JOURNAL OF BACTERIOLOGY, Feb. 1992, p. 1403–1409
0021-9193/92/041403-07$02.00/0

Vol. 174, No. 4

Characterization and Comparison of *Clostridium cellulovorans* Endoglucanases-Xylanases EngB and EngD Hyperexpressed in *Escherichia coli*

FRANCES C.-F. FOONG AND ROY H. DOI*

Department of Biochemistry and Biophysics, University of California, Davis, California 95616

Received 4 October 1991/Accepted 13 December 1991

By the use of a T7 expression system, endoglucanases-xylanases EngB and EngD from *Clostridium cellulovorans* were hyperexpressed and purified from *Escherichia coli*. The two enzymes demonstrated both endoglucanase and xylanase activities. The substrate specificities of both endoglucanases were similar except that EngD had four-times-greater *p*-nitrophenyl β-1,4-cellobiosidase activity. The two proteins were very homologous (80%) up to the Pro-Thr-Thr region which divided the protein into -NH_2- and -COOH- terminals. The -COOH- region of EngB has high homology to the endoglucanases and a xylanase from *Clostridium thermocellum* and to an endoglucanase from *Clostridium cellulolyticum* and did not show strong binding to cellulose (Avicel). However, the -COOH- region of EngD, which had homology to the cellulose-binding domains of *Cellulomonas fimi* exo- and endoglucanases and to *Pseudomonas fluorescens* endoglucanase, demonstrated binding ability to cellulose even when the domain was fused to the N-terminal domain of EngB. By probing the Avicel-purified cellulase complex (F8) with anti-EngB and anti-EngD antibodies, both EngB and EngD were shown to be present on the cellulase complex of *C. cellulovorans*. Many proteins homologous to EngB and EngD were also present on the complex.

Cellulose degradation is an important part of the biological recycling of carbon. Since the initial investigation of the mechanism of cellulose degradation more than 30 years ago, much progress has been achieved, especially with aerobic fungal cellulases. The degradative process can be different for different species, especially between anaerobic and aerobic organisms. With cellulolytic organisms such as *Trichoderma reesei*, *Neocallismastix frontalis*, *Pseudomonas fluorescens*, *Cellulomonas fimi*, *Fibrobacter succinogenes*, and *Ruminococcus flavefaciens*, it is not certain whether the enzymes essential for the degradation of crystalline cellulose function individually or whether they are present on the cell or substrate surface as a complex (1, 15, 20). There is evidence that in *Clostridium thermocellum* as well as in *Clostridium* strain C7, the enzymes are present in a complex, termed the "cellulosome" (3, 18, 19, 22, 23). There is evidence that the enzymes present in *Clostridium cellulovorans* are also present in a complex. It has been shown that the native cellulase complex of *C. cellulovorans* is a multiprotein complex of large molecular mass. The enzymic proteins are probably organized on a nonenzymic scaffolding protein which has been cloned and sequenced (25, 27, 29). The aim of our investigation was to find out whether our cloned enzymes were part of the cellulase complex.

Many cellulolytic enzymes have been found to contain a catalytic domain and a binding domain (5, 8–10, 16, 31, 34). It was found that the specificity of *C. cellulovorans* endoglucanases EngB and EngD was conferred by the NH_2-terminal domain and that the COOH-terminal domains of both genes had homology to enzymes from other species of cellulolytic bacteria (6, 13). Thus, it would also be interesting to compare our cloned endoglucanases with each other as well as with those of other species in terms of their characteristics and cellulose affinities. Cellulose-binding domains (CBD) have not been described in any cellulolytic *Clostridium* species except in one recent report (16).

In order to study the mechanism of crystalline cellulose degradation, we have used *C. cellulovorans*, an anaerobic, mesophilic bacterium (30). We have cloned four endoglucanase genes (*engA*, *engB*, *engC* [26], and *engD* [14]) and a cellulose-binding-protein gene, *cbpA* (27). Since the individual proteins in the cellulase complex are very difficult to separate, our strategy was to clone the genes and hyperproduce and purify each protein separately so that we could study and compare them. We hope eventually to reconstitute an active complex from the subunits and elucidate their functional domains.

We describe in this paper the hyperproduction and purification of *C. cellulovorans* endoglucanases EngB and EngD with the vector pET T7-phage RNA polymerase system and the comparison of the purified enzymes in terms of their substrate specificities and biochemical characteristics. We also studied their affinities to Avicel and analyzed their putative binding domains. The presence of these proteins in Avicel affinity-purified cellulase fraction F8 (25), which contains the multiprotein cellulase complex from the culture growth medium, was also investigated.

MATERIALS AND METHODS

Plasmids and strains. Endoglucanase gene *engD* plasmid pET-D and the signal peptide deletion mutant *engB* plasmid pET-B2 were made by using the appropriate primers, with the *Nco*I restriction site and *Bam*HI site synthesized for the 5′ and 3′ termini, respectively, by the polymerase chain reaction. Signal peptide mutants had to be constructed for *engB* because of problems in the growth of the cells and hyperproduction of the protein. The template used for the insert of pET-D was pEQ52V (14). The template for pET-B2 was pC2-engB (6). The templates for binding mutants pET-D1 and pET-D2, which encoded the binding mutant

* Corresponding author.

proteins Eng-D1 and Eng-D2, were plasmids pEC-9 and pEC-17, respectively (13). Binding mutants pET-D1 and pET-D2 were also made by synthesizing the appropriate primers with the *Nco*I restriction site for the 5′ termini and the *Bam*HI restriction site for the 3′ termini and cloning them into pET-3d, previously known as pET-8C (32, 33). Primer oligonucleotides were synthesized by the Gene Assembler II (Pharmacia). Temperature was cycled 30 times at 94°C for 30 s, 52°C for 30 s, and 72°C for 30 s with a COY temperature cycler. Plasmid construction was always checked by restricting the fragment with the cloning site enzymes and also by restricting one other enzyme site internal to the fragment. For hyperproduction of the proteins, pET-3d was used as the cloning vector and *Escherichia coli* BL21 DE3 was used as the recipient strain. All controls used were the cloning vector itself without any insert transformed into the recipient strain.

