JOURNAL OF BACTERIOLOGY, Apr. 1993, p. 1910–1918
0021-9193/93/071910-09$02.00/0

Vol. 175, No. 7

Cellulose-Binding Polypeptides from *Cellulomonas fimi*: Endoglucanase D (CenD), a Family A β-1,4-Glucanase

A. MEINKE,[1,2] N. R. GILKES,[1] D. G. KILBURN,[1] R. C. MILLER, JR.,[1] AND R. A. J. WARREN[1]*

Department of Microbiology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3,[1] and Institut für Biologie III der Universität Freiburg, Schänzlestrasse 1, D-7800 Freiburg im Breisgau, Germany[2]

Received 25 September 1992/Accepted 19 January 1993

Five cellulose-binding polypeptides were detected in *Cellulomonas fimi* culture supernatants. Two of them are CenA and CenB, endo-β-1,4-glucanases which have been characterized previously; the other three were previously uncharacterized polypeptides with apparent molecular masses of 120, 95, and 75 kDa. The 75-kDa cellulose-binding protein was designated endoglucanase D (CenD). The *cenD* gene was cloned and sequenced. It encodes a polypeptide of 747 amino acids. Mature CenD is 708 amino acids long and has a predicted molecular mass of 74,982 Da. Analysis of the predicted amino acid sequence of CenD shows that the enzyme comprises four domains which are separated by short linker polypeptides: an N-terminal catalytic domain of 405 amino acids, two repeated sequences of 95 amino acids each, and a C-terminal domain of 105 amino acids which is >50% identical to the sequences of cellulose-binding domains in Cex, CenA, and CenB from *C. fimi*. Amino acid sequence comparison placed the catalytic domain of CenD in family A, subtype 1, of β-1,4-glycanases. The repeated sequences are more than 40% identical to the sequences of three repeats in CenB and are related to the repeats of fibronectin type III. CenD hydrolyzed the β-1,4-glucosidic bond with retention of anomeric configuration. The activities of CenD towards various cellulosic substrates were quite different from those of CenA and CenB.

When grown in the presence of cellulose, *Cellulomonas fimi* produces a variety of polypeptides with carboxymethylcellulase activity, some of which may arise by deglycosylation or proteolysis of native enzymes (25). This complicates the purification of the native polypeptides. Gene cloning simplifies the isolation and characterization of the native enzymes because the cloned genes can be expressed in an organism, such as *Escherichia coli*, which is devoid of other β-1,4-glycanases (5). This approach has led to the characterization of four β-1,4-glycanases from *C. fimi*: endoglucanases A, B, and C (CenA, CenB, and CenC) and an exoglycanase (Cex) (14, 32–34, 41, 42). Although it hydrolyzes carboxymethyl cellulose (CM-cellulose), Cex is properly classified as a xylanase (11). All of these enzymes contain discrete cellulose-binding domains (CBDs) which can function independently of the catalytic domains (7, 8, 15, 24, 28). While the CBDs of CenA, CenB, and Cex are closely related, the two N-terminal CBDs of CenC are only distantly related (7). CenA and Cex (44 and 47 kDa, respectively) both comprise a catalytic domain joined to a CBD by a short linker polypeptide; CenB and CenC (106 and 113 kDa, respectively) contain additional domains which contribute to their larger size (7, 8, 15, 24, 28).

The *cenA*, *cenB*, and *cex* genes were first isolated as *E. coli* clones expressing polypeptides which reacted with an antiserum to supernatant proteins from a *C. fimi* culture grown in the presence of cellulose (14, 41). The *cenC* gene was isolated by taking advantage of the capacity of CenC to bind to Sephadex (32). This article describes the cloning and sequencing of the *cenD* gene of *C. fimi* by a different approach. *cenD* encodes endoglucanase D (CenD), which was initially identified as a cellulose-binding polypeptide (Cbp) from *C. fimi* distinct from Cex, CenA, CenB, and CenC.

* Corresponding author.

MATERIALS AND METHODS

Materials. *Torula* yeast extract was from Natural Focus Foods Ltd., Vancouver, Canada. Avicel PH101 (a microcrystalline cellulose preparation derived from wood) was from FMC International, Little Island, Ireland. Bacterial microcrystalline cellulose (BMCC) was prepared from *Acetobacter xylinum* (ATCC 23769) as described previously (13). Regenerated cellulose was prepared from Avicel PH101 as described previously (7). CM-cellulose (sodium salt; low-viscosity grade [15]) and cellulose azure (type I) were from Sigma Chemical Co., St. Louis, Mo. Azurine-cross-linked xylan was from MegaZyme Ltd., North Rocks, Australia. Cellopentaose (<99% pure) was from Seikagaku America Inc., Rockville, Md.

Bacterial strains and plasmids. The *C. fimi* strain used was ATCC 484. *E. coli* DH5α and DH5α F′ were purchased from Bethesda Research Laboratories, Burlington, Canada. *E. coli* RZ1032 was used as the host for plasmids to produce single-stranded DNA for site-directed mutagenesis. Bacteriophage M13KO7 was used for the production of single-stranded DNA for sequencing and mutagenesis. Plasmids pTZ18R and pTZ19R have been described previously (29).

Growth conditions. *E. coli* strains were grown at 30 or 37°C in Luria broth or tryptone-yeast extract-phosphate medium (35), supplemented with 50 to 100 μg of ampicillin per ml. Solid medium contained 1.5% agar. *C. fimi* was grown at 30°C in basal salts medium (41) containing 1% CM-cellulose and 1% *Torula* yeast extract.

Identification of Cbps. The cells from an 8- to 10-day-old *C. fimi* culture were removed by centrifugation. A 5-ml aliquot of the supernatant was adsorbed with 10 mg of Avicel, 1 mg

of BMCC, or 10 mg of Sephadex G-50 by incubation for 30 min on ice with occasional shaking. The adsorbent was recovered by centrifugation and then washed once with 1 ml of 0.5 M NaCl and twice with 1 ml of 10 mM Tris-HCl (pH 8.0). A sample of the washed adsorbent was boiled for 2 min with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer, and the extracted polypeptides were analyzed by SDS-PAGE. For N-terminal amino acid sequencing, polypeptides were resolved by SDS-PAGE, transferred electrophoretically to a polyvinylidene difluoride membrane, and detected by staining with Coomassie blue. Appropriate bands were excised from the membrane, and the polypeptides were sequenced directly with an Applied Biosystems 470A gas phase sequenator. Western immunoblotting was performed as described previously (15).

PCR. Genomic *C. fimi* DNA was isolated as described previously (32). Oligodeoxynucleotide primers were synthesized with an Applied Biosystems 380A automated DNA synthesizer. Each oligodeoxynucleotide had an 8-nucleotide 5′ extension coding for an *Eco*RI site and two extra nucleotides at the 5′ terminus. Mixtures for polymerase chain reaction (PCR) contained 50 mM Tris-HCl (pH 8.0), 0.05% Tween 20, 0.05% Nonidet P-40, 2 mM $MgCl_2$, 10% dimethyl sulfoxide, 0.2 mM each of the four 2′-deoxynucleotide 5′-triphosphates, 25 pM each of the primers, 10 to 100 ng of *C. fimi* DNA, and 1 U of *Taq* DNA polymerase in a total volume of 50 μl. One cycle of 1 min at 94°C, 2 min at 45°C, and 3 min at 72°C was followed by 34 cycles of 30 s at 94°C, 45 s at 50°C, and 2.5 min at 72°C. The final cycle was like the previous 34 cycles except that the incubation time at 72°C was extended to 7 min. The PCR product was digested with *Eco*RI and ligated into pTZ18R which had been digested with *Eco*RI.

