Expression in *Escherichia coli* of the *Cellulomonas fimi* Structural Gene for Endoglucanase B

Joshua B. Owolabi, Pierre Beguin, Douglas G. Kilburn, Robert C. Miller Jr. and R. Antony J. Warren

***Appl. Environ. Microbiol.* 1988, 54(2):518.**

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Feb. 1988, p. 518–523
0099-2240/88/020518-06$02.00/0

Vol. 54, No. 2

Expression in *Escherichia coli* of the *Cellulomonas fimi* Structural Gene for Endoglucanase B

JOSHUA B. OWOLABI,[1] PIERRE BEGUIN,[1,2] DOUGLAS G. KILBURN,[1] ROBERT C. MILLER, JR.,[1] AND R. ANTONY J. WARREN[1]*

Department of Microbiology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5,[1] and Unité de Physiologie Cellulaire, Department de Biochimie et Genetique Moleculaire, Institut Pasteur, F-75724 Paris Cedex 15, France[2]

Received 24 July 1987/Accepted 24 November 1987

Endoglucanase B (EB) of *Cellulomonas fimi* has an M_r of 110,000 when it is produced in *Escherichia coli*. The level of expression of the *cenB* gene (encoding EB) was significantly increased by replacing its normal transcriptional and translational regulatory signals with those of the *E. coli lac* operon. EB was purified to homogeneity from the periplasmic fraction of *E. coli* in one step by affinity chromatography on microcrystalline cellulose (Avicel). Alignment of the NH_2-terminal amino acid sequence with the partial nucleotide sequence of a fragment of *C. fimi* DNA showed that EB is preceded by a putative signal polypeptide of 33 amino acids. The signal peptide functions and is processed correctly in *E. coli*, even when its first 15 amino acids are replaced by the first 7 amino acids of β-galactosidase. The intact EB polypeptide is not required for enzymatic activity. Active polypeptides with M_rs of 95,000 and 82,000 also appear in *E. coli*, and a deletion mutant of *cenB* encodes an active polypeptide with an M_r of 72,000.

The gram-positive bacterium *Cellulomonas fimi* produces a complex mixture of enzymes active against carboxymethyl cellulose (CMC) during growth on microcrystalline cellulose (Avicel) (3, 4, 22). Several of the enzymes bind very tightly to the substrate, and others bind weakly to the substrate and are removed by water or dilute buffer (3, 4, 22; N. R. Gilkes personal communications). Two of the tightly bound enzymes have been characterized in detail. They are eluted from microcrystalline cellulose with 6 M guanidinium hydrochloride. One of them is an exoglucanase (EXG; 1,4-β-D-glucan cellobiohydrolase; EC 3.2.1.91); the other is an endoglucanase, endoglucanase A (EA; 1,4-β-D-glucan glucanohydrolase; EC 3.2.1.4) (11, 12).

The genes for the enzymes, *cex* for the EXG and *cenA* for the EA, have been sequenced previously (31, 46). The amino acid sequences of the enzymes, which were predicted from the nucleotide sequences of their genes, show that each enzyme contains three distinct regions: a short sequence of about 20 amino acids containing only proline and threonine (the Pro-Thr box); an irregular region that is predicted to lack secondary structure, which is rich in hydroxyamino acids but of low charge density; and an ordered region with a secondary structure and a higher charge density. The amino acid sequences of the irregular regions are 50% conserved; those of the regular regions do not appear to be conserved (44). The two proteins are very similar in size. Each protein contains a single polypeptide of 443 amino acids in EXG and 418 amino acids in EA, with molecular weights of 47,000 and 43,800, respectively.

The gene for a second endoglucanase, *cenB* has been also cloned (11, 12). Results of preliminary analyses indicated that the enzyme that it encodes, endoglucanase B (EB), has an M_r of about 110,000, i.e., more than twice those of EXG and EA. The considerable difference in size prompted a closer examination of *cenB* and its encoded polypeptide. In this report we describe the initial characterization of the *cenB* gene, its manipulation to give significantly increased levels of expression in *Escherichia coli*, and the identification of its product in and purification from *E. coli*.

MATERIALS AND METHODS

Bacterial strains, media, and vectors. Except for *E. coli* BD1854, *E. coli* RR1, and *E. coli* JM101, the bacterial strains and the media used for their cultivation have been described previously (11, 45). *E. coli* BD1854 was the strain that was used to identify plasmid-coded proteins by the minicell method (19). *E. coli* RR1 was the host for the pUC-derived plasmids (33). *E. coli* JM101 was the host for M13 phage vectors (47). Plasmids pBR322 (6), pUC19 (27), pEC3 (11), and pDR540 (37) and phages M13 mp11 and M13 mp18 (25, 47) have been described previously.