Enzyme assays. Enzyme assays for the clones were performed with 0.1% (wt/vol) substrate in 1 ml of enzyme solution and PC buffer (50 mM potassium phosphate, 12 mM sodium citrate [pH 6]). The samples were incubated for 12 h at 37°C, and the reducing sugars formed were measured by the Nelson-Somogyi reducing sugar test (35).

Hyperproduction and purification of endoglucanases. Clones used to hyperproduce EngD and EngB contained the plasmids pET-D and pET-B2, respectively. The method of protein purification was modified from that used to purify overproduced *Bacillus subtilis* major sigma factor (4). Cells were grown to 150 Klett units in 2YT medium (twice the amount of nutrients as Luria-Bertani) with 100 μg of ampicillin per ml and 0.4% glucose. They were then induced with 1 mM IPTG (isopropyl-β-D-thiogalactoside), and after 30 min, 200 μg of rifampin per ml was added and the cells were cultured a further 2 h. The cells were then harvested by centrifugation, and the cell pellet was suspended in disruption buffer (0.05 M Tris HCl [pH 8.0], 2 mM EDTA, 0.1 mM dithiothreitol, 1 mM 2-mercaptoethanol, 0.2 M NaCl, 5% glycerol, 130 μg of lysozyme per ml) and incubated on ice for 30 min. Sodium deoxycholate was added to a final concentration of 0.05%, and the cells were lysed by sonic oscillation. The lysate was diluted with 35 ml of TEDG buffer (10 mM Tris HCl [pH 7.0], 0.1 mM EDTA, 0.1 mM dithiothreitol, 5% glycerol). The cell debris was spun down with a Sorvall GSA rotor at 4°C at 8,000 rpm. The cell debris was suspended in 20 ml of TEDG buffer containing 6 M guanidine hydrochloride and spun at 15,000 rpm in a Sorvall type SS-34 rotor for 30 min at 4°C, and the supernatant was diluted gradually to 1,200 ml with TEDG buffer. The protein remaining in the supernatant after dilution was precipitated with 80% ammonium sulfate, centrifuged at 10,000 rpm in a Sorvall GSA rotor for 30 min, and then resuspended in 5 ml of TEDG buffer. It was dialyzed in the same buffer overnight at 4°C, and the debris was removed by centrifugation for 20 min at 8,000 rpm in a Sorvall type SS-34 rotor. The supernatant was used for purification of EngD and EngB. The buffer in the supernatant was replaced by TEA buffer (20 mM triethanolamine [pH 7.2], 0 to 350 mM NaCl gradient) by diluting it and then concentrating it with a Centriprep-10 (Amicon). The protein was applied to a fast-protein liquid chromatography (Pharmacia, LKB) MonoQ anion-exchange column, and the active fractions were collected, pooled, and concentrated with a Centricon-10 (Amicon) microconcentrator.

Binding mutant inserts eng-D1 and eng-D2 were also made by the polymerase chain reaction and cloned into pET-3d vector to form the binding mutants pET-D1 and pET-D2, which coded for the binding mutant proteins Eng-D1 and Eng-D2, respectively. The schematic of construction is shown in Fig. 1. The proteins were hyperproduced and partially purified by solubilizing the inclusion bodies with 6 M guanidine HCl. It was renatured by diluting the solution, precipitating the protein with 80% ammonium sulfate, and dialyzing the solution overnight in TEDG buffer.

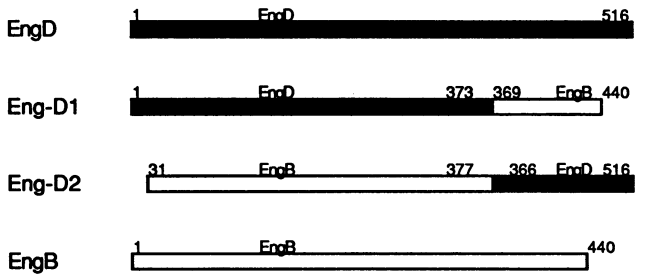

FIG. 1. Schematic of construction of binding mutant proteins Eng-D1 and Eng-D2. The amino acid numbers of each protein are above the lines. Solid black boxes represent EngD portions, and open boxes represent EngB portions. Binding mutants Eng-D1 and Eng-D2 were made by exchanging the -COOH- domains of EngD and EngB. The proteins were hyperproduced by cloning the inserts into pET-3d.

