JOURNAL OF BACTERIOLOGY, June 1992, p. 3522–3531
0021-9193/92/113522-10$02.00/0

Vol. 174, No. 11

celA from *Bacillus lautus* PL236 Encodes a Novel Cellulose-Binding Endo-β-1,4-Glucanase

CHRISTIAN K. HANSEN,[1]†* BØRGE DIDERICHSEN,[2] AND PER L. JØRGENSEN[3]‡

Unité de Biochimie Microbienne, Institut Pasteur, 75724 Paris Cedex 15, France,[1] and Novo Nordisk A/S, Novo Allé 1, DK-2880 Bagsværd,[2] and Department of Microbiology, The Technical University of Denmark, DK-2800 Lyngby,[3] Denmark

Received 31 December 1991/Accepted 1 April 1992

***celA* from the cellulolytic bacterium *Bacillus lautus* PL236 encodes EG-A, an endo-β-1,4-glucanase. An open reading frame of 2,100 bp preceded by a ribosome-binding site encodes a protein with a molecular mass of 76,863 Da with a typical signal sequence. The NH_2-terminal active domain of EG-A is not homologous to any reported cellulase or xylanase and may represent a new family of such enzymes. A 150-amino-acid COOH-terminal peptide is homologous to noncatalytic domains in several other cellulases (A. Meinke, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren, J. Bacteriol. 173:7126–7135, 1991). Upstream of *celA*, a partial open reading frame encodes a 145-amino-acid peptide which also belongs to the family mentioned. Zymogram analysis of extracts from *Escherichia coli* and supernatants of *Bacillus subtilis* and *B. megaterium*, including protease-deficient mutants thereof, which express *celA*, revealed two active proteins, EG-A-L and EG-A-S, with M_rs of 74,000 and 57,000, respectively. The proportion of EG-A-L to EG-A-S depends on the extracellular proteolytic activity of the host organism, indicating that EG-A-S arises from posttranslational proteolytic modification of EG-A-L. Since EG-A-S has an NH_2 terminus corresponding to the predicted NH_2-terminal sequence of EG-A, processing appears to take place between the catalytic and noncatalytic domains described. EG-A-L and EG-A-S were purified to homogeneity and shown to have almost identical characteristics with respect to activity against soluble substrates and pH and temperature dependency. EG-A-L binds strongly to cellulose, in contrast to EG-A-S, and has higher activity against insoluble substrates than the latter. We conclude that the COOH-terminal 17,000-M_r peptide of EG-A-L constitutes a cellulose-binding domain.**

Bacteria are rarely able to degrade crystalline cellulose. Common to the notable exceptions *Clostridium thermocellum* (15), *Ruminococcus albus* (18), *Cellulomonas fimi* (25), and *Bacillus lautus* (20, 37) is their secretion of several endo-β-1,4-glucanases (EC 3.2.1.4) (EGs). Soluble forms of cellulose, such as carboxymethyl cellulose (CMC), are degraded by many species.

To understand better the mechanisms of cellulose degradation and ultimately reconstitute crystalline cellulolytic activity by combination of pure enzymes, it is desirable to clone and analyze genes that encode individual EGs. More than 15 genes from crystalline-cellulose-degrading bacteria and more than 30 genes from CMC-degrading bacteria have been cloned and sequenced. In general, there is little homology between these enzymes, according to conventional primary structure comparison.

However, by hydrophobic cluster analysis, these EGs, their equivalents from eukaryotic organisms, and some other enzymes involved in cellulose and hemicellulose degradation have been grouped into families of homology, in total including approximately 70 members (12, 16). It is remarkable that the product of *celA* from *B. lautus* PL236 (20), the subject of this report, does not belong to any of these families and therefore appears to represent a new family of EGs.

* Corresponding author.

† Present address: c/o P. L. Jørgensen, Novo Nordisk A/S, Novo Allé 1, DK-2880 Bagsværd, Denmark.

‡ Present address: Novo Nordisk A/S, Novo Allé 1, DK-2880 Bagsværd, Denmark.

MATERIALS AND METHODS

Bacterial strains and plasmids. *B. lautus* PL236 was described previously (20). *Escherichia coli* K-12 strain MC1000 (6); *B. subtilis* 168 strains DN1885, DN969 (8), and BG2036 (53); and *B. megaterium* PV291 (50) were used as host organisms. Plasmids are listed in Table 1.

Culture media and growth conditions. Generally, LB medium (28) solidified with 1.5% agar was used for plates and TY medium (8) was used as liquid medium. Selective antibiotic media contained either ampicillin (100 μg/ml), tetracycline (10 μg/ml), chloramphenicol (10 μg/ml), or kanamycin (20 μg/ml). EG^+ colonies were detected on plates containing LB agar medium supplemented with 0.2% CMC which were subsequently stained with Congo red (49). Liquid cultures were grown overnight with vigorous aeration at 37°C.

For purification of EG-A-L, *B. subtilis* DN969(pCH7) was grown for 36 h at 37°C in 2-liter flasks containing LB medium with 10 μg of chloramphenicol per ml and 1 mM $CuSO_4$.