Preparation of proteins, enzymatic assays, and protein determination. Proteins from total cell extracts were prepared by breaking the cells with a French press (45). Periplasmic proteins were isolated by osmotic shock (28). Cytoplasmic proteins were prepared by rupturing the osmotically shocked cells with a French press. CMCase activity was determined as described previously (11). β-Lactamase activity was assayed spectrophotometrically by using nitrocefin as the substrate (30). β-Galactosidase was measured with *o*-nitrophenyl-β-D-galactopyranoside as the substrate (26). One unit of endoglucanase (CMCase) released 1 nmol of glucose equivalent min^{-1}. One unit of β-galactosidase released 1 nmol of *o*-nitrophenol min^{-1}. Zones of hydrolysis around colonies on CMC plates were detected with Congo red (41). Protein was determined by the dye binding assay (7) by using bovine plasma albumin as the standard.

DNA isolation and fractionation. Plasmid DNA for restriction analysis was isolated by the alkaline lysis procedure (5). The plasmid DNA to be sequenced was purified by banding in CsCl-ethidium bromide density gradients (23). The M13 replicative-form and viral DNAs were isolated from infected cultures (25). DNA restriction fragments were resolved by agarose gel electrophoresis (23).

DNA sequence determination. Fusion regions of deletions at the 5′ end of the insert in pJB3, which were generated with

* Corresponding author.

exonuclease III and S1 nuclease (15, 34), were sequenced directly as described previously (17). The 400-base-pair *Bam*HI-*Pst*I fragment of pJB3 and the 5′ end of some deleted inserts were subcloned into M13 mp11, M13 mp18, or both and then sequenced by the dideoxy method (38).

Minicells. Minicells were prepared and labeled with [^{35}S]methionine by a modification of the method described by Jensen et al. (19).

Electrophoretic analysis of proteins. Proteins were separated by electrophoresis in 7.5 or 10% polyacrylamide gels in the presence of 0.1% sodium dodecyl sulfate (SDS) (21). [^{35}S]methionine-labeled proteins were located by autoradiography of the dried gels. Zymograms were obtained by the agarose gel overlay technique (2).

Affinity chromatography of EB. The recombinant EB was purified by a modification of the procedures described by Halliwell and Griffin (16) and Nummi et al. (29). *E. coli* RR1(pJB301) cells were grown in 40 liters of Luria broth supplemented with ampicillin (100 μg ml^{-1}) (11) in a stirred fermentor at 30°C. Exponentially growing cells were harvested with a centrifuge (Sharples) and washed with 2 liters of ice-cold, 0.01 M Tris hydrochloride (pH 7.1)–0.03 M NaCl. The periplasmic fraction of the cells was prepared (28) and kept in buffer A (50 mM potassium phosphate [pH 6.9], 0.02% sodium azide) containing 20 μg of phenylmethylsulfonyl fluoride ml^{-1} at 4°C until needed.

Microcrystalline cellulose (Avicel) was stirred gently in distilled water (20 ml g^{-1}) and centrifuged at 4,000 × *g* for 5 min to remove fine particles. This step was repeated 4 times, and the microcrystalline cellulose was finally suspended in the same volume of water. After it was autoclaved for 30 min (121°C, 15 lb/in^2), it was left to cool to room temperature (22°C). The aqueous phase was replaced and discarded

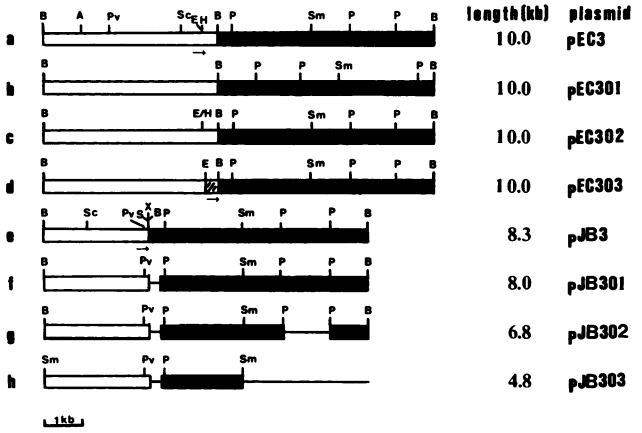

FIG. 1. Diagram of various recombinant plasmids containing *cenB*. The circular plasmids are shown in a linear fashion for clarity. (a) pEC3; (b) pEC301; (c) pEC302; (d) pEC303; (e) pJB3; (f) pJB301; (g) pJB302; and (h) pJB303. The open bar represents pBR322 and pUC19 DNA in the pEC and pJB plasmids, respectively; the solid bar represents *C. fimi* DNA; the hatched bar represents the 0.3-kbp fragment from pDR540 containing the *tac* promoter; the single line represents the regions that were deleted in each derivative. The arrows indicate the functional orientations for the *tet* promoter in rows a and b; the *tac* promoter in row d; and the *lac* promoter in rows e, f, g, and h. The *tet* promoter was deleted in row c. The total length of each plasmid is indicated; the *C. fimi* inserts were 5.6 kbp in rows a, b, c, d, and e; 5.3 kbp in row f; 4.1 kbp in row g; and 2.1 kbp in row h. The deletion in row h extends to the *Sma*I (Sm) site of pUC19. Only relevant restriction sites are shown. Abbreviations: A, *Ava*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; Pv, *Pvu*II; S, *Sph*I; Sc, *Sca*I; Sm, *Sma*I; X, *Xba*I.