Antibody production. For production of anti-EngD and anti-EngB antibody, the protein from the inclusion bodies (cell debris) was obtained in the same way from *E. coli*, but the supernatant after centrifugation was applied to a 10% sodium dodecyl sulfate (SDS)-polyacrylamide gel, and the appropriate band was visualized by soaking the gel in cold 0.25 M KCl. The band was cut out, and the protein was eluted with a Schleicher & Schuell Elutrap electroseparation system in Tris-glycine buffer (25 mM Tris [pH 8.3], 192 mM glycine). The protein (about 500 μg) was emulsified with incomplete adjuvant and injected subcutaneously into two New Zealand White rabbits at multiple injection sites. They were boosted twice with the same amount of protein over a 6-week period. Blood taken from the rabbits was allowed to clot at 37°C for 1 h. Serum was recovered after overnight incubation at 4°C by centrifugation at 10,000 × *g*, and immunoglobulin G was purified by using DEAE Affi-Gel Blue (Bio-Rad) according to the manufacturer's instructions.

Western blot detection. Western blot (immunoblot) analysis was performed by using a 1:2,000 dilution of primary antibody. Anti-rabbit immunoglobulin G alkaline phosphate-conjugated antibodies were used as the secondary antibodies (Bio-Rad). The Western Detection kit of Bio-Rad was used, and the protocol was followed according to the manufacturer's instructions.

Growth of *C. cellulovorans* and cellulase isolation. *C. cellulovorans* (ATCC 35296) was grown anaerobically as described previously (30). The carbon source of 0.1% Avicel cellulose substrate (PH-101; FMC Corp.) was used. The cellulase complex that binds to Avicel was purified by essentially the same method described before (25). The proteins in the culture growth medium of cellulose-grown *C. cellulovorans* cells after 7 days of incubation at 37°C were precipitated with 80% ammonium sulfate. After dialysis in PC buffer at 4°C, the cellulase solution was incubated with 3% Avicel for 10 min at 37°C. The cellulose was precipitated and washed twice with PC buffer–1 M NaCl and twice with PC buffer–double-distilled water. The cellulase complex was then removed from the Avicel cellulose by washing it with

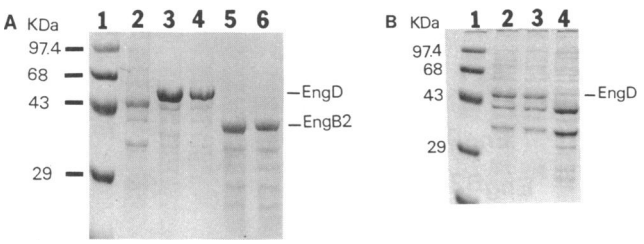

FIG. 2. (A) Coomassie blue staining of SDS-PAGE gel of hyperproducing clones pET-D and pET-B2 (inclusion body fraction). Lanes: 1, molecular weight markers; 2, cell debris (inclusion body) fraction of *E. coli* clone with pET-3d vector; 3 and 4, inclusion body fraction of pET-D clones; 5 and 6, inclusion body fraction of pET-B2 clones. The equivalent of 100 μl of sample from 1 ml of cell was applied. (B) Coomassie blue staining of SDS-PAGE gel of hyperproducing clone pET-D (whole-cell extract). Lanes: 1, molecular weight markers; 2 and 3, whole-cell extract from pET-D clone; 4, whole-cell extract from *E. coli* control with pET-3d vector. The equivalent of 100 μl of sample from 1 ml of cell was applied.

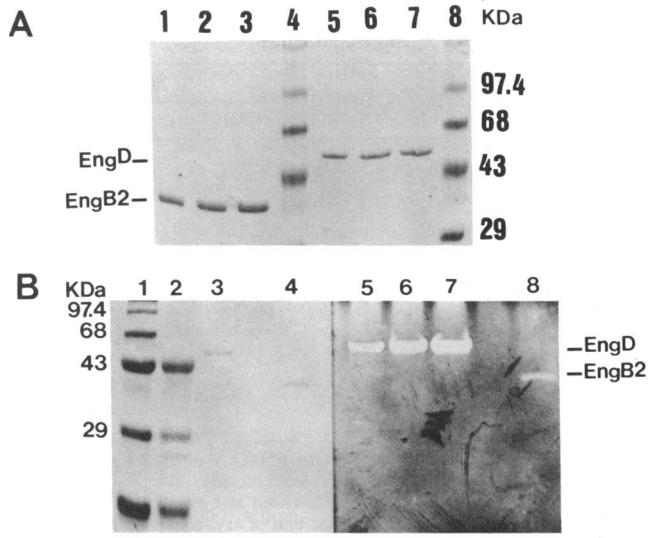

FIG. 3. (A) Coomassie blue staining of SDS-PAGE gel of purified EngB2 and EngD. Lanes: 1, 2, and 3, 1, 2, and 3 μg, respectively, of purified EngB2 from clone pET-B2; 4, molecular weight markers; 5, 6, and 7, 1, 2, and 3 μg, respectively, of purified EngD from clone pET-D; 8, molecular weight markers. (B) Coomassie blue and Congo red staining of CMC–SDS-PAGE gels. Lanes: 1, high-molecular-weight markers; 2, low-molecular-weight markers; 3, purified EngD from clone pET-D; 4, purified EngB2 from clone pET-B2 (Coomassie blue); 5, 6, and 7, 0.5, 1.0, and 1.5 μg, respectively, of purified EngD; 8, 1.5 μg of purified EngB2 (Congo red staining).

400 ml of double-distilled water, precipitated with ammonium sulfate, and dialyzed with PC buffer. This cellulose-affinity-purified cellulase fraction, termed F8, was subjected to SDS-polyacrylamide gel electrophoresis (PAGE), and the proteins were probed with anti-EngB and anti-EngD antibodies.