Cloning of *cenD*. The PCR product was labeled with [α-^{32}P]dATP by the random hexamer primer method (9). *C. fimi* DNA was digested with appropriate restriction endonucleases and then fractionated on a 10 to 40% sucrose gradient (35). The fractions were concentrated by ethanol precipitation and analyzed by Southern hybridization (35, 39) at 65°C in 6× SSC (90 mM sodium citrate, 900 mM sodium chloride)–5× Denhardt's solution–0.5% SDS–10 μg of salmon sperm DNA per ml with the ^{32}P-labeled PCR fragment as a probe. Filters were washed for 15 min in 2× SSC buffer at 55°C, then for 30 min in 2× SSC buffer–0.1% SDS at 55°C, and finally for 10 to 20 min in 0.1× SSC buffer at 55°C. The DNA from positive fractions was ligated into pTZ18R which had been digested with the corresponding restriction endonuclease(s). *E. coli* DH5α F′ was transformed with the ligation mixture. Ampicillin-resistant clones were screened for the presence of the *cenD* gene by colony hybridization with the ^{32}P-labeled PCR fragment as a probe.

DNA sequencing and sequence analysis. The nucleotide sequences of both strands were determined by the dideoxy method (36), with either double-stranded or single-stranded DNA, T7 DNA polymerase, and 7-deaza-dGTP instead of dGTP in the nucleotide mixes. Nucleotide and amino acid sequences were analyzed with the PCGENE program from Intelligenetics.

Expression of *cenD* in *E. coli* DH5α. Site-directed mutagenesis was performed as described previously (29). An oligodeoxynucleotide designated cenDMut1 (5′-GCGCGA TGCGGAGG**CCATGG**GTTCTCCTCG), complementary to the coding strand, was designed to change the GTG start codon to an ATG (shown in boldface) and at the same time to introduce an *Nco*I site (underlined) incorporating the new ATG start codon. pDAM1-1 (see Fig. 2B) was digested with *Apa*I, and the ends were filled in with the Klenow fragment of DNA polymerase I. The linearized plasmid was digested with *Pst*I; the 500-bp fragment containing the first 400 bp of *cenD* and 100 bp of the 5′ noncoding sequence was recovered and ligated into pTZ18R which had been digested with *Sma*I and *Pst*I to give pDAM1-2. Mutation of pDAM1-2 with cenDMut1 produced pDAM1-3. The recognition site for *Apa*LI, encoding the first two amino acids of CenD, was changed through the mutagenesis. The mutation was confirmed by sequencing. Sequencing also revealed the insertion of two short stretches of polylinker DNA of unknown origin, which resulted in the addition of sites for *Bam*HI, *Xba*I, *Sal*I, and *Pst*I on the 5′ side of the insert and sites for *Bam*HI, *Xba*I, and *Sal*I on the 3′ side of the insert. This also explained the reversed orientation of the insert. The extra sites did not influence the subsequent manipulations except that the extra *Pst*I site was used for one of the next cloning steps. The 2.8-kbp *Pst*I-*Sst*I fragment from pDAM1, containing 1.9 kbp of *cenD* at the 3′ end and 900 bp of noncoding sequence, was subcloned into pTZ19R which had been digested with *Pst*I and *Sst*I to give pDAM1-4. Finally, the 500-bp *Pst*I fragment from pDAM1-3 was ligated into the *Pst*I site of pDAM1-4 to produce pDAM2. Since the *lacZ* start codon was in frame with the *cenD* start codon in pDAM2, expression of *cenD* in this construct produced two proteins corresponding to processed CenD and the unprocessed LacZ-CenD fusion (data not shown). Therefore, pDAM2 was digested with *Hin*dIII, the site was filled in with Klenow enzyme, and the plasmid was religated to yield pDAM2-1. Expression of *cenD* from this construct produced only one polypeptide which bound to cellulose, confirming that the fusion of *lacZ* to *cenD* was no longer in frame. N-terminal amino acid sequencing showed that processing occurred at the same site in *C. fimi* and *E. coli* (data not shown).

Purification of CenD. CenD was purified from a 20-liter *E. coli* DH5α(pDAM2-1) culture grown at 30°C; IPTG (isopropylthiogalactopyranoside) was added to a final concentration of 0.1 mM when the culture reached an A_{600} of 1.8; growth was then continued for 16 h. Purification of CenD from crude *E. coli* cell extract was performed as previously described for CenB (31). Protein concentrations were estimated by the far-UV method of Scopes (37).

Enzyme assays. The hydrolysis of Avicel PH101, BMCC, and regenerated cellulose by CenA, CenB, and CenD was assayed by measurement of reducing-sugar groups or total sugar released into solution at 37°C. Reducing-sugar assays were performed in 50 mM sodium citrate, pH 6.8, containing 0.02% NaN_3 and 0.02% bovine serum albumin (citrate buffer-BSA). Soluble reducing sugars were quantitated by assay of the reaction supernatant with hydroxybenzoic acid hydrazide reagent (26). Assays for the release of total sugar were performed in 50 mM potassium phosphate, pH 7.0, containing 0.02% NaN_3 (phosphate buffer). Total soluble sugars were quantitated in the reaction supernatant by the phenol-sulfuric acid assay (6). Avicel (15 mg) was incubated with 0.25 nmol of enzyme in 1.5 ml of buffer for 18 h. BMCC (1.5 mg) was incubated with 1.0 nmol of enzyme in 1.5 ml of buffer for up to 18 h. Regenerated cellulose (10 mg) was incubated with 0.02 nmol of enzyme in 1.45 ml of buffer for 2 h.

The hydrolysis of CM-cellulose was determined from the production of new reducing-sugar groups in the reaction mixture, measured with hydroxybenzoic acid hydrazide reagent. CM-cellulose (2 mg) was incubated with 1.0 pmol of

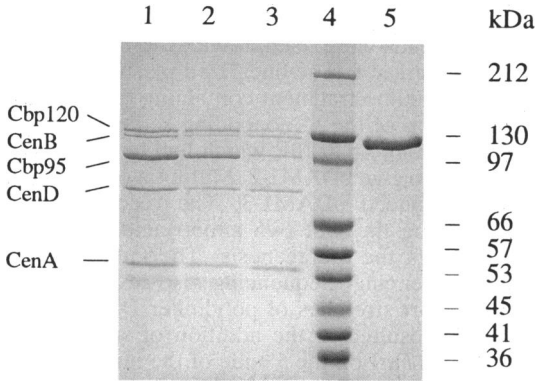

FIG. 1. Cbps in *C. fimi* culture supernatants. Polypeptides from 5-ml aliquots of culture supernatant were adsorbed with Avicel (lane 1), BMCC (lane 2), or Sephadex (lane 3), resolved by SDS-PAGE (7.5% acrylamide), and stained with Coomassie blue. Lane 4 contains molecular mass standards; sizes are indicated on the far right. Lane 5 contains 1.5 μg of purified CenB.