TABLE 1. Endoglucanase activities of various *cenB* clones

Host strain	Plasmid[a]	CMCase sp act[b]
C600	pEC3	7.61
	pEC301	0.28
	pEC302	0.16
	pEC303	51.10
RR1	pJB3	38.9
	pJB301	146.7
	pJB302	167
	pJB303	156.7

[a] For details of plasmid structure, see Fig. 1. pEC3 and pJB3 are the parental clones containing the entire *cenB* gene. pEC301, pEC302, and pEC303 are derivatives of pEC3 that were produced to investigate control of transcription. pJB301 is a gene fusion of *lacZ* and *cenB*. pJB302 and pJB303 are deletion derivatives of pJB301.

[b] CMCase specific activity is expressed as nanomoles of glucose equivalents released per minute per milligram of soluble cell protein.

twice, as described above, and finally replaced with buffer A. After overnight equilibration at 4°C, the liquid was removed. The wet, settled microcrystalline cellulose was suspended in ice-cold buffer A (4 ml g^{-1}) and kept on ice.

A 350-ml volume of periplasmic fluid (from 10 liters of culture) was mixed with 150 ml of the autoclaved microcrystalline cellulose (which contained approximately 25 g of dry microcrystalline cellulose) and kept on ice for 1 h. All subsequent steps were carried out at room temperature. The unadsorbed material was separated from the microcrystalline cellulose by filtration (GF/C filter; Whatman, Inc., Clifton, N.J.) through sintered glass. The microcrystalline cellulose-enzyme complex was washed once with buffer A by filtration as described above, suspended in half its volume of fresh buffer, and packed into a column (2.5 by 13 cm). Proteins bound to microcrystalline cellulose were eluted with a concave descending gradient of buffer A (55 ml) and water (800 ml) at a flow rate of 30 ml/h. Fractions were assayed for endoglucanase activity (11). EB-containing fractions were pooled, lyophilized, suspended in distilled water, and desalted by gel filtration by using Bio-Gel P-6DG.

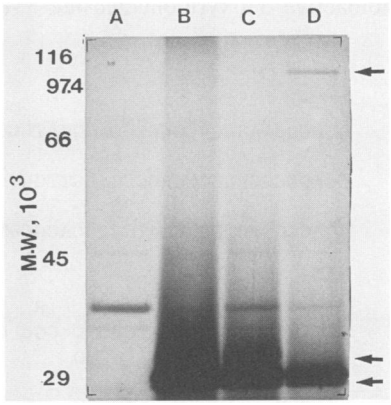

FIG. 2. Autoradiogram of polypeptides encoded by various plasmids. *E. coli* BD1854 was transformed with plasmids pBR322, pEC3, and pEC303. The proteins encoded by the plasmids were labeled in minicells (19). The labeled proteins were analyzed by SDS-polyacrylamide gel electrophoresis. Lane A, BD1854; lane B, BD1854(pBR322); lane C, BD1854(pEC3); lane D, BD1854 (pEC303). The molecular weight (M.W.) standards are as described in the legend to Fig. 3. The positions of the *bla* and *cenB* gene products are indicated by arrows. The exposure was too short to visualize the *tet* gene product.

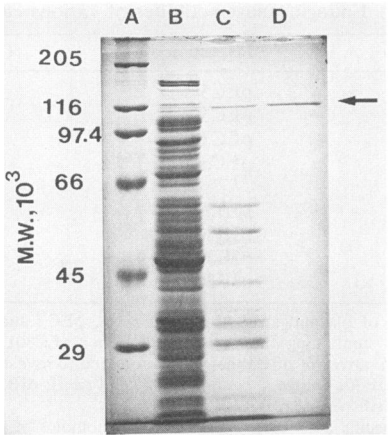

FIG. 3. SDS-polyacrylamide gel electrophoretic analysis of samples during the purification of EB (indicated by the arrow). Lane B, Total cell extract; lane C, periplasmic fraction; lane D, pooled active fraction from the microcrystalline cellulose column; lane A, molecular weight (M.W.) standards (rabbit muscle myosin, 205,000; *E. coli* β-galactosidase, 116,000; rabbit muscle phosphorylase *b*, 97,400; bovine serum albumin 66,000; ovalbumin, 45,000; carbonic anhydrase, 29,000).