For purification of EG-A-S, *B. subtilis* DN1885(pCH57) was grown at 37°C for 166 h in a 2-liter jar fermentor containing 1,500 ml of H medium with 10 μg of kanamycin per ml. The fermentation was carried out with stirring (1,100 rpm) and aeration (1.1 liters/min), while the pH was maintained between 6.2 and 7.2 by addition of NH_3 or H_3PO_4. Dosing of a solution of 60% glucose and 0.06% citric acid at 3.7 ml/h was initiated after 40 h of incubation. H medium contains, per liter, 45 g of potato starch pretreated with 0.1 g of Termamyl (an α-amylase; Novo Nordisk A/S), 100 g of soybean meal, 15 g of corn steep liquor, 25 g of Alburex (potato protein), 2 g of $(NH_4)_2SO_4$, 1.1 g of KH_2PO_4, and 5.3

TABLE 1. Plasmids[a]

Plasmid (reference)	Size (kb)	Allele[b]	Marker	Origin of replication (reference)
pPL517 (20)	7.1	*celA*	Tc^r	pBR322 (3)
pPL217 (TW)[c]	7.1	*celA*	Tc^r	pBR322
pPL212 (TW)	6.6	*celA604/705*	Tc^r	pBR322
pPL216 (TW)	6.6	*celA602/626*	Tc^r	pBR322
pDN2801 (8)	2.8		Cm^r	pUB110 (22)
pCH7 (TW)	5.7	*celA*	Cm^r	pUB110
pPL1759 (21)	3.5		Km^r	pUB110
pCH52 (TW)	6.2	*celA*	Km^r	pUB110
pCH54 (TW)	6.2	*celA*	Km^r	pUB110
pCH57 (TW)	5.6	*celA1/31*	Km^r	pUB110

[a] See Fig. 3 and 4 for plasmid construction details.
[b] *celA* is the wild-type allele. Numbering of other alleles refers to the portion of the substituted gene product which is not encoded by *celA* sequences. For example, *celA604/705* refers to a truncated gene that encodes the first 603 amino acids of EG-A, while amino acids 604 to 705 are encoded by foreign sequences.
[c] TW, this work.

g of Na_2PO_4, all of standard commercial grade. The final culture volume was 1,200 ml.

Transformation. Established procedures for transformation of *E. coli* MC1000 (28), preparation of competent cells of *B. subtilis* DN1885 and BG2036 (9), and transformation of these (10) were used. Protoplasts of *B. subtilis* DN969 and *B. megaterium* PV291 were prepared, transformed, and regenerated as described in references 7 and 51, respectively.

DNA manipulation. Preparation of plasmid DNA was performed by the alkaline lysis procedure (2). Restriction enzyme digestions and ligation procedures were carried out as specified by the enzyme supplier (New England BioLabs). Oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer and manipulated as previously described (28). Agarose gel electrophoresis was used to analyze plasmids, fragments from restriction enzyme digestions, and ligation reactions. Agarose gels were made up and run in borate buffer (28) containing ethidium bromide (1 μg/ml). DNA was detected by illumination with shortwave UV light.

DNA sequencing. DNA sequencing of both strands of the three *Pst*I fragments contained in plasmid pPL517 carrying *celA* was performed by the chemical modification procedure (30) with single-end ^{32}P-labeled DNA fragments. Fragments were created by using convenient restriction sites. The cleavage products were separated on 8 or 20% polyacrylamide gels subjected to autoradiography at −70°C with intensifying screens. The sequence was confirmed on one strand with the method of Sanger et al. (44), by using synthetic oligodeoxynucleotides complementary to the previously determined sequence.

Electrophoretic analysis of expressed EG. Extracts of *E. coli* and *B. subtilis* were made by sonication of 10-fold-concentrated late-logarithmic- and stationary-phase cultures suspended in 100 mM Tris-HCl (pH 7). Protoplasts of *B. subtilis* were prepared by addition of 2 mg of lysozyme per ml to an overnight culture with subsequent incubation at 37°C for 30 min. Protoplasts were harvested at 5,000 × *g* for 5 min, washed twice, suspended in 0.1 volume of 20 mM potassium phosphate (pH 7), and sonicated.

These extracts and supernatants of *B. subtilis* cultures were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (23). Zymogram analysis was performed by denaturing proteins in sample buffer at 70°C for 10 min prior to electrophoresis. Renaturation of proteins was carried out by washing the gels in 100 mM potassium phosphate (pH 7) three times for 30 min each time at 4°C. EG activity in protein bands was detected by overlaying polyacrylamide gels with 0.8-mm-thick 2% agarose gels containing 0.2% CMC and 100 mM phosphate buffer (pH 7), incubating them at 42°C for 4 h, staining the substrate gel with 0.1% Congo red for 15 min, and destaining it with 1 M NaCl for 15 min. Active bands appear as yellow halos on a red background. For greater sensitivity, in some experiments 0.2% CMC was included directly in the polyacrylamide gel. Proteins were denatured and renatured by using the conditions described above, although activity bands usually could be visualized by Congo red staining after only 1 to 2 h of incubation at 42°C.

Pure proteins and proteolytic degradation experiments were monitored by denaturing SDS-PAGE (23), followed by staining with Coomassie brilliant blue R and destaining with a solution of 10% ethanol and 7% acetic acid.

Protein purifications. EG-A-L was purified from the culture of *B. subtilis* DN969(pCH7) in LB medium. The total culture was centrifuged for 10 min at 10,000 × *g* at 4°C, and EDTA was added to the supernatant to a final concentration of 5 mM. Avicel PH-101 (Sigma) was added to a concentration of 2.5 g/liter, and the suspension was incubated with shaking for 2 h at 4°C and centrifuged for 10 min at 10,000 × *g*. The pellet was washed twice with 1 mM EDTA and resuspended in a solution containing 1% triethylamine and 1 mM EDTA and stirred for 1 h at 4°C. Avicel was removed by centrifugation, and the supernatant was neutralized by addition of HCl. Protein was subsequently precipitated with $(NH_4)_2SO_4$ at 70% saturation for 16 h and harvested by centrifugation for 15 min at 30,000 × *g*. The protein pellet was suspended in 20 mM Tris HCl (pH 8) and applied to a Mono Q column (Pharmacia) previously equilibrated with the same buffer and was eluted with a 30-ml 0 to 500 mM NaCl gradient in 20 mM Tris HCl by using a fast protein liquid chromatography unit from Pharmacia. Active fractions were pooled and used for further analysis.

