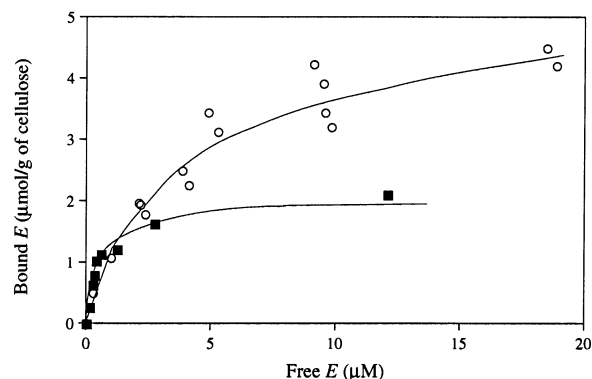

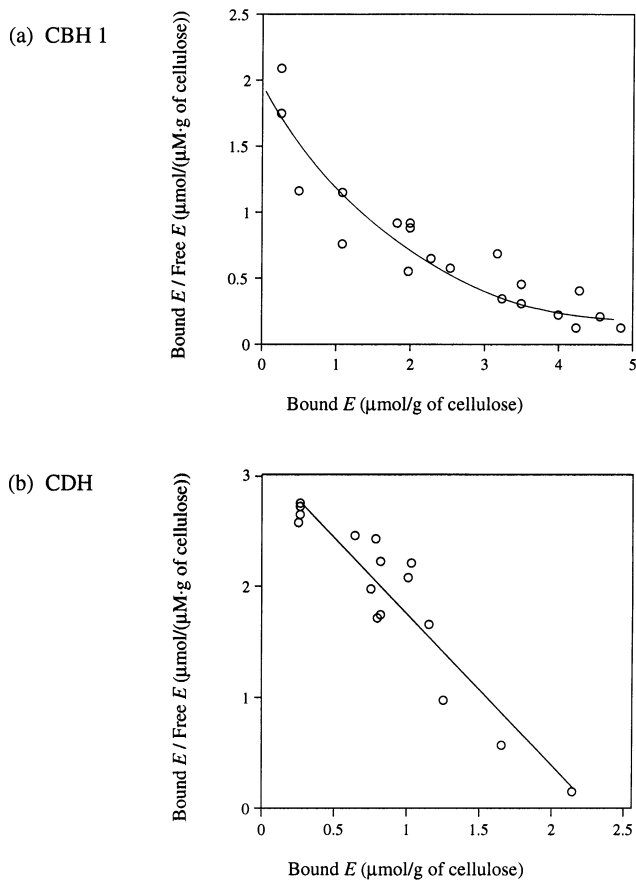**Figure 1   Cellulose-binding isotherm of CBH 1 (○) and CDH (■)**

(a) CBH 1

(b) CDH

Figure 2   Scatchard plots of cellulose-binding isotherm for CBH 1 (a) and CDH (b)

(a)

(b)

(c)

Figure 3   Influence of NaCl (a), $(NH_4)_2SO_4$ (b) and ethylene glycol (c) on the binding of CBH 1 (○) and CDH (□) to cellulose

Figure 4   Kinetics of the binding of CDH

binding was obtained within 10 min (Figure 4). The cellulose binding of CDH was slightly inhibited in the presence of reduced cellodextrins (Figure 5). CDH did not bind to starch, chitin, xylan, mannan or wood powder. Furthermore deglycosylated and non-deglycosylated CDH were bound to cellulose to the same extent and displayed the same activity as that determined by the standard assay [26,50].

Peptides of CDH that had been obtained by proteolytic digestion of CDH and exposed to cellulose were separated by means of SDS/PAGE. The lengths of the peptides ranged from approx. 100 amino acid residues up to the intact enzyme, i.e. 750 residues (the sizes were estimated by comparison with markers of known molecular masses). The Coomassie-stained bands were cut out, subjected to trypsin digestion and sequenced. Two sequences were obtained from a cellulose-binding peptide 350 residues long: VFR (corresponding to residues 118–120) and TGGTYVAPWATS (corresponding to residues 254–265).