Amino acid sequence analysis. The purified EB was sequenced by automated Edman degradation with a gas-phase sequenator (model 470A; Applied Biosystems).

Enzymes and reagents. All restriction endonucleases were purchased from Bethesda Research Laboratories, Burlington, Canada, or Pharmacia, Dorval, Canada. Calf intestinal phosphatase was from Boehringer Mannheim, Dorval, Canada. Isopropyl-β-D-thiogalactopyranoside, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, CMC (low viscosity), phenylmethylsulfonyl fluoride, and *o*-nitrophenyl-β-D-galactopyranoside were from Sigma Chemical Co., St. Louis, Mo. Microcrystalline cellulose (Avicel; type PH-101) was from FMC International, Cork, Ireland; nitrocefin was a gift from Glaxo Group Research Ltd., Greenford, United Kingdom. SDS was from BDH Biochemicals, Toronto, Canada. Radioactive deoxyribonucleoside 5′-triphosphates and [^{35}S]methionine were from New England Nuclear Research Products, Boston, Mass., and Amersham Canada Ltd., Oakville, Canada, respectively.

RESULTS

Expression of *cenB* in *E. coli*. Plasmid pEC3 contains a 5.6-kilobase-pair (kbp) *Bam*HI fragment of *C. fimi* DNA cloned into the *Bam*HI site of pBR322 (11) (Fig. 1). Expression of the endoglucanase activity that it encodes is dependent on the *tet* promoter, because inversion of the *C. fimi* insert with respect to the *tet* promoter (39, 40) (Fig. 1, pEC301) or deletion of the *tet* promoter drastically reduced the expression (Table 1). When the 5.6-kbp *Bam*HI fragment was fused to the *tac* promoter by replacing the small *Eco*RI-*Bam*HI fragment of pEC3 containing the *tet* promoter with the 300-base-pair *Eco*RI-*Bam*HI fragment from pDR540 (37) (Fig. 1, pEC303), expression of *cenB* was increased sevenfold (Table 1). The polypeptide encoded by *cenB* was detected in an *E. coli* minicell system (Fig. 2). It had an M_r of 110,000.

Delineation of the *cenB* gene. The 5.6-kbp *Bam*HI fragment of pEC3 was subcloned into pUC19 such that the 5′ end was adjacent to the *lac* promoter and operator region to give plasmid pJB3 (Fig. 1). The plasmid was linearized at the *Xba*I site, end-labeled with [α-^{32}P]dCTP, and digested to completion with *Sph*I. Then, exonuclease III was used to delete DNA from the 5′ end of the insert as described previously (15, 34). After S1 nuclease digestion, ligation, and transformation, Amp^r colonies were screened for endoglucanase activity. The extents of the deletions in active clones were determined by restriction analyses and by sequencing of the fusion regions of the plasmids.

Deletions of more than 385 base pairs from the 5′ end of the insert prevented expression of *cenB* (data not shown). The most active deletion plasmid, pJB301, lost 319 base pairs from the 5′ end of the insert, resulting in an in-frame fusion between codon 16 for the EB signal peptide and codon 7 for the α-fragment of β-galactosidase (data not shown). The level of activity determined by pJB301 was significantly higher than those observed with pJB3 and pEC303 (Table 1).

Purification of EB. The EB protein that was present in the

```
                                                                GGATCCCGCGCCCGGCGCGA
                                                                1                  20
GCCCGCAACCCACGCGCCCACGGATCGGGCCTCACGAGCCCGACGTTGGCGGCCGGGCCGGGGGGCGACCTCGAGACCGA
                    40                  60                  80                  100
GGAGCCCCCGCGTGAGGCGACGTTGGCCGCGCACGCCGCTGGTGAGCGGGCTGAATCGTTTAGGGCGTTGACCTGCGGAC
                  120                 140                 160                 180
GGACCCGTCTGGACGATGCGCCAGGCGTCGTGCGGGTGCGACTGCGGACAGCACGGGTCGCCGACCACCACTCCCGTGCC
                  120                 220                 240                 260
                -33         -30                                 -20
                 M   L   R   Q   V   P   R   T   L   V   A   G   G   S   A   L
CGGAAGAGGACCCC ATG CTC CGC CAA GTC CCA CGC ACG CTC GTC GCG GGT GGC TCC GCC CTC
                    280                 300                 320
                            -10                         -1  +1
 A   V   A   V   G   V   L   V   A   P   L   A   T   G   A   A   A   P   T
GCC GTC GCC GTC GGG GTG CTC GTC GCC CCG CTC GCG ACC GGC GCG GCC GCC GCG CCC ACC
                    340                 360                 380
                          10                          20
 Y   N   Y   A   E   A   L   Q   K   S   M   F   F   Y   Q   A   Q   G   G
TAC AAC TAC GCC GAG GCC CTG CAG AAG TCG ATG TTC TTC TAC CAG GCG CAC GGC TCC ...
                  400                 420
```

FIG. 4. Nucleotide sequence and deduced amino acid sequence of the 5′ terminus of the *cenB* gene. The putative ribosome-binding site is underlined. The arrow indicates the signal peptide processing site. The underlined amino acids were determined by automated Edman degradation of the purified recombinant EB.