EG-A-S was purified from the culture of *B. subtilis* DN1885(pCH57). The culture was centrifuged at 4°C at 10,000 × *g* for 30 min and further cleared by filtration through a Whatman GF/D filter. Threefold water-diluted supernatant was subsequently concentrated in an Amicon ultrafiltration unit equipped with a PM 10 membrane (Millipore). This process was repeated twice. The solution was brought to a volume of three times the original supernatant volume, and Tris HCl (pH 7) was added to a final concentration of 50 mM. Anion exchange was performed on a column (5 by 15 cm) containing DEAE-Sephacryl (Pharmacia), and protein was eluted with a 0 to 500 mM NaCl gradient in 50 mM Tris HCl (pH 7). Active fractions were pooled, and chromatography was repeated under the same conditions.

Enzyme assays. Assays for activity against CMC (high viscosity; Sigma), Avicel PH-101 (Sigma), 3MM filter paper (Whatman), phosphoric acid-swollen cellulose (20), xylan (larchwood and oatspelt; Sigma), lichenan (from *Cetraria islandica*; Sigma), and laminarin (from *Laminaria digitata*; Sigma) were performed by incubation of various amounts of enzyme solution with 0.5% solutions of the polysaccharides in 50 mM potassium phosphate for various times. For insoluble substrates such as Avicel, filter paper, and ASC, incubation was performed with shaking and residual substrate was removed by centrifugation at 20,000 × *g* for 5 min prior to reducing sugar determination. Liberated reducing

sugar was quantified as glucose by the dinitrosalicylic acid procedure (34).

Assays for hydrolytic activities against 4-nitrophenyl-β-D-glucopyranoside (Sigma) and 4-nitrophenyl-β-D-cellobioside (Sigma) were done with 1 mM substrate and various concentrations of enzyme in 50 mM potassium phosphate (pH 7) for up to 8 h. Reactions were stopped by addition of 0.4 reaction volume of 1 M Na_2CO_3, and the liberated 4-nitrophenol was quantified at 410 nm by using 4-nitrophenol as the standard.

One unit of enzyme activity was defined as the amount of protein that produces 1 μmol of product per min. All assays were performed at 37°C and pH 7, except pH and temperature optimum experiments.

Plate assays were used to monitor purifications and were performed by incubation of enzyme samples in wells in plates containing 2% agarose, 100 mM potassium phosphate (pH 7), and 0.2% CMC. Plates were incubated for 2 to 16 h at 37°C, and clearing zones on CMC-containing plates were visualized with Congo red as previously described (20).

Assay for cellulose-binding capacity. Samples (2 μg) of pure EG-A-L and EG-A-S were incubated with shaking at 37°C for 30 min in 50-μl solutions containing 2 mg of Avicel and 50 mM potassium phosphate (pH 7). Avicel and adsorbed protein were removed by centrifugation at 5,000 × *g* for 2 min, and the supernatant was removed. The pellet was subsequently suspended in 50 μl of a solution containing 1% triethylamine and 1 mM EDTA and incubated with shaking for 30 min. Avicel was again removed by centrifugation at 5,000 × *g* for 2 min. Supernatants of the adsorption and desorption experiments were analyzed by SDS-PAGE followed by staining with Coomassie brilliant blue R.

Proteolytic degradation of EG-A-L. Reactions (50 μl) containing 2 μg of pure EG-A-L in 50 mM potassium phosphate (pH 7) were incubated with 0.2 to 1 μg of *B. subtilis* DN1885 supernatant protein or with 0.01 to 0.2 μg of proteinase K and trypsin for various times at 37°C. Reaction products were subsequently analyzed by SDS-PAGE and Coomassie brilliant blue R staining.

Determination of total protein. Total protein was determined as described by Bradford (4), with a Bio-Rad protein assay kit using bovine serum albumin as the standard.

NH_2-terminal determination. Automatic Edman degradation with an Applied Biosystems 470A sequencer was used for determination of NH_2-terminal amino acids (29).

Nucleotide sequence accession number. The nucleotide and deduced amino acid sequences reported here have been submitted to GenBank (accession number M76588).

RESULTS

Sequence and activity analysis. (i) Nucleotide sequence of *celA*. The entire nucleotide sequence (Fig. 1) of the 2.8-kb insert originating from *B. lautus* PL236 on pPL517 was determined. An open reading frame from nucleotides 531 to 2630 is preceded by a putative ribosome-binding site (31). It encodes a protein of 700 amino acids with a calculated molecular mass of 76,863 Da. The first 33 amino acids resemble a typical signal peptide of a gram-positive bacterium (1, 36, 40). No sequences upstream of the proposed ribosome-binding site have discernible homology to any reported *Bacillus* promoter sequence (5, 14, 35, 41), and no inverted repeat sequence was observed. Likewise, no significant inverted repeat sequence which could be proposed as a ρ-independent terminator (43) is present downstream of the coding region.

Upstream of *celA*, from nucleotides 1 to 435, an open reading frame encodes 145 amino acids. We speculate that it constitutes the COOH terminus of a protein encoded by a gene upstream of *celA*.

(ii) GC content and predominant amino acids. The GC content of the *B. lautus* genomic DNA cloned in pPL517 is 53 mol%, while the fragment containing *celB* (20) has a GC content of 49 mol%. These values correspond well to the total GC content of the *B. lautus* genome of 50 to 52 mol% reported by Nakamura (37).

EG-A and EG-B from *B. lautus* have hydroxyamino acid contents of 22 and 21%, respectively. Such high contents of hydroxyamino acids are characteristic of many glycosidases (32).

(iii) Homology of EG-A and other proteins. The deduced amino acid sequence of EG-A was compared to the cellulase and xylanase activity domain families proposed by Henrissat et al. (16) and Gilkes et al. (12). Interestingly, it was not possible to demonstrate sequence similarity to the members of any of these families by primary structure comparison or hydrophobic cluster analysis (11). The latter method compares predicted secondary protein structures and has been used successfully to demonstrate low degrees of homology between cellulases (16).