DISCUSSION

The binding of CDH to cellulose is clearly different from that of CBH 1 (Figures 1 and 2). First, the capacity of the cellulose to bind CDH is lower than for CBH 1, as expressed both in amount and in mass of bound protein. Secondly, the binding of CDH is stronger, as reflected in a lower K_d. The K_d of CDH is in the same range as that of the CBD of *C. fimi* endoglucanase A [54], which, to our knowledge, has the strongest affinity for cellulose reported for a single domain. Therefore it is highly unlikely that the
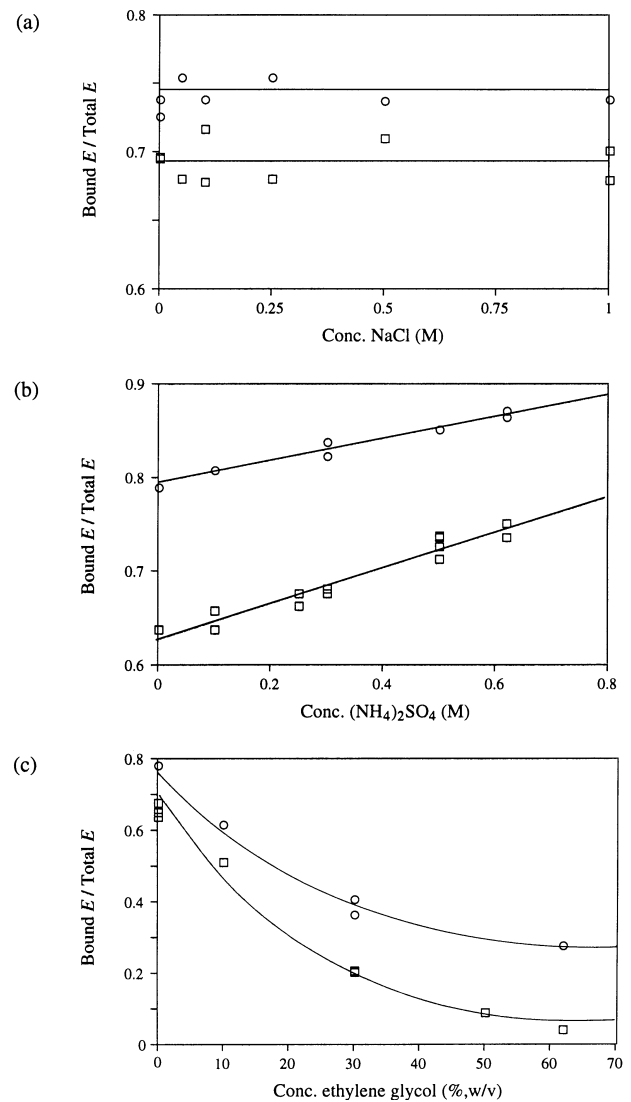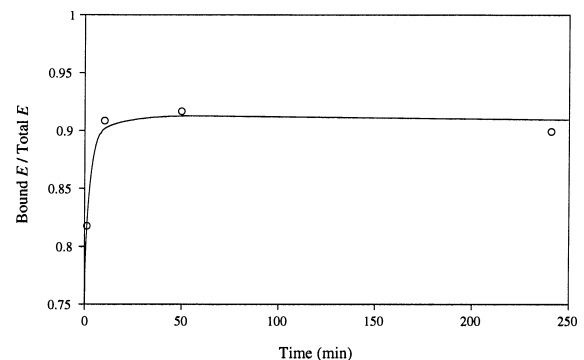
cellulose-binding function in CDH is dependent on a fungal cellulase CBD of type 2. Furthermore no such CBD is present in the amino acid sequence of CDH [29,30].
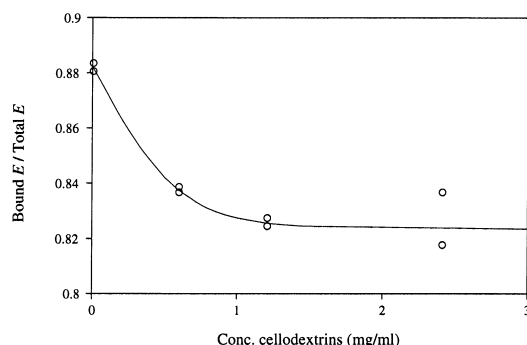
Figure 5   Influence of reduced cellodextrins on the binding of CDH to cellulose

The fit to the one-binding-site model in the classical Scatchard analysis is much better for CDH than for CBH 1. This is surprising because it is expected that the binding sites overlap on cellulose [24]. One explanation might be that CDH binds more sparsely to cellulose than CBH 1, resulting in less overlap. The cellulose microfibril has a heterogeneous surface and it might be that CDH and CBH 1 bind to different sides of the microfibril, or that they bind to regions of cellulose with different degrees of crystallinity. The fact that CBH 1 contains two domains, both of which bind to cellulose, might account for the poor fit to the one-class binding-site model.