PAGE and zymogram. SDS-PAGE and native PAGE were performed in 10 and 12.5% polyacrylamide gels, respectively. Both Mini-Protein II (Bio-Rad) and PhastSystem (Pharmacia) were used. Carboxymethyl cellulase (CMCase) zymograms were prepared by using 0.1% carboxy methyl cellulose (CMC) that was copolymerized with polyacrylamide. CMCase activity was visualized with 0.1% Congo red after 6 h of incubation at 37°C in PC buffer (pH 7.0).

Cellulose-binding assays. Different dilutions of purified EngB and EngD as well as of binding mutant proteins Eng-D1 and Eng-D2 were tested for estimating the best concentrations to use for the binding assays. Concentrations of ~5 μg for EngB and EngD and ~10 μg for the binding mutant proteins were finally chosen. The protein solutions were incubated with 200 μl of Avicel solution (PH101; FMC; 5% [wt/vol] suspended in 100 mM Tris HCl buffer, pH 7) as described by Gilbert et al. (7), with mild agitation at 4°C for 1 h. Two methods were used in separating the supernatant from the Avicel after incubation. One method was gravity filtration by glass filter, and the other was centrifugation for 2 min in a microcentrifuge. The bound and unbound proteins were analyzed by Western blot analysis, and reducing sugar assays for enzymic activity were performed with the fraction prior to incubation with Avicel and in the supernatant after incubation with Avicel.

RESULTS

Hyperproduction and purification of endoglucanases. In order to hyperexpress EngB and EngD, genes *engB* and *engD* were amplified by the polymerase chain reaction method and cloned into the T7 RNA polymerase hyperproduction system of the vector pET-3d (33). Since cells containing *engB* grew poorly and no hyperproduction was achieved with EngB, a signal peptide mutant, EngB2, with 22 of its N-terminal amino acids deleted was used instead, and hyperexpression was achieved with this mutant. Hyperexpressed proteins were found mainly in the inclusion body fraction (approximately 90%), and SDS-PAGE analysis showed that the protein was quite free from other contaminating proteins, a fact which facilitated the purification process (Fig. 2A). The whole-cell soluble extract also contained the hyperexpressed protein, and it was estimated that the protein was expressed to about 40% of the total cell protein (Fig. 2B).

Purified EngB had a molecular mass of approximately 40 kDa, and EngD had a molecular mass of approximately 50 kDa (Fig. 3A). The substrate specificities of purified EngD and EngB2 were compared (Table 1). They were very similar except that EngD had about four-times-greater *p*-nitrophenyl β-1,4-cellobiosidase activity, indicating the presence of cellobiosidase activity as originally reported (14). The purified enzymes retained the substrate specificities of the original clones (6, 13, 14). The specific activity of EngD was about 80 times greater than that of the unpurified periplasmic fraction of the *E. coli* clone, and the specific activity of EngB2 purified protein was about 70 times greater than that of the periplasmic fraction of the pC2-engB *E. coli* clone (6). Isoelectric focusing showed that the isoelectric point of EngD was around 8.8 and that of EngB2 was 7.0 (data not shown).

In order to determine the activity bands and molecular weights of the purified endoglucanases-xylanases, the purified enzymes were applied to SDS-PAGE denaturing gels copolymerized with 0.1% CMC. The samples applied to the portion for Coomassie staining were boiled with 2× SAB (9 parts of 0.125 M Tris HCl [pH 6]–4% SDS–20% gycerol–0.002% bromophenol blue with 1 part of mercaptoethanol) for 5 min, whereas the samples applied to the portion for CMCase visualizing were mixed with 2× SAB and heated to 50°C for 3 min. The gel was then divided into two portions. One portion was stained with Coomassie Blue, and the other

TABLE 1. Specific activity and substrate specificities of purified EngB and EngD

Substrate[a]	Sp act[b]	
	EngB	EngD
CMC	82.3	86.4
Cellulose, microgranular	51.4	37.0
Cellulose, fibrous	30.5	12.2
Avicel, microcrystalline	20.4	16.6
Polygalacturonic acid	18.5	11.1
Mannan	46.3	36.7
Laminarin	50.9	27.7
Lichenan	853.9	555.5
Xylan	118.0	77.2
PNPC	7.4	31.5
PNPG	15.6	7.1
PNPX	7.8	2.9

[a] PNPC, *p*-nitrophenyl β-D-cellobioside; PNPG, *p*-nitrophenyl β-D-glucopyranoside; PNPX, p-nitrophenyl β-D-xylopyranoside.

[b] One unit of polymer hydrolysis is expressed as 1 μmol of reducing sugar liberated per min per mg of protein. One unit of aryl glycoside hydrolysis is expressed as 1 μmol of *p*-nitrophenyl released per min per mg of protein.

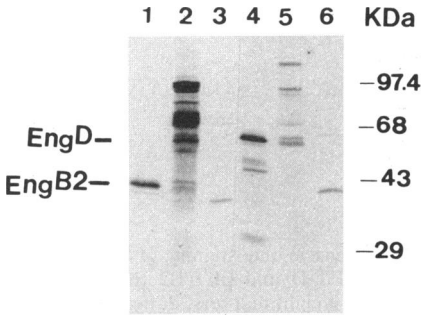

FIG. 5. Western blot analysis of Avicel cellulose-affinity-purified cellulase fraction (F8) from *C. cellulovorans* growth medium (cultured in 0.1% Avicel). Lanes: 1, 2.0 ng of purified EngB2; 2, Avicel cellulose-affinity-purified cellulase (F8) fraction from the growth medium probed with anti-EngB2 antibody; 3, concentrated proteins from *E. coli* control growth medium; 4, 0.2 ng of purified EngD (additional bands are degradation products of EngD); 5, Avicel cellulose-affinity-purified cellulase; 6, concentrated proteins from *E. coli* control growth medium (lanes 4, 5, and 6 were probed with anti-EngD antibody).

portion was washed and incubated in PC buffer (pH 7.0) for 6 h at 37°C and then stained with 0.1% Congo red to visualize the CMCase activities. The Coomassie-stained bands of the purified enzymes showed the same mobilities and molecular weights as the activity bands (Fig. 3B) when stained with Congo red, i.e., ~40 kDa for EngB2 and 50 kDa for EngD.