The COOH-terminal 150-amino-acid sequence of EG-A is highly homologous to (i) the COOH-terminal peptides of the three reported *B. subtilis* EGs (26, 36, 42), (ii) the noncatalytic central domain of the *Caldocellum saccharolyticum celB* product (45), (iii) a domain in the *C. fimi cenB* gene product (32), (iv) two domains (C and C′) in *Clostridium stercorarium* avicelase I (19), (v) a domain in CelF of *C. thermocellum* (38), and (vi) the 145-amino-acid peptide encoded by the partial open reading frame upstream of *celA* from *B. lautus* (Fig. 2). Ten amino acids are conserved in all seven sequences and are probably essential for the structure and/or function of the domain (33).

A high degree of homology (84%) was observed between the peptide encoded by the sequence upstream of *celA* and the COOH-terminal peptide of EG-A (Table 2). These peptides are moderately (40 to 48%) homologous to the *B. subtilis* and *C. saccharolyticum* sequences and to the C domain of *C. stercorarium* avicelase I. Finally, the *C. fimi* and *C. thermocellum* sequences and the C′ domain from *C. stercorarium* avicelase I are internally well related (28 to 46%) although only distantly related to the other five sequences (12 to 34%).

Cloning and expression of *celA* in different hosts. (i) Construction of deletion clones in *E. coli*. pPL517 was described earlier (20) as arising from ligation of partially *Pst*I-digested genomic DNA with *Pst*I-digested pBR322 (3). In the same cloning experiment, three other plasmids that confer CM Case activity on the host organism and contain *celA* or parts of it were isolated and extracts from these cells were submitted to zymogram analysis (Fig. 3).

Restriction enzyme analysis showed that pPL217 and pPL517 both contained the three consecutive *Pst*I fragments, but in opposite orientation with respect to the pBR322 (3) vector. Earlier it was shown that these three *Pst*I fragments are consecutive on the *B. lautus* PL236 chromosome (20). Plasmids pPL212 and pPL216 contain the two 5′-proximal *Pst*I fragments in opposite orientation relative to the vector fragment. These two plasmids contain 3′-truncated forms of *celA* which encode COOH-terminally modified versions of EG-A.

(ii) Expression of deletion clones in *E. coli*. EG activity in MC1000(pPL517) was 40 times stronger than in MC1000 (pPL217), although both contain the complete *celA* gene

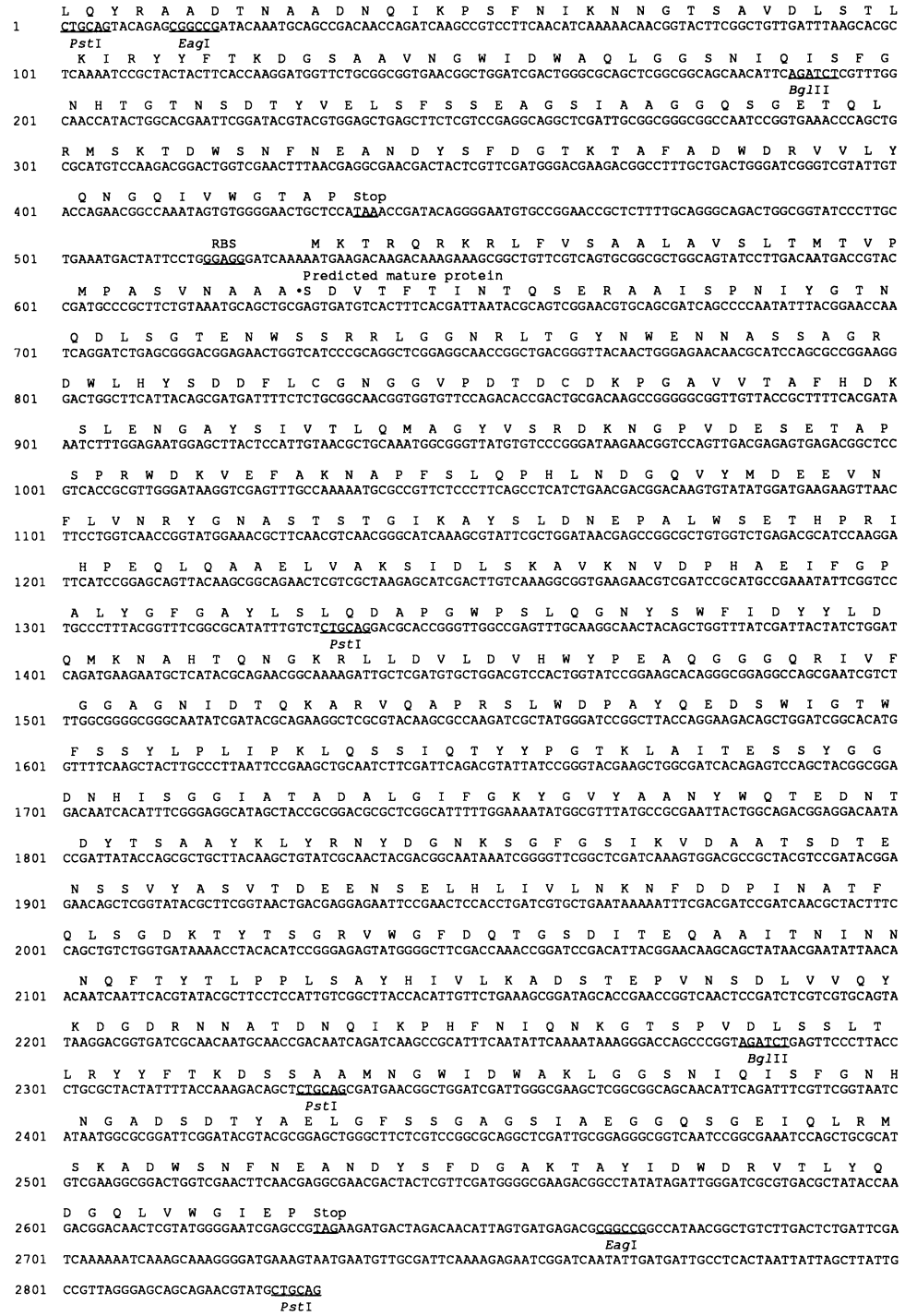


FIG. 1. Nucleotide sequence of the three *Pst*I fragments carrying the *celA* gene and an upstream open reading frame. The proposed ribosome-binding site (RBS), the predicted NH_2 terminus of EG-A (black dot), and relevant restriction sites are shown.