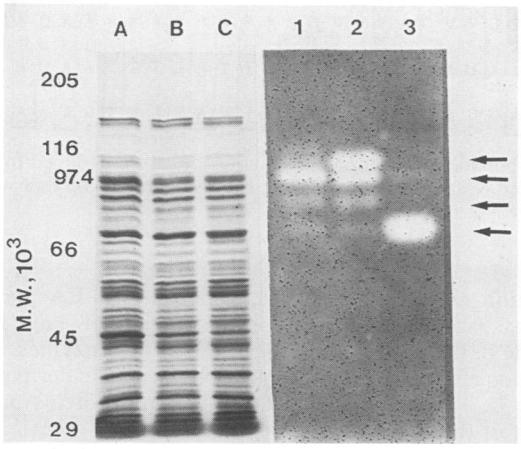

FIG. 5. SDS-polyacrylamide gel electrophoresis and zymograms of total cellular proteins from *E. coli* RR1 containing pJB301 or its deletion derivatives. Lanes A, B, and C are total cellular proteins stained with Coomassie blue. Lane A, RR1(pJB301); lane B, RR1(pJB302); lane C, RR1(pJB303). Lanes 1, 2, and 3 are zymograms of lanes A, B, and C, respectively. Molecular weight (M.W.) standards are as described in the legend to Fig. 3. The arrows indicate active endoglucanase components.

osmotic shock fluid from *E. coli* RR1(pJB301) was purified approximately 50-fold to homogeneity by a one-step procedure in which affinity chromatography on microcrystalline cellulose was used (see above) (Fig. 3). Recovery of activity from the microcrystalline cellulose column was 60%. The purified enzyme had a specific activity of 94,400 U mg^{-1} protein (defined as nanomoles of glucose equivalents released per minute).

Structure of the 5'-terminal region of the *cenB* gene. The nucleotide sequence of the 5' end of the *C. fimi* DNA fragment containing *cenB* is shown in Fig. 4. The sequence corresponding to the amino terminus of the mature enzyme was located by using the amino acid sequence of the protein purified from *E. coli* (see above). This sequence was preceded by one that encoded a putative signal peptide of 33 amino acids, with a hydrophilic NH_2 terminus of 7 amino acids, including 2 arginines, followed by a hydrophobic sequence of 26 amino acids. The translational start codon at nucleotide 275 was preceded by a stretch of nucleotides (GGAAGAGGA) closely resembling other ribosome-binding sites (13, 24, 42). The region upstream of the ribosome-binding site contains a sequence that resembles other *C. fimi* promoter sequences (14).

Localization of the 3' end of the *cenB* gene. pJB301 was cleaved partially with *Sma*I or *Pst*I and religated. This resulted in the deletion of various lengths of DNA on the 3' side of the *cenB* gene. Transformants were screened for endoglucanase activity, and plasmids from active clones were characterized by restriction mapping (Fig. 1). The cells from selected clones were assayed quantitatively for endoglucanase activity (Table 1). The shortest, uninterrupted fragment of the 5.3-kbp insert which expressed endoglucanase activity equal to that of pJB301 was 2.1 kbp long; it was found in pJB303 (Fig. 1 and Table 1). It should be pointed out that a protein with an M_r of 110,000 requires a nucleotide sequence of at least 3 kbp.

Cellular location of EB. Exponential-phase cells of *E. coli* RR1 containing pJB3, pJB301, pJB302, or pJB303 were fractionated by the osmotic shock method (28). The endoglucanase activity of pJB3 appeared in the periplasm (Table 2). In the fusion plasmids pJB301, pJB302, and pJB303, the first 15 amino acids of the EB signal peptide were replaced with the first 7 amino acids of β-galactosidase; and in pJB303, the carboxy-terminal region of the mature polypeptide was deleted. Nonetheless, endoglucanase activity appeared in the periplasms of the cells that carried those plasmids (Table 2).