The observation that the binding of both proteins is decreased by ethylene glycol and increased by $(NH_4)_2SO_4$ suggests that hydrophobic interaction/charge transfer is involved in the binding of both enzymes. Purely electrostatic forces probably play a minor role because the binding was not affected by sodium chloride (Figure 3). In this context it should be noted that aromatic residues are conserved in both the bacterial type 1 [18] and the fungal type 2 CBDs [16], and that earlier studies have shown that NaCl does not significantly influence the binding of the *T. reesei* cellulases to cellulose [55]. The activity of CBH 1 decreased by approx. 15 % after preincubation in 60 % ethylene glycol. The possibility cannot be excluded that the decrease in binding is at least partly due to denaturation of the protein. In CDH the cellulose binding was partly inhibited by reduced cellodextrins (Figure 5). Although the inhibition is relatively weak, this suggests that the cellulose-binding site of this enzyme is able to recognize a single cellulose chain. In line with this, Samejima and Eriksson have suggested that CDH binds mainly to amorphous cellulose.

Because CDH did not bind to any of the other polysaccharides tested, the enzyme seems to have a distinct and specific affinity for cellulose rather than a general non-specific affinity for carbohydrates. Furthermore the binding phenomenon is independent of the glycosylation state of the protein because removal of surface-attached carbohydrates did not affect the binding properties of CDH. This supports the hypothesis that, in nature, the activity of CDH is limited to exposed regions of naked cellulose [41,56], and makes it less likely that CDH is cell wall-bound *in vivo*.

We have previously suggested that the binding of CDH to cellulose has an immobilizing function [26]. The combination of low K_d values and relatively low capacity supports this hypothesis. Absorption to the cellulose surface is likely to be a prerequisite for two of the suggested functions for CDH: the chemical reduction of aromatic cation radicals produced by lignin and manganese peroxidases to prevent polymerization by

radicals [37], and the generation of hydroxyl radicals, which can attack cellulose, by a Fenton reaction [39,41]. To exert these functions CDH would be expected to act preferentially on naked cellulose regions. For the latter function this would be required in order to enable direct cellulose degradation and to expand the naked regions continuously by the elimination of lignin and hemicellulose.

It has been shown that the cellulose-binding function of CDH is contained within the FAD domain (residues 216–754) [27] and that it does not coincide with the active site [26]. The first sequence (starting at residue 118) obtained from the 350-residue cellulose-binding peptide is located in the haem domain [29,30], whereas the second (starting at residue 256) is located close to the beginning of the FAD-binding domain. This suggests that the cellulose-binding site of CDH is located between residues 118 and 470; however, because the haem domain lacks specific affinity for cellulose and because the FAD domain starts at residue 216, the cellulose binding is likely to be located within residues 216–470. By comparing the amino acid sequence of CDH with those of other proteins, a relationship between the FAD domain of CDH and the glucose–methanol–choline (GMC) oxidoreductase family [57] of flavoproteins has been established [29]. The three-dimensional structures for two members of this family have been determined, namely that of glucose oxidase (GOX) from *Aspergillus niger* [58] and that of cholesterol oxidase (COX) from *Brevibacterium sterolicum* [59,60]. In contrast with CDH, the GMC oxidoreductases do not require haem for their function and consequently do not possess a haem-binding domain.

The structures of GOX and COX consist of two domains: a highly conserved FAD-binding domain and a substrate-binding domain with little or no sequence conservation within the family. The FAD-binding domain includes the canonical $\beta\alpha\beta$ mononucleotide-binding motif [61,62], which corresponds to residues 216–250 in the CDH sequence [29]. From the alignment of the available amino acid sequences of GMC oxidoreductases [57] with that of CDH [29], it was evident that regions that constitute the FAD-binding domain are highly conserved and that they are interspersed with regions that belong to the less well-conserved substrate-binding domain. Superposition of the structures of GOX and COX (Protein Data Bank [63] accession codes 1GAL and 3COX) showed that the FAD-binding domain is also conserved structurally, whereas the substrate-binding domains, although topologically similar [58,60], are more diverse.