(Fig. 3). This indicates that transcription takes place primarily from a vector-encoded promoter sequence, for example, the β-lactamase promoter of pBR322 which is present 700 bp upstream of the *Pst*I site, and is in agreement with the lack of anything recognizable as a promoter sequence in the cloned fragment.

EG activities in MC1000(pPL517) and MC1000(pPL212) are almost the same, although pPL212 encodes only the NH_2-terminal 603 of the 700 amino acids of EG-A. Apparently, activity of EG-A against CMC is not dependent on the presence of the extreme COOH-terminal amino acids.

Zymogram analysis of extracts from MC1000(pPL517), which expresses the complete *celA* gene, showed EG activity bands with M_rs of 74,000 and 57,000 (EG-A-L and EG-A-S, respectively). MC1000(pPL212), MC1000(pPL216), and MC1000(pPL217) contain 3′-differing versions of *celA*

```
B.l.EGA-CBD   PVNSD-LVVQYKDGDRNNATDWQIKPHFNIQWKGTSPVDLSSLTLRYYFTKDSSAAMWGW
B.l. UPSTREAM GGNHP-FRLQYRAAD-TNAADWQIKPSFNIKWNGTSAVDLSTLKIRYYFTKDGSAAVWGW
B.s.DLG EG    TQEKG-VSVQYKAGD-GRVNSWQIRPQLHIKWNGNATVDLKDVTARYWYNVK-NKGQWFD
C.s.EGB-CD    ATGGQ-IKVLYANKE-TNSTTWTIRPWLKVVWSGSSSIDLSRVTIRYWYTVDGERAQSAV
C.st.C-DOM    ---GVIQIQWFWGNTSDKTNGIMPRYRLTNTGTTPIRLSDV--KIRYYYTIDGEKDQWFW
C.th.EGF-CBD  PGEEFYVEAAVWA--AG-PGFVNIKASIINKSGWPARGSDK-LSAKYFVDISEAVAKGIT
C.st.C'-DOM   TNDEFFVWAGIWA--SG-QNFIEIKALLHNQSGWPARVADK-LSFRYFVDLTELIEAGYS
C.f. EGB      DGDQLFVEAMLWQPPSG-T-FTEVKAMIRNQSAFPARSLKN-AKVRYW-----FTTDGFA

CONSENSUS           V  YNA  G   N IKP L  N G A VDL  L  RX  T D   A N


B.l.EGA-CBD   IDWAKLGGSWIQISFGNHNGA--D--SDTYAELGFSSGAGSIAEGGQSG---EIQLRMSK
B.l. UPSTREAM IDWAQLGGSWIQISFGNHTGT--N--SDTYVELSFSSEAGSIAAGGQSG---EIQLRMSK
B.s.DLG EG    CDYAQWGCGWLTHKFVTLHKP--KQGADTYLELGFKTGT--LSPGASTG---NIQLRLHN
C.s.EGB-CD    SDWAQIGASWVTFKFVKLSSS--VSGADTYLEIGFKSGAGQLQPGKDTG---EIQIRFNK
C.st.C-DOM    CDWSSVGSNWITGTFVKMAEP--KEGADTYLETGFDTGAGYLQP-NQSI---EVQNRFSK
C.th.EGF-CBD  LDQITVQSTTNGGAKVSQLLW-DPDNHIYYVNIDFTGIN--IFPGGINEYKRDVYFTITA
C.st.C'-DOM   ASDVTITTNYNAGAKVTGLHPWNEAENIYYVNVDFTGTK--IYPGGQSAYRKEVQFRIAA
C.f. EGB      ASDVTLSANYSECGAQSGKGV-SAGGTLGYVELSCVGQD--IHPGGQSQHRREIQFRLTG

CONSENSUS      DWA  G  N    FV        G DYXVELGF  GAG I PGGQSG   EIQ R  K


B.l.EGA-CBD   ADWSNF-WEAWDYSFDG----AKTAYIDWDR-VTLYQDGQLVWGIEP----
B.l. UPSTREAM TDWSNF-WEAWDYSFDG----TKTAFADWDR-VVLYQNGQMVWGTAP----
B.s.DLG EG    DDWSNY-AQSGDYSFFQ----SNT-FKTTKK-ITLYHQGKLIWGTEPN---
C.s.EGB-CD    SDWSNY-WQGWDWSWLQ----SMTSYGENEK-VTAYIDGVLVWGQEPSGAT
C.st.C-DOM    ADWTDY-IQTWDYSF-----STNTSYGSNDR-ITVYISGVLVSGIEP----
C.th.EGF-CBD  PYGEGNWDNTWDFSFQGLEQGFTSKKTE---YIPLYDGNVRVWGKVPDGGS
C.st.C'-DOM   PQNTWFWWNDWDYSFRDIK-GVTSGNTVKTVYIPVYDDGVLVFGVEP----
C.f. EGB      PAG---WWPAWDPSYTGLTQTALAKASA----ITLYDGSTLVWGKEP----

CONSENSUS      DWSN  N  WDYSF G      T        ITLX  GVLVWG EP
```