CenA or CenB or 20 pmol of CenD in 0.5 ml of citrate buffer-BSA for 30 min at 30°C. The hydrolysis of cellulose azure (a Remazol brilliant blue-dyed cellulose preparation) was determined from the release of dyed product into solution, measured by its A_{595}. Cellulose azure (40 mg) was incubated with 0.1 nmol of enzyme in 2.0 ml of citrate buffer-BSA at 37°C. Xylanase activity was determined with insoluble azurine-cross-linked xylan as the substrate. The reaction mixture contained 1 mg of substrate and 0.5 μg of Cex or 50 μg of CenA, CenB, or CenD in 1 ml of phosphate buffer. After incubation at 37°C for up to 6 h, the rate of hydrolysis was determined from the release of dyed product into solution, measured by its A_{595}. One unit of xylanase activity results in a change of 1 absorbance unit in the reaction supernatant per hour.

Determination of stereochemical course of hydrolysis. Preliminary experiments (data not shown) established the conditions under which cellopentaose was rapidly hydrolyzed by CenD, allowing the anomeric configuration of the products to be determined before and after mutarotation. A reaction mixture containing 50 μl of 2.5 mM cellopentaose and 2 μl of a 6.17-mg/ml CenD solution was analyzed, after incubation at room temperature for 15 s or at 37°C for 1 min and then at 100°C for 2 min, with a Waters high-performance liquid chromatography system (Millipore Corp., Milford, Mass.) fitted with a Waters Dextro-Pak Cartridge and operated at a column flow rate of 1 ml/min. Sugars were detected by refractometry.

Nucleotide sequence accession number. The reported nucleotide sequence has been assigned GenBank accession number L02544.

RESULTS

Cbps from *C. fimi*. Five Cbps were recovered from *C. fimi* culture supernatants by adsorption to cellulose or Sephadex (Fig. 1). Their molecular masses and N-terminal amino acid sequences showed that the 53- and 110-kDa polypeptides corresponded to CenA and CenB, respectively, but the 120-, 95-, and 75-kDa polypeptides did not correspond to any of the known *C. fimi* β-1,4-glycanases (Table 1). The N termini of the 120- and 95-kDa polypeptides did not correspond to any N-terminal or internal sequences in the protein data base; the N terminus of the 75-kDa Cbp was similar to those of β-1,4-glycanases in family A, subtype 1 (4, 21). These data provided evidence for the first identification of a family A enzyme in *C. fimi*. The enzyme was tentatively designated endoglucanase D and its gene *cenD*, because all but one of the family A enzymes are endoglucanases (5, 12, 20).

TABLE 1. N-terminal amino acid sequences of Cbps[a]

Cycle	Cbp				
	Cbp120	CenB[b]	Cbp95	CenD[b]	CenA[b]
1	A	A	A	A	A
2	V	P	P	T	X
3	T	T	V	G	X
4	T/G	Y	H	D	V
5	E	N	V	D	D
6	Y	Y	D	W	Y
7	A	A	N	L	A
8	Q	E	P	X	V
9	X	A	Y	V	T
10	F	L	A	E	N/I
11	L		G	G	Q
12	A		A	N	W
13	Q		V	T	P
14	Y		Q	I	
15	D		Y	V	
16	K		V	D	
17	I		N	S	
18	K		P	T	
19	R			G	
20	P			K	
21	A				
22	N				

[a] Amino acid sequences were determined by Edman degradation, as described in Materials and Methods. Amino acid residues are represented by the single-letter code; residues which could not be determined are indicated by the letter X.

[b] N-terminal sequence has been shown to be the same for the polypeptides produced in *C. fimi* and *E. coli*.

Cloning of *cenD*. Alignment of the amino acid sequences of the enzymes in family A, subtype 1, showed that two stretches of six amino acids each, corresponding to W238 to L243 and P264 to P269 of mature CenD, are highly conserved (see Fig. 4). Degenerate oligodeoxynucleotides (2 and 3; see Fig. 4) were synthesized which corresponded to these conserved amino acid sequences. A third oligodeoxynucleotide (1; see Fig. 4) was synthesized which corresponded to the six N-terminal amino acids of mature CenD, using the codon bias of other sequenced *C. fimi* genes (8, 28, 33, 42). The combination of oligodeoxynucleotides 1 and 3 allowed amplification of an 800-bp fragment of *C. fimi* DNA by PCR.

The amplified fragment was cloned into pTZ18R and then sequenced. The amino acid sequence predicted by the nucleotide sequence confirmed that CenD belongs to family A, subtype 1. The cloned PCR product was labeled with ^{32}P and used as a probe to detect a *Pvu*II fragment of *C. fimi* DNA containing the entire *cenD* gene (Fig. 2A). The fragment was cloned into pTZ18R to give pDAM1.

Sequence of *cenD*. Restriction enzyme analysis showed that *cenD* overlapped a 1.9-kbp *Pvu*II-*Pst*I fragment and a 2.8-kbp *Pst*I fragment from the 5.6-kbp *Pvu*II fragment (Fig. 2B). Sequencing of the entire 2.8-kbp *Pst*I fragment and about half of the 1.9-kbp *Pvu*II-*Pst*I fragment revealed the sequence of the entire *cenD* gene (Fig. 2 and 3). The gene encodes a polypeptide of 747 amino acids. The codon usage of the *cenD* gene is very similar to that of *cex*, *cenA*, *cenB*, and *cenC* (8, 28, 33, 42). Only 35 codons are used in *cenD*.

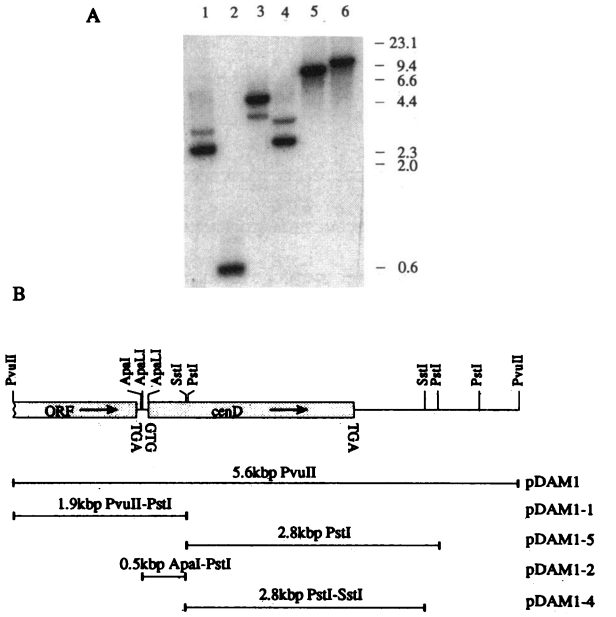

FIG. 2. Southern analysis and schematic presentation of the *C. fimi* genomic DNA fragment containing *cenD*. (A) *C. fimi* DNA was digested for 8 h at 37°C with *Sst*I (lane 1), *Sal*I (lane 2), *Pvu*II (lane 3), *Pst*I (lane 4), *Kpn*I (lane 5), or *Bgl*II (lane 6) and resolved by agarose gel electrophoresis. DNA fragments were screened with a ^{32}P-labeled, cloned PCR product designed to hybridize to *cenD*. Sizes are shown in kilobase pairs. (B) Schematic representation of the 5.6-kbp *Pvu*II fragment containing *cenD*. The restriction sites shown were determined by restriction analysis and DNA sequence analysis. The 1.9-, 2.8-, 0.5-, and 2.8-kbp restriction fragments shown below were subcloned into pTZ18R or pTZ19R for sequencing or expression of *cenD*; the corresponding recombinant plasmids were designated as indicated on the right. ORF, open reading frame.