Homology and cellulose-binding activities of EngB and EngD. EngB and EngD have 80% homology up to the proline-threonine-threonine "hinge" region of EngD (13), and their antibodies cross-reacted. The carboxyl domains of the two proteins, which had no homology, were exchanged with each other to investigate their cellulose-binding activities.

The cellulose-binding activities of EngB and EngD as well as of chimeras of EngB and EngD (binding mutants Eng-D1 and Eng-D2) were tested by mixing the enzymes with Avicel. EngD could bind to Avicel, as determined by the assay conditions used (7). The supernatant from the EngD enzyme-Avicel mixture after 1 h of incubation at 4°C did not contain the protein, but the supernatant of the Avicel mixture contained the protein after being boiled with 2× SAB, as shown by the Western blot analysis (Fig. 4A). The activity remaining in the supernatant after incubation with Avicel was (on average) 2.1 U (1 U is defined as 1 μmol of reducing sugar liberated per min per ml), whereas the activity of the enzyme prior to incubation with Avicel was 13.9 U. Some binding was also detected with EngB2, although Western probing of the binding showed that it was also found in the supernatant (Fig. 4A). The remaining activity in the supernatant was 12.7 U compared with 21.1 U prior to incubation with Avicel. For mutant proteins Eng-D1 and Eng-D2, Western blot analyses and enzymic activity assays showed that Eng-D1, which had the EngB carboxyl terminal, was detected mainly in the supernatant, whereas Eng-D2, which had the EngD carboxyl terminal, was found mostly in the bound fraction (Fig. 4B). The average enzymic activities for Eng-D1 and Eng-D2 were respectively 23.4 and 1.55 U before binding and 23.8 and 0.35 U after binding.

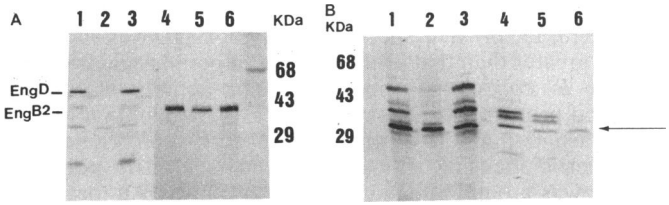

FIG. 4. (A) Western blot analysis of EngD and EngB2 binding to Avicel. Lanes: 1, EngD prior to incubation with Avicel; 2, supernatant of Avicel mixture after incubation with EngD; 3, after incubation with EngD, Avicel was mixed with 2× SAB and boiled for 5 min, and the supernatant was then applied; 4, EngB2 prior to incubation with Avicel; 5, supernatant of Avicel mixture after incubation with EngB2; 6, supernatant of incubated Avicel after being boiled with 2× SAB. (B) Western blot analysis of Avicel binding of binding mutant proteins Eng-D1 and Eng-D2. Lanes: 1, Eng-D2 prior to incubation with Avicel; 2, Eng-D2 after incubation with Avicel; 3, bound fraction from Avicel; 4, Eng-D1 prior to incubation with Avicel; 5, Eng-D1 after incubation with Avicel; 6, fraction of Eng-D1 bound to Avicel. Arrow indicates *E. coli* protein.

Western detection of cellulose-purified fraction. Since considerable amino acid homology existed between EngB and EngD, the cross-reactivities of their antibodies were tested. Antibodies raised against cloned EngB cross-reacted with EngD. Anti-EngB also reacted with other proteins in the *C. cellulovorans* cellulase complex, although it did not cross-react with proteins precipitated and concentrated with ammonium sulfate from the control *E. coli* growth medium. Western blot analysis showed that the F8 (cellulase complex) fraction contained EngB and EngD. Other proteins of higher molecular masses were also detected by anti-EngB and anti-EngD antibodies. EngB as well as proteins in the 50- to 100-kDa range were detected by anti-EngB antibodies. Proteins of around 50 kDa (presumably EngD) and 60, 100, and 150 kDa were also detected by EngD antibodies. This implied that there are high-molecular-mass proteins homologous to EngB and EngD within the cellulase complex (Fig. 5). When a native PAGE of the F8 fraction was probed, the band detected had a much higher molecular weight than EngB and EngD, indicating that the enzymes were present in a complex. EngB was also detected in the total protein from the growth medium of cellulose-grown cells, indicating that it could also be found as an independent enzyme (Fig. 6).

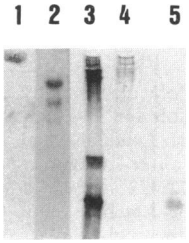

FIG. 6. Western blot analysis of native PAGE. Lanes: 1, Avicel cellulose-affinity-purified cellulase complex (F8) from *C. cellulovorans* growth medium; 2, purified EngD; 3, total protein from *C. cellulovorans* growth medium (0.1% cellulose carbon source); 4, Avicel cellulose-affinity-purified cellulase (F8); 5, purified EngB2. Lanes 1 and 2 were probed with anti-EngD antibody, and lanes 3, 4, and 5 were probed with anti-EngB2 antibody.