Mapping of the amino acid sequence of CDH to the aligned structures of GOX and COX suggests that the residues in CDH that are likely to form the conserved FAD-binding domain are: 216–250, 300–341, 399–484, 502–515, 613–620 and 686–754 (Figure 6). Because of the high conservation of the sequence and structure of the FAD-binding domain, segments that map to this domain are not expected to account for the cellulose-binding properties of CDH. Therefore, on the assumption that the cellulose-binding function of CDH is contained somewhere between residues 216 and 470 in the sequence, only two segments remain that might be involved in cellulose binding: 251–299 (segments LOOP and LID in Figure 6) and 342–398 (segment SUB in Figure 6).

Of the two regions, region 251–299 is particularly interesting. From the alignment in Figure 6 we find that this segment is not conserved throughout the family, and that it seems to be a natural hotspot for deletions and insertions. In GOX and COX, the corresponding regions form a long loop (residues 56–76 in COX; 47–95 in GOX) that protrudes somewhat from the surface of the molecule and then folds back to form a lid (residues 77–97 in COX; 96–109 in GOX) over the entrances to the active site
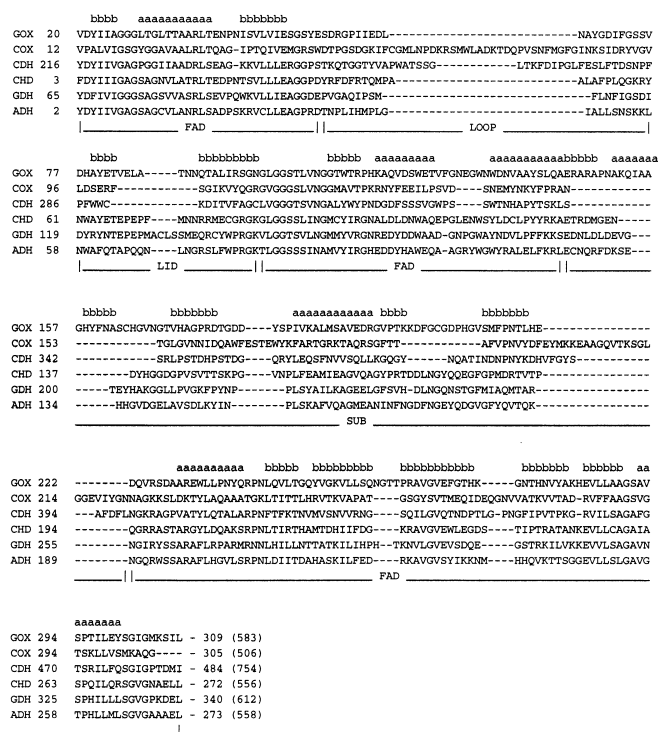
```
          bbbb  aaaaaaaaaaa   bbbbbbbb
GOX  20 VDYIIAGGGLTGLTTAARLTENPNISVLVIESGSYESDRGPIIEDL---------------------------NAYGDIFGSSV
COX  12 VPALVIGSGYGGAVAALRLTQAG-IPTQIVEMGRSWDTPGSDGKIFCGMLNPDKRSMWLADKTDQPVSNFMGFGINKSIDRYVGV
CDH 216 YDYIIVGAGPGGIIAADRLSEAG-KKVLLLERGGPSTKQTGGTYVAPWATSSG------------LTKFDIPGLFESLFTDSNPF
CHD   3 FDYIIIGAGSAGNVLATRLTEDPNTSVLLLEAGGPDYRFDFRTQMPA--------------------------ALAFFLQGKRY
GDH  65 YDFIVIGGGSAGSVVASRLSEVPQWKVLLIEAGGDEPVGAQIPSM-------------------------------FLNFIGGSI
ADH   2 YDYIIVGAGSAGCVLANRLSADPSKRVCLLEAGPRDTNPLIHMPLG----------------------------IALLSNSKKL
        |_____  FAD  _____||_____   LOOP   _____|
```

```
          bbbb    bbbbbbbbbb    bbbbb aaaaaaaaa      aaaaaaaaaabbbbb aaaaaaa
GOX  77 DHAYETVELA-----TNNQTALIRSGNGLGGSTLVNGGTWTRPHKAQVDSWETVFGNEGWNWDNVAAYSLQAERARAPNAKQIAA
COX  96 LDSERF-----------SGIKVYQGRGVGGGSLVNGGMAVTPKRNYFEEILPSVD----SNEMYNKYFPRAN-----------