FIG. 2. Alignment of proposed cellulose-binding domains from several sources. *B. lautus* PL236 EG-A COOH-terminal domain (B.l.EGA-CBD), the domain encoded by the sequence upstream of *celA* (B.l. UPSTREAM), the COOH-terminal domain of an EG from *B. subtilis* DLG (B.s.DLG EG) (42), the central domain of EG-B from *C. saccharolyticum* (C.s.EGB-CD) (45), the C′ and C domains of *C. stercorarium* avicelase I (C.st.C′-DOM and C.st.C-DOM, respectively) (19), a domain of *C. thermocellum* CelF (C.th.EGF-CBD) (38), and a central domain of *C. fimi* CenB (C.f. EGB) (32) are aligned. The consensus sequence has been constructed by showing amino acids conserved in four sequences in normal letters, amino acids conserved in seven sequences in boldface letters, and residues conserved in all sequences in underlined boldface letters. Residues in the individual sequences identical to the consensus sequence are shown in boldface letters.

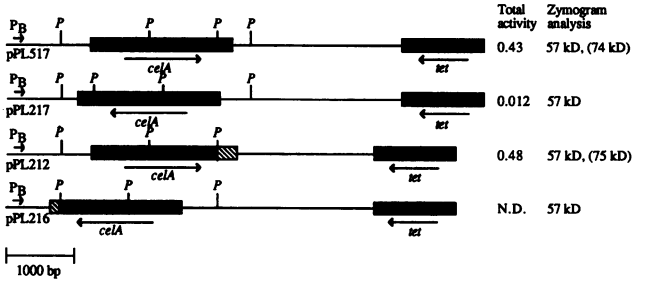


FIG. 3. Subcloning of *celA* in *E. coli*. The tetracycline resistance gene from pBR322 (*tet*) (3), the *celA* gene of *B. lautus* PL236 (black boxes), extended open reading frames originating from vector sequences (cross-hatched boxes), *Pst*I sites (*P*), and the β-lactamase promoter from pBR322 (P_B) are shown. Total activity refers to the units of CMCase per milliliter of culture extract in cells carrying the respective plasmids. One unit is defined as the amount of enzyme which liberates 1 μmol of reducing sugar from CMC per min. Zymogram analysis refers to the molecular masses of EGs observed in extracts of cells carrying the respective plasmids. The values in parentheses indicate less abundant species of EG. N.D., not done.

that encode COOH-terminally differing proteins with M_rs of 78,000, 69,000, and 77,000, respectively. However, extracts from these cells all contain active EG with an M_r of 57,000 (data not shown). This shows that EG-A-S does not arise from NH_2-terminal modification of a precursor, since this would result in different mature proteins.

(iii) Cloning and expression of the *celA* gene in *B. subtilis*. Plasmids pCH7 and pCH57, containing native *celA* and *celA* fused to the signal peptide-encoding region of an α-amylase-encoding gene (Fig. 4), respectively, gave rise to EG-A production in *B. subtilis*.

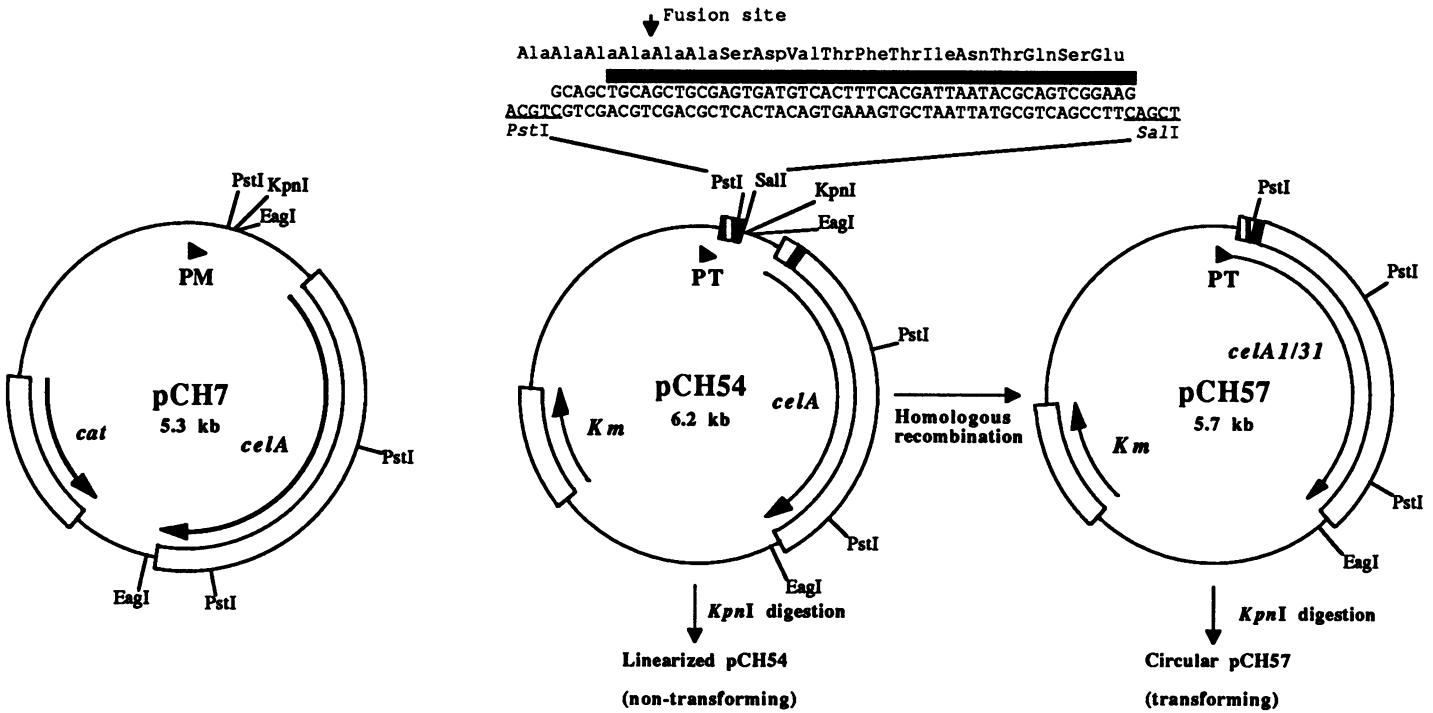
Supernatant from *B. subtilis* DN1885(pCH7) contained approximately 50 times more EG activity than that of *B. subtilis* DN1885 (Table 3), demonstrating that *B. subtilis* actively expresses *celA* and secretes EG-A. *B. subtilis* EGs did not interfere in the zymogram analysis, presumably owing to their very low levels of activity (Table 3). Zymogram analysis of supernatants from late-logarithmic- and stationary-phase cultures showed that EG-A-L and EG-A-S are present regardless of the growth phase. However, while stationary-phase supernatant contains only trace amounts of EG-A-L (<1/10 of the amount of EG-A-S), late-logarithmic-phase supernatant contains approximately equal amounts of the two species. Furthermore, zymogram analysis of protoplast extract showed EG activity at 77 and 74 kDa (Table 3). These proteins correspond to EG-A before and after signal peptide cleavage and indicate that EG-A-S arises from extracellular proteolytic degradation of EG-A-L and not from differential transcription of the *celA* gene.