DISCUSSION

The hyperproduction and purification of endoglucanases-xylanases EngB and EngD were carried out by cloning the genes into the T7 phage hyperproduction system (vector pET-3d) by using restriction sites at the *Nco*I (5′ terminus) and *Bam*HI (3′ terminus) sites of the vector that were created by polymerase chain reaction. The hyperexpression of EngD but not of EngB was achieved easily. Also, attempts to purify EngB from 10 liters of cell culture by using the original clone, pC2 (26), which had a 3.2-kb fragment containing the gene in vector pUC19, were unsuccessful. It was also noticed that clones of pC2 and those with the entire *engB* gene grew very slowly (6). On closer inspection of the signal peptides between *engB* and *engD*, it was noticed that their signal peptides differed from the homologous portions of the genes. Signal deletion mutants were then constructed for *engB*, and these mutant clones were then able to grow normally. Hyperexpression and inclusion body production could then be achieved with this deletion mutant. The putative signal peptide of EngB was found to cause inner membrane lysis of the host cell (unpublished results).

EngD protein could easily be detected by Western blot of the cytoplasmic and periplasmic fractions (unpublished results) of pUC19 vector clones. We therefore attempted to detect the binding mutant proteins Eng-D1 and Eng-D2 in the crude cell extract as well as in the periplasmic fractions of the clones, but the proteins produced by these mutants either were produced in very small quantities or were easily degraded by *E. coli* proteases so that they could not be detected. The proteins were then hyperproduced by pET vectors, solubilized from the inclusion body, and renatured for binding and enzymic assays. The binding mutant proteins were then detected easily.

Both the nucleotide and amino acid sequences of EngB and EngD are 80% homologous up to the Pro-Thr-Thr linker of EngD. The homology is 75% for the entire gene sequence (13). EngD has the Pro-Thr-Thr linker region seen previously in *C. fimi* genes (24) which is absent from *engB*. This type of linker has never been seen in *Clostridium* enzymes. There is no substantial homology between *engB* and *engD* after this linker region. EngB had homology to *C. thermocellum* endoglucanase genes *celA*, *celB*, and *celD* and xylanase gene *xynZ* as well as to *Clostridium cellulolyticum* endoglucanase at the carboxyl-terminal region. The role of this conserved region is still unknown, but the region is not needed for catalytic activity. It has been suggested that this structure may interact with a scaffolding or organizational protein of the cellulase complex (6, 11, 12). The EngD carboxyl-terminal region had homology to *C. fimi* exoglucanase and endoglucanase as well as to *P. fluorescens* endoglucanase and xylanase (13). These regions were shown to be CBD (7–9), and since *engD* has some homology to these regions, it was suggestive that EngD had affinity to cellulose as well. Binding assays demonstrated that EngD could bind to Avicel, as shown by Western blot analysis and enzyme assays before and after incubation with cellulose. EngB bound to Avicel considerably less efficiently but was able to bind to Avicel under the conditions used since the enzyme was detected in the bound fraction as well as in the supernatant after incubation.

The mutant protein Eng-D1, which has an EngD N terminus and an EngB C terminus, could not bind to Avicel. Mutant protein Eng-D2, which has an EngB N terminus and an EngD C terminus, was able to bind to cellulose with the characteristics of the native EngD protein. This indicated that the carboxyl region of 138 amino acids of EngD, which has homology with the CBD of endo- and exoglucanases from other bacterial species, was the binding domain of EngD. CBD have been proven to exist in a few bacterial species and in eukaryotic enzymes from *T. reesei* but have not been demonstrated frequently in *Clostridium* species. Recently, there was a report of a *Clostridium stercorarium* endoglucanase having a CBD (16). Our results with EngD confirm that a *Clostridium* enzyme can indeed possess a CBD.

The binding ability of EngB was not clear, since it bound weakly to Avicel and the binding mutant with an EngB C terminus was unable to bind to Avicel. Therefore, it may have a very weak binding domain. No binding activity has been reported with other endoglucanases which have homology to EngB at the C terminus.

Proteins found on the cellulase complex ideally should have the same regulational control so that the proper ratios of proteins are made, but there exists the possibility that some genes that code for proteins on the cellulase complex may not be under the same regulational influences or on the same operon, so ratios of these proteins may not be compatible at certain times. Some enzymes may also be expressed constitutively in order to maintain a basal level of enzymes, since it is energy wasting to express the whole cellulase complex at all times. Enzymes may also have started out as individual proteins and gradually evolved to organization as a complex in anaerobes. This may increase the enzyme's degradative efficiency and specific activity in an environment which has no oxygen to aid the degradative process. The scaffolding protein has no enzymic activity but very high affinity to cellulose (25). Therefore, enzymes found on the complex probably undergo more mutations on their residual binding domains, and this may have resulted in enzymes with various affinities to their substrate.

C. cellulovorans secretes a protein which has no enzymic activity but is essential for the degradation of crystalline cellulose (25). The gene coding for this protein, *cbpA*, has been cloned and sequenced in our laboratory. CbpA is probably a scaffolding or organizational protein in *C. cellulovorans* that organizes and concentrates the enzymes so that crystalline cellulose can be degraded efficiently (29). The cellulase complex containing the organizational protein has much higher binding affinity to cellulose than the individual enzymes do. EngB and EngD can be detected in the F8 cellulose-affinity-purified cellulase that binds to Avicel (Fig. 5). This fraction has been shown to contain a cellulase complex of high molecular mass (25). The cellulase multi-

protein complex also contains endoglucanases of higher molecular masses which are homologous to EngB and EngD. This is not surprising, since the cellulase complex of *Clostridium* strain C7 consists of at least 15 proteins (2), and the cellulosome of *C. thermocellum* may contain even more.