The putative GTG translational start codon is preceded by a ribosome-binding site. Nine nucleotides after the TGA translational stop codon, there is an inverted repeat of 19 nucleotides which may be a transcriptional stop signal (Fig. 3). Promoters have been characterized for the previously isolated *C. fimi* β-1,4-glycanase genes (16, 17, 32); 50 nucleotides upstream of the postulated GTG translational start codon are two sequences of six nucleotides each which are 17 bp apart and resemble typical *C. fimi* promoter sequences (Fig. 3).

There is an open reading frame 129 nucleotides upstream of the *cenD* gene, and 33 nucleotides after the TGA translational stop codon of this open reading frame, there is an inverted repeat of 15 nucleotides which overlaps the postulated promoter sequence of the *cenD* gene (Fig. 3).

Amino acid sequence of CenD. Comparison of the predicted amino acid sequence of CenD with the primary structures of other previously characterized bacterial β-1,4-glycanases showed that CenD comprises four domains. The leader peptide is 39 amino acids long and similar in sequence to those of Cex, CenA, CenB, and CenC (not shown). This is followed by a catalytic domain of 405 amino acids. The amino acid sequence of the catalytic domain is 49 to 68% similar to those of the other catalytic domains in family A, subtype 1 (Fig. 4). The sequence N168-E169-P170 of CenD is found in most family A enzymes, and it has been shown that the strictly conserved Glu residue corresponding to E169 of CenD is essential for the catalytic activity of at least some of them (2).

The catalytic domain is connected to the rest of the polypeptide by a short linker of three Gly residues (Fig. 3). The linker is followed by two repeats of about 95 amino acids each, which are also joined by a linker of three Gly residues (Fig. 3). The repeats are 45 to 50% identical in sequence to three repeats of 98 amino acids in CenB (28) (Fig. 5).

The second repeat sequence is connected to a C-terminal domain of 105 amino acids by the short linker Pro-Thr-Thr (Fig. 3). The amino acid sequence of the C-terminal domain is about 50% identical to those of CBDs in Cex, CenA, CenB, and a number of other β-1,4-glucanases and -xylanases (30) (Fig. 6).

Expression of *cenD* in *E. coli* and purification of CenD. The wild-type *cenD* gene was not expressed in *E. coli*(pDAM1) regardless of the promoter used to transcribe the gene from this plasmid. This was reminiscent of the poor expression of the wild-type *cenC* gene in *E. coli* (32). Both *cenC* and *cenD* have a GTG rather than an ATG translational start codon. Furthermore, there is an inverted repeat less than 70 bp upstream of the *cenD* start codon; this might serve as a transcriptional stop signal for the upstream open reading frame and cause the lack of expression of *cenD* in *E. coli*(pDAM1). *cenD* was expressed by changing the GTG start codon to an ATG and by deleting most of the noncoding *C. fimi* DNA upstream of the gene, including more than half of the inverted repeat (see Materials and Methods and Fig. 3).

The yield of CenD in *E. coli* DH5α(pDAM2-1) was estimated from Coomassie blue-stained SDS-polyacrylamide gels to be about 5% of total cell protein (data not shown).

Catalytic activities of CenD. The catalytic activities of CenD on various cellulosic substrates were determined and compared with those of *C. fimi* CenA and CenB (Table 2). The rate of Avicel hydrolysis was similar for all three endoglucanases, but there were differences in the rates of hydrolysis of the other insoluble substrates. BMCC was hydrolyzed at comparable rates by CenD and CenB, but the rate of hydrolysis catalyzed by CenA was approximately 20-fold lower. The rates of hydrolysis of regenerated cellulose and cellulose azure by CenA were significantly higher than those of CenD or CenB. CenD catalyzed the hydrolysis of CM-cellulose at a 20-fold lower rate than CenB.

The hydrolysis of BMCC (a highly crystalline form of cellulose; see Discussion) by CenB and CenD was examined further in order to determine the extent to which these enzymes are capable of hydrolyzing this substrate to soluble sugars. This was done by determining the total sugar released into the reaction supernatant, because quantitation from the amount of reducing-sugar groups released (expressed as glucose) underestimates the extent of hydrolysis if the product contains significant levels of cellobiose or higher cellooligosaccharides. Solubilization of BMCC by CenD amounted to 85% of the total substrate during a 24-h incubation. Comparable results (87% solubilization) were obtained with CenB.

The xylanase activities of CenD, CenA, CenB, and Cex on azurine-cross-linked xylan were 0.125, 0.005, 0.052, and 195 U/nmol, respectively.

Stereochemical course of hydrolysis catalyzed by CenD. Under the chromatographic conditions described (see Materials and Methods), the retention times of α- and β-cellotriose (7.2 and 7.5 min, respectively) and α- and β-cellotetraose (19.6 and 21.3 min, respectively) were sufficiently different to allow resolution of the anomeric forms; the anomers of cellobiose were not resolved. The major hydrolysis products from cellopentaose were cellobiose and cellotriose. The