(iv) *celA* expression in protease-deficient bacilli. pCH7 was expressed in *B. subtilis* BG2036 (Apr^- Npr^-) and DN969 (very weak protease phenotype) and in *B. megaterium* PV291 (Npr^-) to produce high levels of 74-kDa EG-A (Table 3). Although the EG-A-S–to–EG-A-L ratio (determined by zymogram analysis) was lower in these strains (Fig. 5), EG-A-S was observed in all cases. This indicates that the presumed proteolytic modification of EG-A-L does not have an absolute requirement for the major extracellular proteases of *B. subtilis* and *B. megaterium*. Thus, several minor proteases apart from Apr and Npr were reportedly produced

TABLE 2. Percentages of identical amino acids in aligned sequences[a]

Sequence	% of amino acids identical						
	B.l. UPSTREAM	B.s.DLG EG	C.s.EGB-CD	C.st.C′-DOM	C.st.C-DOM	C.f. EGB	C.th.EGF-CBD
B.l.EGA-CBD	84	44	47	22	46	34	14
B.l. UPSTREAM		40	42	23	40	30	17
B.s.DLG EG			48	19	42	27	14
C.s.EGB-CD				22	48	27	12
C.st.C′-DOM					22	42	46
C.st.C-DOM						24	14
C.f. EGB							28

[a] All values are given as percentages of identical amino acids of the total number of aligned amino acids. The sequence abbreviations are defined in the legend to Fig. 2.



(1) MKTRQRKRLFVSAALAVSLTMTVPMPASVNAAASDVTFTINTQSERAAISPNIYGTN

(2) MKQQKRLYARLLTLLFALIFLLPHSAAAAAASDVTFTINTQSERAAISPNIYGTN

(p1) AAASDVTFTINTQSERAAISPNIY

(p2) AASDVTFTINTQSERAAISPNIYG

FIG. 4. Plasmids pCH7, pCH54, and pCH57. *Eag*I-digested pDN2801 (8) was ligated to *Eag*I-digested pPL517 (20) to form pCH7. The promoter of the maltogenic α-amylase-encoding gene (PM) from *B. stearothermophilus* (8) and the chloramphenicol resistance gene (*cat*) from pC194 (17) are shown. pCH54 was constructed in two steps. First, a synthetic linker (top) was inserted into the *Pst*I and *Sal*I sites of pPL1759 (21). The resulting plasmid, pCH52 (Table 1), was then digested with *Eag*I and ligated to the *Eag*I fragment of pPL517 containing *celA*, resulting in pCH54. In vivo homologous recombination between the short direct repeats (black areas) resulted in pCH57 as previously described (21). The promoter (PT) and signal peptide-encoding region of a *B. licheniformis* α-amylase-encoding gene (cross-hatched region) (21) and the kanamycin resistance gene from pUB110 (22) are shown. Deduced NH_2-terminal amino acid sequences of *celA*-encoded EG-A (no. 1) and the hybrid protein encoded by pCH57 (no. 2), and the determined NH_2 termini of purified EG-A-S (p1 and p2) are shown below. Sequences p1 and p2 are present in approximately equal amounts. The predicted potential NH_2-terminal amino acids are marked with dots.

TABLE 3. *celA* expression in bacilli

Strain	Protease phenotype	Plasmid	Supernatant activity (U/ml of culture)	Zymogram analysis[a]	
				Supernatant	Extract
B. subtilis					
DN1885	WT[b]	None	0.011	(36)	ND[c]
DN1885	WT	pCH7	0.23	57 (74, 36)	(74, 77)
BG2036	Apr$^-$ Npr$^-$	pCH7	0.26	57 (74, 36)	ND
DN969	Very weak	pCH7	0.08	74 (57, 36)	ND
B. megaterium PV291	Npr$^-$	pCH7	0.20	57 (74, 36)	ND

[a] M_rs (10^3) of observed EGs are shown. The values in parentheses refer to minor activities.
[b] WT, wild type.
[c] ND, not determined.

by *B. subtilis* (46–48, 52), while no other proteases have been reported in *B. megaterium*.

Purification and characterization of EG-A-L and EG-A-S. (i) Purification of EG-A-L and EG-A-S. The various steps of purification of EG-A-L and EG-A-S are summarized in Table 4.