CDH 286 PFWWC-------------KDITVFAGCLVGGGTSVNGALYWYPNDGDFSSSVGWPS----SWTNHAPYTSKLS-----------
CHD  61 NWAYETEPEPF----MNNRRMECGRGKGLGGSSLINGMCYIRGNALDLDNWAQEPGLENWSYLDCLPYYRKAETRDMGEN-----
GDH 119 DYRYNTEPEPMACLSSMEQRCYWPRGKVLGGSTSVLNGMMYVRGNREDYDDWAAD-GNPGWAYNDVLPFFKKSEDNLDLDEVG---
ADH  58 NWAFQTAPQQN----LNGRSLFWPRGKTLGGSSSINAMVYIRGHEDDYHAWEQA-AGRYWGWYRALELFKRLECNQRFDKSE---
        |_____  LID  _____||_____  FAD  _____||_____|
```

```
          bbbbb    bbbbbbb     aaaaaaaaaaaa bbbb    bbbbbbb
GOX 157 GHYFNASCHGVNGTVHAGPRDTGDD----YSPIVKALMSAVEDRGVPTKKDFDGCGDPHGVSMFPNTLHE---------------
COX 153 ------------TGLGVNNIDQAWFESTEWYKFARTGRKTAQRSGFTT------------AFVPNVVDFEYMKKEAAGQVTKSGL
CDH 342 ------------SRLPSTDHPSTDG----QRYLEQSFNVVSQLLKGQGY------NQATINDNPNYKDHVFGYS-----------
CHD 137 ---------DYHGGDGPVSVTTSKPG----VNPLFEAMIEAGVQAGYPRTDDLNGYQQEGFGPMDRTVTP---------------
GDH 200 ----TEYHAKGGLLPVGKFPYNP-------PLSYAILKAGEELGFSVH-DLNGQNSTGFMIAQMTAR-----------------
ADH 134 ------HHGVDGELAVSDLKYIN--------PLSKAFVQAGMEANINFNGDFNGEYQDGVGFYQVTQK----------------
        _____  SUB  _____
```

```
              aaaaaaaaaa    bbbbb bbbbbbbbbb    bbbbbbbbbbb      bbbbbbb bbbbbb aa
GOX 222 --------DQVRSDAAREWLLPNYQRPNLQVLTGQYVGKVLLSQNGTTPRAVGVEFGTHK-----GNTHNVYAKHEVLLAAGSAV
COX 214 GGEVIYGNNAGKKSLDKTYLAQAAATGKLTITTLHRVTKVAPAT----GSGYSVTMEQIDEQGNVVATKVVTAD-RVFFAAGSVG
CDH 394 ---AFDFLNGKRAGPVATYLQTALARPNFTFKTNVMSNVVRNG-----SQILGVQTNDPTLG-PNGFIYVTPKG-RVILSAGAFG
CHD 194 --------QGRRASTARGYLDQAKSRPNLTIRTHAMTDHIIFDG----KRAVGVEWLEGDS----TIPTRATANKEVLLCAGAIA
GDH 255 --------NGIRYSSARAFLRPARMRNNLHILLNTTATKILIHPH--TKNVLGVEVSDQE-----GSTRKILVKKEVVLSAGAVN
ADH 189 --------NGQRWSSARAFLHGVLSRPNLDIITDAHASKILFED----RKAVGVSYIKKNM----HHQVKTTSGGEVLLSLGAVG
        _____||_____  FAD  _____
```

```
          aaaaaaa
GOX 294 SPTILEYSGIGMKSIL - 309 (583)
COX 294 TSKLLVSMKAQG---- - 305 (506)
CDH 470 TSRILFQSGIGPTDMI - 484 (754)
CHD 263 SPQILQRSGVGNAELL - 272 (547)
GDH 325 SPHILLLSGVGPKDEL - 340 (612)
ADH 258 TPHLLMLSGVGAAAEL - 273 (558)
        _____|
```

Figure 6   Alignment of a subset of sequences for members of the GMC oxidoreductase family