Characterization of the endoglucanase activities expressed by *E. coli* RR1(pJB301) and deletion derivatives. Cell extracts of *E. coli* containing pJB301, pJB302, and pJB303 were subjected to SDS-polyacrylamide gel electrophoresis. The separated polypeptides were screened for endoglucanase activity. This activity was easily detected in SDS gels prepared from samples that were heated in loading buffer for 2 min at 65°C prior to loading. Control experiments showed no noticeable difference in migration of the polypeptides from gently heated or boiled samples (data not shown). Three active polypeptides with M_rs of 110,000, 95,000, and 82,000 were observed for pJB301 and pJB302; pJB303 encoded an active polypeptide with an M_r of 72,000 (Fig. 5). The numbers and

TABLE 2. Localization of EB, β-lactamase, and β-galactosidase in *E. coli* RR1 cultures

Plasmid	Fraction	Enzyme activity[a] (sp act[b]) of:		
		EB	β-Lactamase	β-Galactosidase
pJB3	Periplasmic	1.9 (148.3)	770 (53,600)	0.85 (67.7)
	Cytoplasmic	3.5 (53.5)	108 (1,120)	355 (3,669)
	Whole cells	5.4 (38.9)	980 (7,210)	427 (4,611)
pJB301	Periplasmic	11.6 (644)	260 (14,330)	7.9 (441)
	Cytoplasmic	5.6 (52.3)	12.1 (92)	123 (1,156)
	Whole cells	18.1 (146.7)	320 (3,400)	142 (1,150)
pJB302	Periplasmic	13.6 (689)	323 (16,400)	11.7 (594)
	Cytoplasmic	7.9 (100)	16 (200)	132.2 (1,679)
	Whole cells	22.1 (167)	371 (3,960)	142 (1,515)
pJB303	Periplasmic	14.1 (726)	420 (21,600)	9.1 (467)
	Cytoplasmic	6.9 (68.3)	20 (200)	208 (2,037)
	Whole cells	22.6 (156.7)	453 (3,700)	265 (1,595)

[a] Enzyme activity is nanomoles of products (see text) released per minute per milliliter of culture.
[b] Specific activity is nanomoles of products released per minute per milligram of protein. Specific activities were computed by using protein concentrations in shock fluid for the periplasmic fractions and soluble cell protein for the whole-cell fractions.

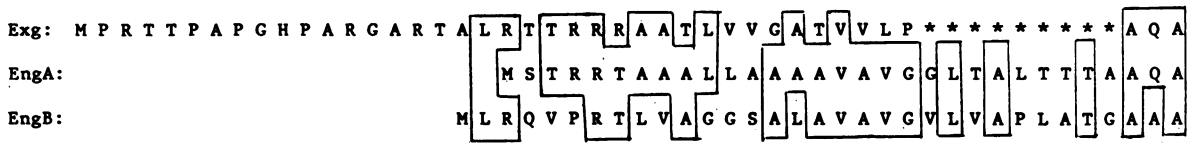

FIG. 6. Comparison of the EXG (Exg), EA (EngA), and EB (EngB) signal peptides. The amino acid sequences were deduced from the following DNA sequences: *cex* (31), *cenA* (46), and *cenB* (this study). Conserved residues are boxed. The asterisks note a gap that was left in the sequence.

sizes of active endoglucanase polypeptides did not differ when *E. coli* PAM163, which is deficient in La protease (20), was transformed with these plasmids (data not shown).

DISCUSSION

To our knowledge, this is the first report of the application of affinity chromatography on cellulose to the purification of the product of a cloned cellulase gene. The procedure is relatively simple, fast, inexpensive, and efficient. Autoclaving of the microcrystalline cellulose (Avicel) resulted in a more specific adsorption of the recombinant cellulase.

The *cenB* gene encoded an endoglucanase with an M_r of 110,000 in *E. coli*. A polyclonal antibody raised against the EB purified from *E. coli* recognizes a *C. fimi* extracellular protein with an M_r of 110,000, as well as six polypeptides with lower molecular weights (J. B. Owolabi, unpublished results). Furthermore, a polypeptide with an M_r of 110,000, which binds weakly to the substrate in *C. fimi* cultures grown with microcrystalline cellulose, has been purified but not sequenced (N. R. Gilkes, personal communication). The relationship of these polypeptides to each other and the identity of the *C. fimi cenB* gene product remain to be determined.

The sequence upstream from the translational start site of *cenB* does not contain an *E. coli*-like promoter (36). It does contain sequences similar to those of two other *C. fimi* promoters (14), however. The *cenB* promoter is weak or nonfunctional in *E. coli*, like those of the *cenA* and *cex* genes of *C. fimi* (31, 46).

Only 0.2% of the total bacterial protein constituted EB when the *cenB* gene was cloned in phase with *lacZ* gene. The apparent low EB expression is in agreement with our earlier data on the expression of the *cex* gene in *E. coli* (32). A *lacZ-cex* fusion plasmid yielded a 15-fold increase in EXG expression when compared with the expression of EXG under its own translational initiation signals. However, fusion of the *cex* gene to a portable translation initiation signal and placement of these sequences under the transcriptional control of the lambda p_L promoter contained in plasmid pCP3 resulted in overproduction of EXG (32).