Both EngB and EngD have endoglucanase and endoxylanase activities. EngD also shows significant cellobiosidase activity (Table 1). This multiple-activity spectrum may be important for the degradation of a substrate as heterogeneous as plant cellulosic material. The xylanase activity may be important for breaking down the hemicellulose and lignin matrix surrounding the cellulose microfibrils of plant cell wall. Since *C. cellulovorans* can also utilize xylan as a growth substrate (30), the initial breakdown of xylan can also provide the cells with a carbon and energy source. Polypeptides of *C. thermocellum*, *Clostridium* strain C7 (3), and *C. cellulovorans* found on the multiprotein complex also have xylanase activities whose functions may enhance cellulose hydrolysis in nature by breaking down the hemicellulose matrix. The cellulosomal subunits of *C. thermocellum* possess considerable xylanase activity even though *C. thermocellum* cannot utilize xylan. Zymogram analysis demonstrated that the xylanase activity bands corresponded to CMCase activity bands for some of these subunits in *C. thermocellum* (22).

EngB has very high homology to EngD as well as some homology to endoglucanases from other species. Endoglucanases secreted by *C. cellulovorans* are homologous to each other, since antibodies against EngB cross-reacted with EngD as well as with other proteins in the molecular mass range of 50 to 100 kDa, which coincided with the zymogram of CMCase bands of the cellulase complex (25). Endoglucanases which do not have homology to EngB and EngD probably also exist in *C. cellulovorans*, since an endoglucanase gene, *engC*, has also been sequenced, and it has no homology to either *engB* or *engD* (28). The multiplicity of endoglucanases has been the subject of investigation, as it is not known whether this variety is due to expression of different genes or to posttranslational modification of the same gene product (20–22). Gene *engB* has high homology to *engD* and perhaps to other endoglucanases in *C. cellulovorans*, as evidenced by their cross-reactivities. Gene *engD* has homology to *C. fimi* and *P. fluorescens* genes. Two genes from *P. fluorescens* contain identical CBD (17). Homologous recombination and domain shuffling within their own species as well as interspecific exchange of genetic material must have occurred between cellulolytic organisms to give rise to these patterns of homology (8).

It is hoped that the availability of large amounts of purified EngB and EngD and of the scaffolding protein CbpA will enable us to do in vitro reconstitution experiments. Although the complete native cellulase complex may never be achieved in vitro, a partial reconstitution to obtain "true" cellulase activity is very plausible.

The report that the multiprotein cellulase complex of *Clostridium* strain C7 may be assembled entirely extracellularly (2) gives us an optimistic outlook on achieving reconstitution and true cellulase activity. It is also possible to investigate the hypothesis of whether the reiterated conserved carboxyl region found in EngB and many *Clostridium* enzymes is the domain which interacts with the scaffolding protein.

ACKNOWLEDGMENTS

The work was supported in part by grant DE-FG03-87ER13705 from the U.S. Department of Energy to R.H.D.

We thank Tetsuo Hamamoto for his expert advice.