The sequences shown are: GOX (*Aspergillus niger*), COX (*Brevibacterium sterolicum*), CDH (*Phanerochaete chrysosporium*), CHD (choline dehydrogenase; *Escherichia coli*), GDH (glucose dehydrogenase; *Drosophila melanogaster*) and ADH (alcohol dehydrogenase; *Pseudomonas oleovorans*). Only the first 269 residues (216–484) of the 539 residues belonging to the FAD domain (216–754) in CDH have been aligned. An initial structural alignment was made by superimposing the Cα carbon atoms of GOX and COX (Protein Data Bank entries 1GAL [58] and 3COX [59]) with the program O [65], after which the other sequences were added manually to the alignment. Elements of secondary structure, α helix (aaa) and β strand (bbb), in the known three-dimensional structures of GOX and COX were assigned by using the YASSPA command in O [66]. Sequence regions that belong to the FAD-binding domain and the substrate-binding domain are denoted FAD and SUB respectively. The first FAD segment contains the canonical ADP-binding βαβ motif [61,62]. The segment containing the loop-and-lid structure in GOX and COX is referred to as LOOP and LID. Gaps in the sequences of GOX and COX correspond to deletions and insertions, typically in loops and reverse turns. The total length of the sequences is given at the end in parentheses. Accession codes for the sequences are: GOX, Swiss-Prot P13006; CDH, GenBank X88897; CHD, Swiss-Prot P17444; GDH, Swiss-Prot P18173; ADH, Swiss-Prot Q00593; COX, Protein Data Bank 3COX.

and the FAD-binding cavity [58,60]. In CDH these two regions together span 45 residues, which would be sufficient to form an additional domain or an (at least partly) independently folded structural module. Nine aromatic residues, of which three are tryptophan residues, are present in this region and might, by analogy with the CBD of *C. fimi* endoglucanase A, be involved in cellulose binding [18]. Furthermore there are two cysteine residues at the end of the segment (Cys-290 and Cys-299) that have been predicted to form a disulphide bridge [29]. There are no disulphide bridges in this region in either GOX or COX, which suggests that this segment does indeed have a completely different structure in CDH and that the disulphide bond might serve to stabilize that structure. The second segment in the sequence, residues 342–398, is part of the substrate-binding domain (segment SUB in Figure 6). Although the sequence identity between CDH and the other enzymes is low in this region, there are weak similarities in parts defined as α helix or

β sheet, which decreases the probability of finding a unique sequence in this region that might account for cellulose binding.

Because the segment 251–299 displays a natural variability in sequence and length, and because it might be predicted to reside in close proximity to both the active site and the FAD-binding cavity [58–60], this region is the most likely candidate for the role of CBD. CDH is a relatively large protein with overall dimensions of 180 Å × 50 Å × 40 Å (18 nm × 5 nm × 4 nm), as estimated from low-angle X-ray scattering [64]. Assuming that CDH actively takes part in cellulose and/or lignin degradation, the enzyme needs to bind to the surface in a way that facilitates the uptake of substrate into the active site without hampering the catalytic activity. Thus, with no structural data available for CDH and only limited similarity with other enzymes, our best guess is that the cellulose-binding functionality of CDH is due to the peptide 251–299 and, if this is true, that this segment is structurally unique compared with known bacterial and fungal CBDs. Whether this sequence might fold into a distinct domain is uncertain. However, because some small peptides (of approx. 100 residues) were bound to the cellulose, we cannot rule out the possibility of a CBD-like structure in CDH. Because both termini are part of the FAD-binding domain, a putative CBD has to be located internally in the sequence.