Although it is longer than most *E. coli* signal peptides, the EB signal peptide allowed export of EB from *E. coli*. Furthermore, replacement of its basic NH_2-terminal section with the NH_2-terminal amino acids of β-galactosidase did not block the processing and export of EB. Deletion of the basic amino acids in the signal peptide of *E. coli* lipoprotein, or their replacement with neutral amino acids, had little effect on lipoprotein export (18, 43). Their replacement with negatively charged amino acids, however, reduced lipoprotein export drastically.

The significance of the hydrophobic region of the EB signal peptide remains to be determined. The EA, EB, and EXG signal peptides exhibit extensive homology in their hydrophobic carboxy-terminal sequences (Fig. 6). The conservation of these sequences, especially in EA and EB, suggests an essential role for this region of each leader peptide in the export and processing of the enzymes. Some mutations in the hydrophobic regions of the leader peptides of the *E. coli* LamB and maltose-binding proteins prevent export of these proteins (1, 9).

The properties of the *cenB* deletion plasmid pJB303 indicate that the intact EB is not required for enzymatic activity. A similar observation has recently been reported for an alkaline cellulase gene from an alkalophilic *Bacillus* sp. (10). It is possible that the active site of EB lies in the NH_2-terminal region and that the multiplicity of the activities encoded by the complete gene is a result of secondary proteolytic processing from the carboxy-terminal region. The limited proteolysis of EB does not appear to result from La protease activity since there was no difference in the endoglucanase activity pattern obtained from an *E. coli lon* mutant (data not shown). Multiple active polypeptides are determined by other genes cloned in *E. coli*: the *celA* gene from *Clostridium thermocellum* (8), a β-1,4-glucanase gene from *Bacillus subtilis* DLG (35), and the *cex* and *cenA* genes from *C. fimi* (Z. M. Guo and N. Arfman, personal communications). We are presently carrying out more detailed biochemical characterization of purified EB.

ACKNOWLEDGMENTS

We thank M. Sandy Kielland and Bob Olafson for the amino acid sequencing and Neil R. Gilkes for information on the purification of *C. fimi* endoglucanases weakly bound to microcrystallin cellulose.

J.B.O. and P.B. are recipients of a Canadian Commonwealth scholarship and a North Atlantic Treaty Organization fellowship, respectively. This research was supported by strategic grant 67-0941 (to D.G.K., R.C.M., and R.A.J.W.) and operating grant 5-83246 (to D.G.K.) from the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

1. **Bedouelle, H., P. J. Bassford, Jr., A. V. Fowler, I. Zabin, J. Beckwith, and M. Hofnung.** 1980. Mutations which alter the function of the signal sequence of the maltose binding protein of *Escherichia coli*. Nature (London) **285**:78–81.
2. **Beguin, P.** 1983. Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. Anal. Biochem. **131**:333–336.
3. **Beguin, P., and H. Eisen.** 1978. Purification and partial characterization of three extracellular cellulases from *Cellulomonas* sp. Eur. J. Biochem. **87**:525–531.
4. **Beguin, P., H. Eisen, and A. Roupas.** 1977. Free and cellulose-bound cellulases in a *Cellulomonas* sp. J. Gen. Microbiol. **101**: 191–196.
5. **Birnboim, H. G., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
6. **Bolivar, F., R. L. Rodriguez, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
7. **Bradford, M. M.** 1976. A rapid and sensitive method for the