REFERENCES

1. **Beguin, P.** 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. **44:**219–248.
2. **Cavedon, K., S. B. Leschine, and E. Canale-Parola.** 1990. Cellulase system of a free-living, mesophilic *Clostridium* (strain C7). J. Bacteriol. **172:**4222–4230.
3. **Cavedon, K., S. B. Leschine, and E. Canale-Parola.** 1990. Characterization of the extracellular cellulase from a mesophilic *Clostridium* (strain C7). J. Bacteriol. **172:**4231–4237.
4. **Chang, B.-Y., and R. H. Doi.** 1990. Overproduction, purification, and characterization of *Bacillus subtilus* RNA polymerase δ^A factor. J. Bacteriol. **172:**3257–3263.
5. **Ferreira, L. M. A., A. J. Durrant, J. Hall, G. P. Hazlewood, and H. J. Gilbert.** 1990. Spatial separation of protein domains is not necessary for catalytic activity or substrate binding in xylanases. Biochem. J. **269:**261–264.
6. **Foong, F., T. Hamamoto, O. Shoseyov, and R. H. Doi.** 1991. Nucleotide sequence and characteristics of endoglucanase gene *engB* from *Clostridium cellulovorans*. J. Gen. Microbiol. **137:** 1729–1736.
7. **Gilbert, H. J., J. Hall, G. P. Hazlewood, and L. M. A. Ferreira.** 1990. The N-terminal region of an endoglucanase from *Pseudomonas fluorescens subspecies cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. Mol. Microbiol. **4:**759–767.
8. **Gilkes, N. R., B. Henrissat, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Domains in microbial β-1,4-glycanases: sequence conservation function and enzyme families. Microbiol. Rev. **55:**303–315.
9. **Gilkes, N. R., D. G. Warren, R. C. Miller, Jr., and R. A. J. Warren.** 1989. Structural and functional analysis of a bacterial cellulase by proteolysis. J. Biol. Chem. **264:**17802–17808.
10. **Gilkes, N. R., R. A. J. Warren, R. C. Miller, Jr., and D. G. Kilburn.** 1988. Precise excision of the *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. J. Biol. Chem. **263:**10401–10407.
11. **Grepinet, O., M. Chebrou, and P. Beguin.** 1988. Nucleotide sequence and deletion analysis of the xylanase gene (*xynZ*) of *Clostridium thermocellum*. J. Bacteriol. **170:**4582–4588.
12. **Hall, J., G. P. Hazlewood, P. J. Barker, and H. J. Gilbert.** 1988. Conserved and reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. Gene **69:**29–38.
13. **Hamamoto, T., F. Foong, O. Shoseyov, and R. H. Doi.** 1991. Analysis of functional domains of endoglucanases from *Clostridium cellulovorans* by gene cloning, nucleotide sequencing and chimeric protein construction. Mol. Gen. Genet., in press.
14. **Hamamoto, T., O. Shoseyov, F. Foong, and R. H. Doi.** 1990. A *Clostridium cellulovorans* gene, *engD*, codes for both endo-beta-1,4-glucanase and cellobiosidase activities. FEMS Microbiol. Lett. **72:**285–288.
15. **Hazlewood, G. P., K. Davidson, J. H. Clarke, A. J. Durrant, J. Hall, and H. J. Gilbert.** 1990. Endoglucanase E, produced at high level in *Escherichia coli* as a *lac* Z fusion protein, is part of the *Clostridium thermocellum* cellulosome. Enzyme Microb. Technol. **12:**656–662.
16. **Jauris, S., K. P. Rucknagel, W. H. Schwarz, P. Kratzsch, K. Bronnenmeir, and W. L. Staudenbauer.** 1990. Sequence analysis of the *Clostridium stercorarium celZ* gene encoding a thermoactive cellulase (Avicelase I). Identification of catalytic and cellulose binding domains. Mol. Gen. Genet. **233:**258–267.
17. **Kellett, L. E., D. M. Poole, L. M. A. Ferreira, A. J. Durrant, G. P. Hazlewood, and H. J. Gilbert.** 1990. Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens subsp. cellulosa* contain identical cellulose domains and are encoded by adjacent genes. Biochem. J. **272:**369–376.
18. **Lamed, R., and E. A. Bayer.** 1988. The cellulosome concept: exocellular and extracellular enzyme reactor centers for efficient binding and cellulolysis, p. 101–116. *In* J.-P. Aubert, P. Beguin, and J. Millet (ed.), Biochemistry and genetics of cellulose degradation. Academic Press, Inc., San Diego, Calif.

19. **Mayer, F., M. P. Coughlan, Y. Mori, and L. G. Ljungdahl.** 1987. Macromolecular organization of the cellulolytic enzyme complex of *Clostridium thermocellum* as revealed by electron microscopy. Appl. Environ. Microbiol. **53**:2785–2792.
20. **McGavin, M., and C. W. Forsberg.** 1988. Isolation and characterization of endoglucanases 1 and 2 from *Bacteroides succinogenes* S85. J. Bacteriol. **170**:2914–2922.
21. **Mischak, H., F. Hofer, R. Messner, E. Weissinger, and M. Hayn.** 1989. Monoclonal antibodies against different domains of cellobiohydrolase I and II from *Trichoderma reesei*. Biochim. Biophys. Acta **990**:1–7.
22. **Morag, E., E. A. Bayer, and R. Lamed.** 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose-degrading enzymes. J. Bacteriol. **172**:6098–6105.
23. **Morag, E., I. Halevy, E. A. Bayer, and R. Lamed.** 1991. Isolation and properties of major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. J. Bacteriol. **173**: 4155–4162.
24. **O'Neill, G., S. H. Goh, R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1986. Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*. Gene **44**:325–330.
25. **Shoseyov, O., and R. H. Doi.** 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87**:2192–2195.
26. **Shoseyov, O., T. Hamamoto, F. Foong, and R. H. Doi.** 1990. Cloning of *Clostridium cellulovorans* endo-1,4-beta-glucanase genes. Biochem. Biophys. Res. Commun. **169**:667–672.
27. **Shoseyov, O., T. Hamamoto, F. Foong, and R. H. Doi.** 1990. Cloning of 170 kDa *Clostridium cellulovorans* cellulase subunit: an essential protein for the degradation of crystalline cellulose, abstr. O-24, p. 267. Abstr. 90th Annu. Meet. Am. Soc. Microbiol. 1990.
28. **Shoseyov, O., L. Shoseyov, M. Dekel-Reichenbach, M. Goldstein, F. Foong, and R. H. Doi.** 1991. Unpublished data.
29. **Shoseyov, O., M. Takagi, M. Goldstein, and R. H. Doi.** Primary sequence of *Clostridium cellulovorans* cellulose binding protein A (CbpA). Proc. Natl. Acad. Sci. USA, in press.
30. **Sleat, R., R. A. Mah, and R. Robinson.** 1984. Isolation and characterization of an anaerobic, cellulolytic bacterium, *Clostridium cellulovorans* sp. nov. Appl. Environ. Microbiol. **48**:88–93.
31. **Stahlberg, J., G. Johansson, and G. Petersson.** 1988. A binding-site-deficient, catalytically active, core protein of endoglucanase III from the culture filtrate of *Trichoderma reesei*. Eur. J. Biochem. **173**:179–183.
32. **Studier, F. W., and B. A. Moffat.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
33. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
34. **Tomme, P., H. Van Tilbeurgh, G. Petersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens.** 1988. Studies of the cellulolytic system of *Trichoderma reesei* QM9414. Eur. J. Biochem. **170**:575–581.
35. **Wood, W. A., and K. M. Bhat.** 1988. Methods for measuring cellulase activities. Methods Enzymol. **160**:87–112.