REFERENCES

1   Ferreira, L., Wood, T., Williamsson, G., Faulds, C., Hazlewood, G., Black, G. and Gilbert, H. (1993) Biochem. J. **294**, 349–355
2   Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, Jr., R. C. and Warren, R. A. J. (1991) Microbiol. Rev. **55**, 303–315
3   Ståhlberg, J., Johansson, G. and Pettersson, G. (1991) Bio/Technology **9**, 286–290
4   Gilkes, N. R., Warren, R. A. J., Miller, Jr., R. C. and Kilburn, D. G. (1988) J. Biol. Chem. **263**, 10401–10407
5   Lamed, R. and Bayer, E. A. (1988) Adv. Appl. Microbiol. **33**, 1–46
6   Béguin, P. and Aubert, J.-P. (1994) FEMS Microbiol. Rev. **13**, 25–58
7   Lymar, E. S., Li, B. and Renganathan, V. (1995) Appl. Env. Microbiol. **61**, 2976–2980
8   Libby, C. B., Cornett, C. A. G., Reilly, P. J. and Ford, C. (1994) Prot. Eng. **7**, 1109–1114
9   Fuji, T. and Miyashita, K. (1993) J. Gen. Microbiol. **139**, 677–686
10 Wilson, C. A. and Wood, T. M. (1992) Appl. Microbiol. Biotechnol. **37**, 125–129
11 Wilson, C. A. and Wood, T. M. (1992) Enzyme Microbiol. Technol. **14**, 258–264
12 Ali, B. R. S., Zhou, L., Graves, F. M., Freedman, R. B., Black, G. W., Gilbert, H. J. and Hazlewood, G. P. (1995) FEMS Microbiol. Lett. **125**, 15–22
13 Poole, D. M., Morag, E., Lamed, R., Bayer, E. A., Hazlewood, G. P. and Gilbert, H. J. (1992) FEMS Microbiol. Lett. **99**, 181–186
14 Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G. and Miller, Jr., R. C. (1992) Mol. Microbiol. **6**, 1243–1252
15 Kraulis, P. J., Clore, G. M., Nilges, M. J., Jones, T. A., Pettersson, G., Knowles, J. and Gronenborn, A. M. (1989) Biochemistry **28**, 7241–7257
16 Linder, M., Mattinen, M.-L., Kontteli, M., Lindeberg, G., Ståhlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G. and Annila, A. (1995) Prot. Sci. **4**, 1056–1064
17 Din, N., Gilkes, N. R., Tekant, B., Miller, Jr., R. C., Warren, A. J. and Kilburn, D. G. (1991) Bio/Technology **9**, 1096–1099
18 Poole, D. M., Hazlewood, G. P., Huskisson, N. S., Virden, R. and Gilbert, H. J. (1993) FEMS Microbiol. Lett. **106**, 77–83
19 Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R. and Pettersson, L. G. (1986) FEBS Lett. **204**, 223–227
20 Johansson, G., Ståhlberg, J., Lindeberg, G., Engström, Å. and Pettersson, G. (1989) FEBS Lett. **243**, 389–393
21 Srisodsuk, M. (1994) Ph.D. thesis, VTT, Espoo, Finland
22 Nidetzky, B., Steiner, W. and Claeyssens, M. (1994) Biochem. J. **303**, 817–823
23 Tomme, P., van Tilbeurgh, H., Pettersson, G., van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeyssens, M. (1988) Eur. J. Biochem. **170**, 575–582

24 Sild, V., Ståhlberg, J., Pettersson, G. and Johansson, G. (1996) FEBS Lett. **378**, 51–56

25 Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T. and Jones, A. (1994) Science **265**, 524–528

26 Henriksson, G., Pettersson, G., Johansson, G., Ruiz, A. and Uzcategui, E. (1991) Eur. J. Biochem. **196**, 101–106

27 Renganathan, V., Usha, S. N. and Lindenburg, F. (1990) Appl. Microbiol. Biotechnol. **32**, 609–613

28 Perera, I. K., Uzcategui, E., Håkansson, P., Brinkmalm, G., Pettersson, G., Johansson, G. and Sundquist, B. U. R. (1990) Rapid Commun. Mass Spectrom. **4**, 285–289

29 Raices, M., Paifer, E., Cremata, J., Montesino, R., Ståhlberg, J., Divne, C., Szabó, I. J., Henriksson, G., Johansson, G. and Pettersson, G. (1995) FEBS Lett. **369**, 233–238