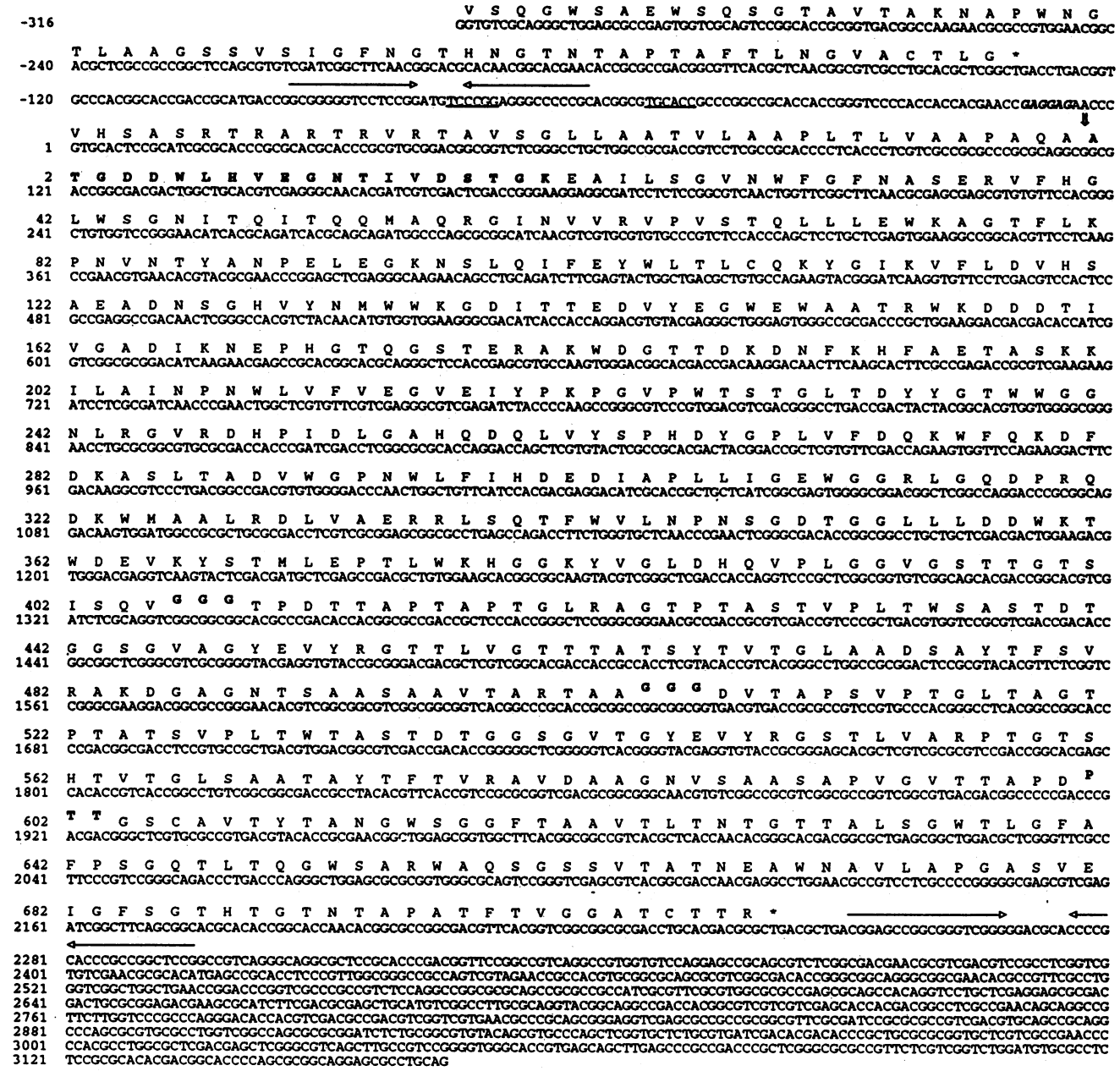

FIG. 3. Nucleotide sequence of *cenD* and deduced amino acid sequence of CenD. The nucleotide sequence is numbered from the first nucleotide (+1) of the translation initiation codon. The stop codon is indicated by an asterisk. Possible transcriptional stop signals are indicated by horizontal arrows. Putative promoter sequences are underlined. The putative ribosome-binding site is shown in boldface italics. Amino acid residues are numbered from the first residue of mature CenD. The N-terminal sequence determined for mature CenD produced in *E. coli* or *C. fimi* is shown in boldface, and the processing site is indicated by a vertical arrow. Putative linker polypeptides are shown in boldface letters raised above the line.

β-anomer of cellotriose was the major product after 15 s of hydrolysis; a small proportion of the β-anomer of cellotetraose was also present (Fig. 7A). After hydrolysis for 1 min at 37°C and incubation at 100°C for 2 min, the 1:2 equilibrium ratio of α- and β-anomers was established by mutarotation (Fig. 7B). Therefore, hydrolysis of the β-1,4-glucosidic bond by CenD proceeds with retention of anomeric configuration, characteristic of a double-displacement reaction mechanism (38).

Immunological evidence for fibronectin type III-like repeats in other Cbps. Polyvalent antiserum directed against CenB appears to react strongly with the fibronectin type III-like repeats because the polypeptide, expressed from pTAL5.1, in which these repeats and the CBD are deleted (29), has a greatly reduced interaction with the antiserum relative to intact CenB, and because the polypeptide p41, comprising the three fibronectin type III-like repeats and CBD of CenB (31), reacted strongly with the antiserum (data not shown).