quantitation of microgram quantities of proteins utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
8. **Cornet, P., J. Millet, P. Beguin, and J. P. Aubert.** 1983. Characterization of two *cel* (cellulose degradation) genes of *Clostridium thermocellum* coding for endoglucanases. Bio/-Technology **1**:589–594.
9. **Emr, S. D., J. Hedgpeth, J. M. Clement, T. J. Silhavy, and M. Hofnung.** 1980. Sequence analysis of mutations that prevent export of λ receptor, an *Escherichia coli* outer membrane protein. Nature (London) **285**:82–85.
10. **Fukumori, F., T. Kudo, and K. Horikoshi.** 1987. Truncation analysis of an alkaline cellulase from an alkalophilic *Bacillus* species. FEMS Microbiol. Lett. **40**:311–314.
11. **Gilkes, N. R., D. G. Kilburn, M. L. Langsford, R. C. Miller, Jr., W. Wakarchuk, R. A. J. Warren, D. J. Whittle, and W. R. Wong.** 1984. Isolation and characterization of *Escherichia coli* clones expressing cellulase genes from *Cellulomonas fimi*. J. Gen. Microbiol. **130**:1377–1384.
12. **Gilkes, N. R., M. L. Langsford, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1984. Mode of action and substrate specificities of cellulases from cloned bacteria genes. J. Biol. Chem. **259**:10455–10459.
13. **Gold, L., D. Pribnow, T. Schneider, S. Schindedling, B. Swebilius-Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
14. **Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr.** 1987. Regulation, size and startsite determination of the *cenA* and *cex* transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:646–653.
15. **Guo, L.-H., and R. Wu.** 1983. Exonuclease III: use for DNA sequence analysis and in specific deletions of nucleotides. Methods Enzymol. **100**:60–97.
16. **Halliwell, G., and M. Griffin.** 1978. Affinity chromatography of the cellulase system of *Trichoderma koningii*. Biochem. J. **169**: 713–715.
17. **Hattori, M., and Y. Sakaki.** 1986. Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. **152**: 232–238.
18. **Inouye, S., X. Soberon, T. Franceschini, K. Nakamura, K. Itakura, and M. Inouye.** 1982. Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. Proc. Natl. Acad. Sci. USA **79**:3438–3441.
19. **Jensen, K. F., J. N. Larsen, L. Schack, and A. Sivertsen.** 1984. Studies on the structure and expression of *Escherichia coli pyrC*, *pyrD* and *pyrF* using the cloned genes. Eur. J. Biochem. **140**:343–352.
20. **Johnson, B. F.** 1977. Fine structure mapping and properties of mutations suppressing the *lon* mutation in *Escherichia coli* K-12 and B strains. Genet. Res. **30**:273–286.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
22. **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.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **McLaughlin, J. R., C. L. Murray, and J. C. Rabinotwitz.** 1981. Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* β-lactamase gene. J. Biol. Chem. **254**:11283–11291.
25. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
26. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Norrander, J., T. Kempe, and J. Messing.** 1983. Improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**: 101–106.
28. **Nossal, N. G., and L. A. Heppel.** 1966. The release of enzymes by osmotic shock from *Escherichia coli* in exponential phase. J. Biol. Chem. **241**:3055–3062.
29. **Nummi, M., M.-L. Niku-Paavola, T.-M. Enari, and V. Raunio.** 1980. Isolation of cellulases by means of biospecific sorption on amorphous cellulose. Anal. Biochem. **116**:137–141.
30. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Singler.** 1972. Novel method for detection of β-lactamase by using a chromogenic cephalosporin substrate. Antimicrob. Agents Chemother. **1**:283–288.
31. **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.
32. **O'Neill, G. P., D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1986. Overproduction from a cellulase gene with a high guanosine-plus-cytosine content in *Escherichia coli*. Appl. Environ. Microbiol. **52**:737–743.
33. **Peacock, S. L., C. M. McIver, and J. J. Monohan.** 1981. Transformation of *E. coli* using homopolymer-linked plasmid chimeras. Biochim. Biophys. Acta **655**:243–250.
34. **Roberts, T. M., and G. D. Lauer.** 1979. Maximizing gene expression on a plasmid using recombination *in vitro*. Methods Enzymol. **68**:473–482.
35. **Robson, L. M., and G. H. Chambliss.** 1986. Cloning of the *Bacillus subtilis* DLG β-1,4-glucanase gene and its expression in *Escherichia coli* and *B. subtilis*. J. Bacteriol. **165**:612–619.
36. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
37. **Russel, D. R., and G. N. Bennett.** 1982. Construction and analysis of *in vitro* activity of *E. coli* promoter hybrids and promoter mutants that alter the −35 to −10 spacing. Gene **20**: 231–243.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Stüber, D., and H. Bujard.** 1981. Organization of transcriptional signals in plasmids pBR322 and pACYC184. Proc. Natl. Acad. Sci. USA **78**:167–171.
40. **Sutcliffe, J. G.** 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43**:77–90.
41. **Teather, R. M., and P. J. Wood.** 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from bovine rumen. Appl. Environ. Microbiol. **43**:777–780.
42. **Vasantha, N., L. D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula.** 1984. Genes for alkaline protease and neutral protease from *Bacillus amyloliquefaciens* contain a large open reading frame between the regions coding for signal sequence and mature protein. J. Bacteriol. **159**:811–819.
43. **Vlasuk, G. P., S. Inouye, H. Ito, K. Itakura, and M. Inouye.** 1983. Effects of the complete removal of basic amino acid residues from the signal peptide on secretion of lipoprotein in *Escherichia coli*. J. Biol. Chem. **258**:7141–7148.
44. **Warren, R. A. J., C. F. Beck, N. R. Gilkes, D. G. Kilburn, M. L. Langsford, R. C. Miller, Jr., G. P. O'Neill, M. Scheufens, and W. K. R. Wong.** 1986. Sequence-conservation and region shuffling in an endoglucanase and an exoglucanase from *Cellulomonas fimi*. Proteins **1**:335–341.
45. **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**:139–154.
46. **Wong, W. K. R., B. Gerhard, Z. M. Gou, 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.
47. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.