30 Li, B., Nagalla, S. R. and Renganathan, V. (1996) Appl. Environ. Microbiol. **62**, 1329–1335

31 Ander, P. and Eriksson, K.-E. (1977) Physiol. Plant. **41**, 239–248

32 Ander, P. (1994) FEMS Microb. Rev. **13**, 297–312

33 Eriksson, K.-E., Habu, N. and Samejima, M. (1993) Enzyme Microb. Technol. **15**, 1002–1008

34 Morpeth, F. F. (1985) Biochem. J. **228**, 557–564

35 Ayers, A. R., Ayers, S. B. and Eriksson, K.-E. (1978) Eur. J. Biochem. **90**, 171–181

36 Bao, W., Usha, S. N. and Renganathan, V. (1992) Arch. Biochem. Biophys. **300**, 705–713

37 Ander, P., Mishra, C., Farrell, R. and Eriksson, K.-E. (1990) J. Biotechnol. **13**, 189–198

38 Bao, W. and Renganathan, V. (1991) FEBS Lett. **279**, 30–32

39 Kremer, S. and Wood, P. (1992) Eur. J. Biochem. **205**, 133–138

40 Bao, W. and Renganathan, V. (1992) FEBS Lett. **302**, 77–80

41 Henriksson, G., Ander, P., Pettersson, B. and Pettersson, G. (1995) Appl. Microbiol. Biotechnol. **42**, 790–796

42 Roy, B. P., Paice, M. G., Archibald, F. S., Misra, S. K. and Misiak, L. (1994) J. Biol. Chem. **269**, 19745–19750

43 Dekker, R. H. F. (1988) Methods Enzymol. **160**, 454–463

44 Henriksson, G. (1995) Ph.D. thesis, Uppsala University, Almqvist and Wiksell International, Stockholm, Sweden

45 Bhikhabhai, R., Johansson, G. and Pettersson, G. (1984) J. Appl. Biochem. **6**, 336–345

46 Sakamoto, R., Arai, M. and Murao, S. (1989) Agric. Biol. Chem. **53** (5), 1407–1409

47 Herstin, S. (1963) in Methods in Carbohydrate Chemistry, vol. 3, (Whistler, R. L., ed.), pp. 4–9, Academic Press, NY

48 Gilkes, N. R., Jervis, E., Henrissat, B., Tekant, B., Miller, Jr., R. C., Warren, R. A. J. and Kilburn, D. G. (1992) J. Biol. Chem. **267**, 6743–6749

49 Hofsten, B., van Kley, XX and Eaker, D. (1965) Biochim. Biophys. Acta **110**, 585–598

50 Westermark, U. and Eriksson, K.-E. (1975) Acta Chem. Scand. B Org. Chem. Biochem. **29**, 419–424

51 Maizel, J. V. (1969) in Fundamental Techniques in Virology, (Habel, K. and Salzman, N. P., eds.), pp. 334–362, Academic Press, NY

52 Hewick, R. M., Hunkapiller, M. W., Hood, L. E. and Dreyer, W. J. (1981) J. Biol. Chem. **256**, 7990–7997

53 Deshpande, M., Eriksson, K.-E. and Pettersson, G. (1984) Anal. Biochem. **138**, 481-487

54 Gilkes, N. R., Kilburn, D. G., Miller, Jr., R. C., Warren, R. A. J., Sugiyama, J., Chanzy, H. and Henrissat, B. (1993) Int. J. Biol. Macromol. **15**, 347–351

55 Kyriacou, A., Neufeld, R. J. and MacKenzie, C. R. (1988) Enzyme Microb. Technol. **10**, 675–681

56 Henriksson, G., Johansson, G. and Pettersson, G. (1993) Biochim. Biophys. Acta **1144**, 184–190

57 Cavener, D. R. (1992) J. Mol. Biol. **223**, 811–814

58 Hecht, H. J., Kalisz, H. M., Hendle, J., Schmid, R. D. and Schomburg, D. (1993) J. Mol. Biol. **229**, 153–172

59 Vrielink, A., Lloyd, L. F. and Blow, D. M. (1991) J. Mol. Biol. **219**, 533–554

60 Li, J., Vrielink, A., Brick, P. and Blow, D. M. (1993) Biochemistry **32**, 11507–11515

61 Wierenga, R. K., Drenth, J. and Schulz, G. E. (1983) J. Mol. Biol. **167**, 725–739

62 Wierenga, R. K., Terpstra, P. and Hol, W. G. J. (1986) J. Mol. Biol. **187**, 101–107

63 Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, Jr., M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. **112**, 535–542

64 Lehner, D., Zipper, P., Henriksson, G. and Pettersson, G. (1996) Biochim. Biophys. Acta **1293**, 161–169

65 Jones, T. A., Zou, J.-Y., Cowan, S. W. and Kjeldgaard, M. (1991) Acta Crystallogr. **A47**, 110–119

66 Kleywegt, G. and Jones, T. A. (1996) Methods Enzymol., in the press