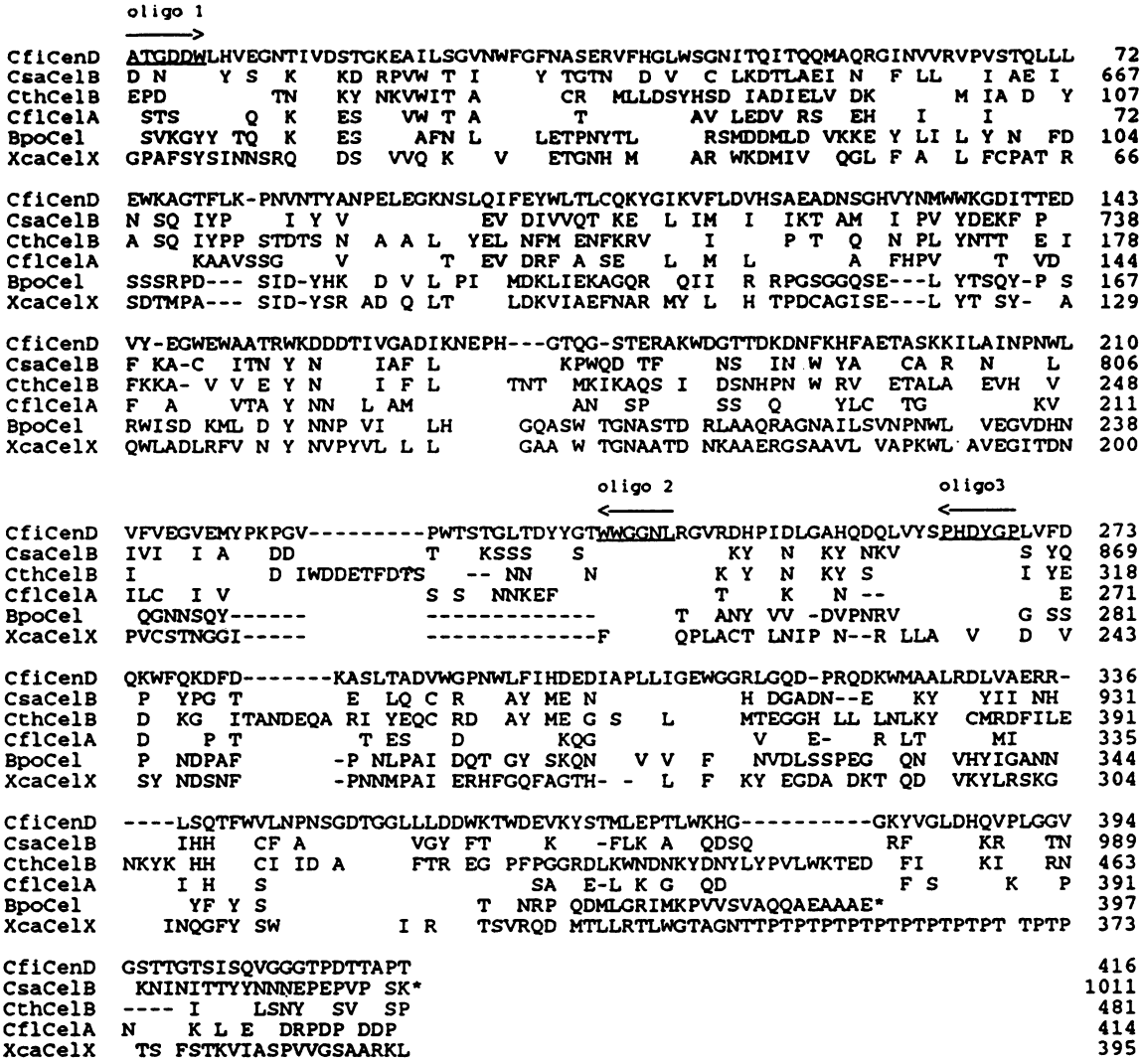

FIG. 4. Comparison of the amino acid sequences of the catalytic domains of family A, subtype 1, β-1,4-glucanases. Sequences are numbered from the first residue of the mature polypeptide. CfiCenD, CenD from *C. fimi*; CsaCelB, CelB from *Caldocellum saccharolyticum*; CthCelB, CelB from *Clostridium thermocellum*; CflCelA, CelA from *Cellulomonas flavigena*; BpoCel, endoglucanase from *Bacillus polymyxa*; XcaCelX, CelX from *Xanthomonas campestris*. Amino acid residues identical to those of CenD are not shown; hyphens indicate gaps left to improve the alignment. A C terminus deduced from a stop codon in the corresponding DNA sequence is marked with an asterisk. Underlined residues are those chosen to synthesize degenerate oligodeoxynucleotides for PCR; the orientations of the oligodeoxynucleotides are indicated by arrows.

The antiserum does not appear to react strongly with the *C. fimi*-type CBD because it does not recognize Cex and CenA (Fig. 8, lanes 2 and 3, respectively), both of which contain this domain. The CenB antiserum reacted with CenB (Fig. 8, lane 1) and with the 95- and 75-kDa Cbps from *C. fimi* culture supernatant, which presumably correspond to Cbp95 and CenD, respectively (Fig. 8, lanes 4 and 5). These data suggest that the fibronectin type III-like repeats of CenB and CenD contain related epitopes and that Cbp95 may also contain similar repeats. DNA sequence analysis has now confirmed the occurrence of an open reading frame encoding three such repeats in the gene for Cbp95 (27). There was no apparent reaction of the antiserum with Cbp120 (Fig. 8, lanes 4 and 5).

DISCUSSION

CenD is the first family A β-1,4-glucanase identified in *C. fimi*. It is an endoglucanase with a catalytic domain of 405 amino acid residues. All the known family A enzymes are endo-β-1,4-glucanases, with the exception of a β-1,4-mannanase from *Caldocellum saccharolyticum* (20), and all have a catalytic domain of 350 to 500 amino acid residues. CenD, like all family A enzymes whose stereochemistry has been analyzed (3, 10), catalyzes hydrolysis with retention of anomeric carbon configuration. It is anticipated that all enzymes within a given family of catalytic domains will have the same stereochemical course of hydrolysis and similar active-site topologies (10, 12).

C. fimi CenB and CenD, but not CenA, CenC, or Cex, contain fibronectin type III-like repeats. There are three copies of the repeat in CenB and two copies in CenD. In both enzymes, the repeats join a C-terminal CBD to a catalytic domain. There is strong evidence for three similarly organized repeats in *C. fimi* Cbp95. One copy of the repeat, followed by a *C. fimi*-type CBD, is encoded by an open

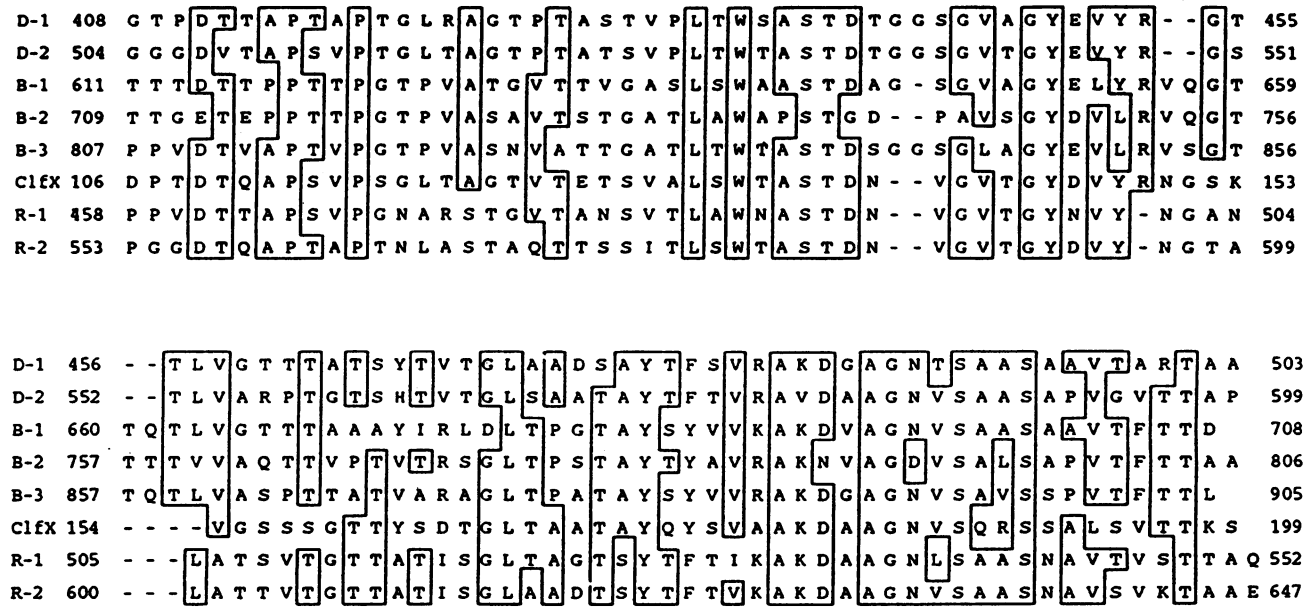

FIG. 5. Comparison of fibronectin type III-like sequences from bacterial polypeptides. D-1 and D-2 are the repeats in CenD of *C. fimi*; B-1, B-2, and B-3 the repeats in CenB of *C. fimi*. ClfX is the single fibronectin type III-like sequence in an open reading frame from *C. flavigena*; R-1 and R-2 are the repeats from chitinase A1 from *B. circulans* WL-12. Residues that are identical in at least five sequences are boxed.

reading frame from *Cellulomonas flavigena* (1), but fibronectin type III-like repeats are absent in more than 100 other bacterial and fungal β-1,4-glucanases whose genes have been sequenced (12, 20). The only other enzymes known to contain such repeats are *Bacillus circulans* chitinase A1 (two copies) (40), *Clostridium thermohydrosulfuricum* α-amylase-pullulanase (two copies) (19), and *Alcaligenes faecalis* poly-3-hydroxybutyrate depolymerase (one copy) (19). The function(s) of these structures is presently unknown, but they are not essential for catalytic activity in CenB or chitinase A1 (15, 29, 31, 40).

The catalytic activities of CenD distinguish the enzyme from other *C. fimi* endo-β-1,4-glucanases (Table 2). CenD has a relatively high activity towards partially crystalline and highly crystalline cellulose preparations (Avicel and BMCC, respectively) and a relatively low activity towards regenerated cellulose and soluble CM-cellulose. The extent of BMCC hydrolysis by CenD after 24 h amounted to 85% of the total cellulose in the reaction mixture, based on total soluble sugar released into the reaction supernatant; the extent of hydrolysis by CenB was similar (87%). BMCC is reported to be approximately 76% crystalline, relative to highly crystalline algal cellulose (23). These data indicate that both CenD and CenB are able to hydrolyze crystalline cellulose to a significant extent. Some family A β-1,4-glucanases are reported to have a low level of β-1,4-xylanase activity (4, 20); a low level of xylanase activity was also found for CenD. The rate of xylan hydrolysis catalyzed by

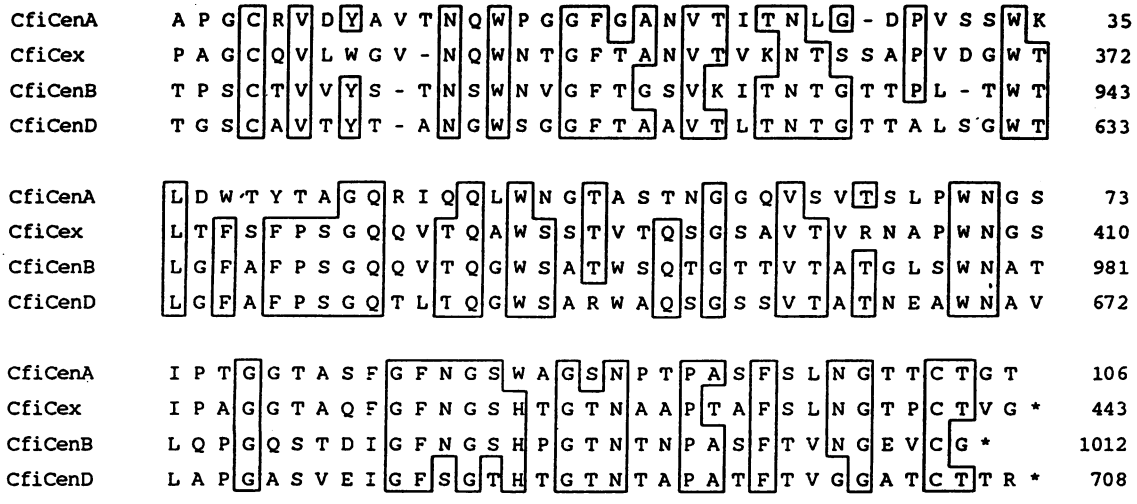

FIG. 6. Comparison of CBDs from *C. fimi* β-1,4-glycanases. The CBDs from CenA, CenB, CenD, and Cex (CfiCenA, CfiCenB, CfiCenD, and CfiCex, respectively) are shown. Residues that are identical in at least three sequences are boxed.

TABLE 2. Comparison of the catalytic activities of CenA, CenB, and CenD on five cellulosic substrates

Cbp	Activity				
	Avicel[a]	BMCC[a]	CM-cellulose[a]	Cellulose azure[b]	Regenerated cellulose[a]
CenA	1.98	0.23	280	107	303
CenB	1.98	10.17	1,030	16	29
CenD	1.97	10.17	77	35	28

[a] Activity expressed as moles of reducing sugar (as glucose) per mole of enzyme per minute.

[b] Activity expressed as units per micromole; one unit of activity results in a change of one absorbance unit in the reaction supernatant per hour.

CenD was higher than that for CenA or CenB but approximately 1,500-fold lower than that for Cex, a *C. fimi* β-1,4-xylanase and β-1,4-glucanase.

Many bacterial and fungal β-1,4-glycanases implicated in the degradation of plant biomass have been shown to contain structurally and functionally independent CBDs which may or may not be related to the *C. fimi*-type CBD (12). These include enzymes other than β-1,4-glucanases: for example, two xylanases and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* (18, 22). The present investigation demonstrates that adsorption on cellulose or Sephadex provides a useful screen for the identification of such enzymes in culture supernatants and facilitates the subsequent cloning of their corresponding genes. CenD was shown to be an endo-β-1,4-glucanase; the substrate specificities of Cbp120 and Cbp95 remain to be determined. However, not all β-1,4-glycanases within a given system are necessarily recovered by this strategy: *C. fimi* CenC and Cex, both of which contain CBDs, were not recovered from the culture supernatant, presumably, because they were not induced under the conditions employed.

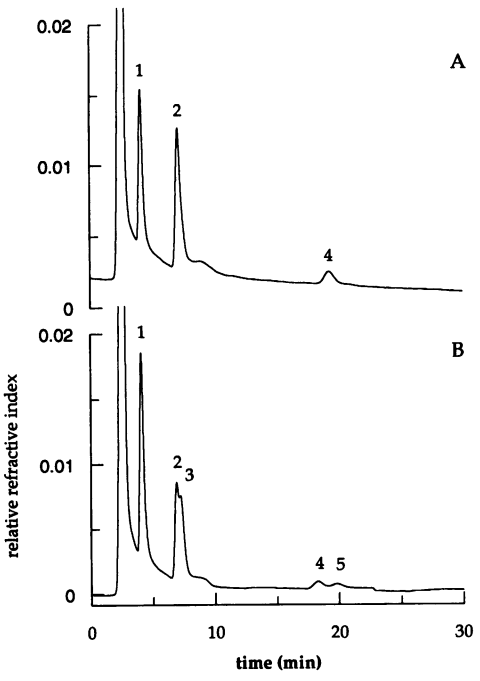

FIG. 7. High-performance liquid chromatographic analysis of the hydrolysis of cellopentaose catalyzed by CenD to determine the stereochemical outcome. (A) Hydrolysis products from cellopentaose after 15 s of incubation with CenD at room temperature; (B) products after incubation at 37°C for 1 min and then at 100°C for 2 min, conditions which allow mutarotation at the anomeric carbon. Details of the analysis are as described in Materials and Methods. The α and β anomers of cellotriose (peaks 3 and 2, respectively) and cellotetraose (peaks 5 and 4, respectively) are resolved; those of cellobiose (peak 1) are not.

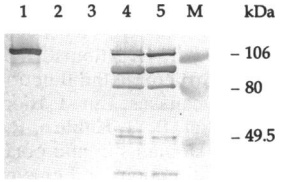

FIG. 8. Reaction of polyvalent antiserum directed against CenB with CenA, CenB, Cex, and Cbps from *C. fimi* culture supernatant. Following SDS-PAGE (7.5% acrylamide), polypeptides were transferred to a nitrocellulose membrane and detected with polyvalent antiserum directed against CenB. Lane 1, 1 μg of CenB; lane 2, 1 μg of Cex; lane 3, 1 μg of CenA; lane 4, polypeptides adsorbed to Avicel from 5 ml of *C. fimi* culture supernatant; lane 5, polypeptides adsorbed to BMCC from 5 ml of *C. fimi* culture supernatant; lane M, molecular mass standards, with sizes indicated on the right.

ACKNOWLEDGMENTS

We thank Curtis Braun and Emily Kwan for helpful advice and expert technical assistance.

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada. A.M. thanks the Landesgraduiertenförderungsgesetz, Land Baden-Württemberg, for a scholarship.

REFERENCES

1. **Al-Tawheed, A. R.** 1988. M.Sc. thesis, Trinity College, Dublin, Ireland.
2. **Baird, S. D., M. A. Hefford, D. A. Johnson, W. L. Sung, M. Yaguchi, and V. L. Seligy.** 1990. The Glu residue in the conserved Asn-Glu-Pro sequence of two highly divergent endo-β-1,4-glucanases is essential for enzymatic activity. Biochem. Biophys. Res. Commun. **169:**1035–1039.
3. **Barras, F., I. Bortoli-German, M. Bauzan, J. Rouvier, C. Gey, A. Heyraud, and B. Henrissat.** 1992. Stereochemistry of hydrolysis catalysed by endoglucanase Z from *Erwinia chrysanthemi.* FEBS Lett. **300:**145–148.
4. **Béguin, P.** 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. **44:**219–248.
5. **Béguin, P., N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., G. P. O'Neill, and R. A. J. Warren.** 1987. Cloning of cellulase genes. Crit. Rev. Biotechnol. **6:**129–162.
6. **Chaplin, M. F.** 1986. Monosaccharides, p. 1–36. *In* M. F. Chaplin and J. F. Kennedy (ed.), Carbohydrate analysis: a practical approach. IRL Press, Oxford.
7. **Coutinho, J. B., N. R. Gilkes, R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1992. The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and Sephadex is mediated by the N-terminal repeats. Mol. Microbiol. **6:**1243–1252.
8. **Coutinho, J. B., B. Moser, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1991. Nucleotide sequence of the endoglucanase C gene (*cenC*) of *Cellulomonas fimi*, its high-level expression in *Escherichia coli*, and characterization of its products. Mol. Microbiol. **5:**1221–1233.
9. **Feinberg, A., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction and endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
10. **Gebler, J., N. R. Gilkes, M. Claeyssens, D. B. Wilson, P. Béguin, W. W. Wakarchuk, D. G. Kilburn, R. C. Miller, Jr., R. A. J. Warren, and S. G. Withers.** 1992. Stereoselective hydrolysis catalyzed by related β-1,4-glucanases and β-1,4-xylanases. J.

Biol. Chem. **267**:12559–12561.
11. **Gilkes, N. R., M. Claeyssens, R. Aebersold, B. Henrissat, A. Meinke, H. D. Morrison, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1991. Structural and functional relationships in two families of β-1,4-glycanases. Eur. J. Biochem. **202**:367–377.
12. **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**:305–315.
13. **Gilkes, N. R., E. Jervis, B. Henrissat, B. Tekant, R. C. Miller, Jr., R. A. J. Warren, and D. G. Kilburn.** 1992. The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. J. Biol. Chem. **267**:6743–6749.
14. **Gilkes, N. R., D. G. Kilburn, M. L. Langsford, R. C. Miller, Jr., W. W. Wakarchuk, R. A. J. Warren, D. J. Whittle, and W. R. K. Wong.** 1984. Isolation and characterization of *Escherichia coli* clones expressing cellulase genes from *Cellulomonas fimi*. J. Gen. Microbiol. **130**:1377–1384.
15. **Gilkes, N. R., R. A. J. Warren, R. C. Miller, Jr., and D. G. Kilburn.** 1988. Precise excision of the cellulose-binding domains from two *Cellulomonas fimi* cellulases by an homologous protease and the effect on catalysis. J. Biol. Chem. **263**:10401–10407.
16. **Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1987. Regulation and initiation of *cenB* transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:4674–4677.
17. **Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1987. Regulation, initiation, and termination of the *cenA* and *cex* transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:646–653.
18. **Hall, J., G. P. Hazlewood, N. S. Huskisson, A. J. Durrant, and H. J. Gilbert.** 1989. Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subsp. *cellulosa*: internal signal sequence and unusual protein processing. Mol. Microbiol. **3**:1211–1219.
19. **Hansen, C. K.** 1992. Fibronectin type III-like sequences and a new domain in prokaryotic depolymerases with insoluble substrates. FEBS Lett. **305**:91–96.
20. **Henrissat, B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280**:309–316.
21. **Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle, and J.-P. Mornon.** 1989. Cellulase families revealed by hydrophobic cluster analysis. Gene **81**:83–95.
22. **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-binding domains and are encoded by adjacent genes. Biochem. J. **272**:369–376.
23. **Kulshreshtha, A. K., and N. E. Dweltz.** 1973. Paracrystalline lattice disorder in cellulose. I. Reappraisal of the application of the two-phase hypothesis to the analysis of powder X-ray diffractograms of native and hydrolysed cellulosic materials. J. Polymer Sci. **11**:487–497.
24. **Langsford, M. L., N. R. Gilkes, B. Singh, B. Moser, R. C. Miller, Jr., R. A. J. Warren, and D. G. Kilburn.** 1987. Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett. **225**:163–167.
25. **Langsford, M. L., N. R. Gilkes, W. W. Wakarchuk, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1984. The cellulase system of *Cellulomonas fimi*. J. Gen. Microbiol. **130**:1367–1376.
26. **Lever, M.** 1972. A new reaction for colorimetric determination of carbohydrates. Anal. Biochem. **47**:273–279.
27. **Meinke, A.** Unpublished data.
28. **Meinke, A., C. Braun, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Unusual sequence organization in CenB, an inverting endoglucanase from *Cellulomonas fimi*. J. Bacteriol. **173**:308–314.
29. **Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: functions and relatedness to domains in other polypeptides. J. Bacteriol. **173**:7126–7135.
30. **Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Bacterial cellulose-binding domain-like sequences in eucaryotic polypeptides. Protein Sequence Data Anal. **4**:349–353.
31. **Meinke, A., M. Schmuck, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1992. The tertiary structure of endo-β-1,4-glucanase B (CenB), a multidomain cellulase from the bacterium *Cellulomonas fimi*. Glycobiology **2**:321–326.
32. **Moser, B., N. R. Gilkes, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1989. Purification and characterization of endoglucanase C of *Cellulomonas fimi*, the cloning of its gene, and analysis of in vivo transcripts of the gene. Appl. Environ. Microbiol. **55**:2480–2487.
33. **O'Neill, G. P., 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.
34. **Owolabi, J. B., P. Béguin, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1988. Expression in *Escherichia coli* of the *Cellulomonas fimi* structural gene for endoglucanase B. Appl. Environ. Microbiol. **54**:518–523.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
37. **Scopes, R. K.** 1974. Measurement of protein by spectrophotometry at 205 nm. Anal. Biochem. **59**:277–282.
38. **Sinnott, M. L.** 1988. Glycosyl group transfer, p. 259–296. *In* M. I. Page and A. Williams (ed.), Enzyme mechanisms. Royal Society of Chemistry, London.
39. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
40. **Watanabe, T., K. Suzuki, K. Oyanagi, K. Ohnishi, and H. Tanaka.** 1990. Gene cloning of chitinase A1 from *Bacillus circulans* WL-12 revealed its evolutionary relationship to *Serratia* chitinase and to the type III homology units of fibronectin. J. Biol. Chem. **265**:15659–15665.
41. **Whittle, D. J., D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1982. Molecular cloning of a *Cellulomonas fimi* cellulase gene in *Escherichia coli*. Gene **17**:139–145.
42. **Wong, W. K. R., B. Gerhard, Z. M. Guo, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1986. Characterization and structure of an endoglucanase gene *cenA* of *Cellulomonas fimi*. Gene **44**:315–324.