B. subtilis DN969(pCH7) was chosen for purification of EG-A-L, which is very sensitive to proteolytic attack, because of the low extracellular protease activity of this strain. Since only relatively low levels of EG-A-L-production are obtained in the selected medium, purification of this enzyme was greatly facilitated by the affinity of EG-A-L for Avicel. Avicel-adsorbed protein was liberated with 1% triethylamine, and EG-A-L was further purified by anion exchange on a Mono Q HR column. SDS-PAGE revealed a single band with an M_r of 74,000 (Fig. 6).

EG-A-S was purified from the culture supernatant of *B. subtilis* DN1885(pCH57) fermented for 166 h in an industrial medium. Very high levels of extracellular protein and specific activity were obtained (Table 4), facilitating purification of EG-A-S which does not bind to cellulose. Approximately 10% of the total extracellular protein was constituted by EG-A-S, and pure protein was obtained after dialysis and anion exchange on a DEAE Sephacryl column (repeated twice). SDS-PAGE showed a single band with an M_r of 57,000 (Fig. 6).

(ii) NH_2-terminal sequences of EG-A-S. NH_2-terminal sequencing (Fig. 4) showed that EG-A-S has an NH_2-terminal sequence that closely resembles the predicted sequence for native EG-A (Fig. 1), demonstrating that it lacks approximately 150 amino acids at the COOH terminus.

(iii) Substrate specificities. The purified preparations of EG-A-L and EG-A-S were examined for the ability to hydrolyze various carbohydrates and aryl glycosides at 37°C and pH 7 in 100 mM potassium phosphate buffer (Table 5). Of the carbohydrate substrates tested, the soluble substrates lichenan (β-1,3;1,4 linkage) and CMC (β-1,4 linkage) were efficiently hydrolyzed by both EG-A-L and EG-A-S, the enzymes having practically identical specific activities. However, the insoluble substrates phosphoric acid-swollen cellulose and Avicel were hydrolyzed more efficiently by EG-A-L, although the specific activity against these substrates was significantly lower than that against lichenan and CMC. Laminarin (β-1,3;1,6 linkage), xylan, 4-nitrophenyl-β-D-cellobioside, and 4-nitrophenyl-β-D-glucopyranoside were not hydrolyzed by EG-A-L and EG-A-S.

(iv) pH optima. The effect of pH on the activities of EG-A-L and EG-A-S against CMC was determined at 37°C in various buffers ranging from pH 3.5 to pH 11 (Fig. 7). The enzymes had practically identical activities at the pH values tested. Maximal activity was observed over a broad range of pH values from 5 to 8.5, and 50% activity was observed at pH 9.5.

(v) Temperature optima and stability. The temperature dependence of EG-A-L and EG-A-S activities toward CMC was determined by measuring activity at various temperatures in potassium phosphate buffer (pH 7). As with pH dependence, EG-A-L and EG-A-S were identically dependent on temperature (Fig. 7). Maximal activity in the 60-min assay was observed at 60°C. More than 50% of the maximal activity was observed from 46 to 68°C.

EG-A-L and EG-A-S were incubated for 0 to 5 h in 50 mM potassium phosphate buffer (pH 7) at 45 to 65°C. Determination of the residual activity against CMC showed that both enzymes were stable for up to 5 h at 55°C, and a half-life of 70 min was determined at 60°C.

(vi) Cellulose-binding ability. EG-A-L and EG-A-S were incubated with microcrystalline Avicel cellulose, and supernatants of these mixtures were analyzed by SDS-PAGE. Only EG-A-S was observed in the supernatant, while EG-A-L remained adsorbed to Avicel. The cellulose pellet was subsequently treated with 1% triethylamine, which liberated

TABLE 4. Purification of EG-A-L and EG-A-S

Protein	Purification step	Vol (ml)	Total protein (mg)	Total activity[a] (U)	Sp act[a] (U/mg)	Yield (%)	Purification (fold)
EG-A-L	Cell-free culture	1,000	370	17.1	0.046	100	1
	Liberated from Avicel	20	4	6.2	1.54	36	33
	Pool from Mono Q	1.4	2.3	4.9	2.2	29	47
EG-A-S	Cell-free culture	400	4,800	1,280	0.264	100	1
	PM 10 membrane retentate	1,050	2,100	715	0.344	56	1.3
	Pool from DEAE column	119	226	543	2.4	42	9.1

[a] Activity against CMC.

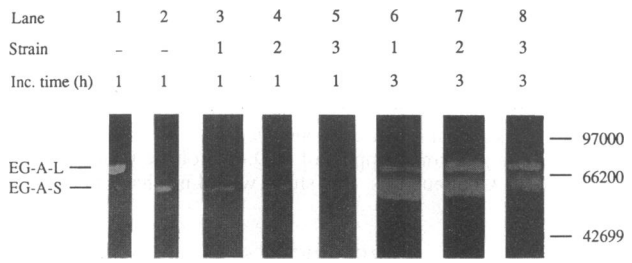


FIG. 5. Zymogram analysis of early-stationary-phase supernatants of *B. subtilis* DN1885(pCH7) (1) (lanes 3 and 6), BG2036(pCH7) (2) (lanes 4 and 7), and DN969(pCH7) (3) (lanes 5 and 8), which express *celA*. Zymogram analysis of pure EG-A-L (lane 1) and EG-A-S (lane 2) is also shown. After electrophoresis and washing, gels with lanes 1 to 5 were incubated for 1 h at 37°C while gels with lanes 6 to 8 were incubated (Inc.) for 3 h at 37°C. Protein (5 μg) was applied to lanes 3 to 8, while lanes 1 and 2 were loaded with 0.5 μg of protein. The numbers on the right are molecular weights.

bound EG-A-L from Avicel and demonstrated that EG-A-S was not adsorbed on the crystalline matrix.

(vii) Proteolysis of EG-A-L. To substantiate the hypothesis that EG-A-S arises from posttranslational proteolysis of EG-A-L, purified EG-A-L was incubated with various amounts of *B. subtilis* DN1885 stationary-phase supernatant, proteinase K (Fig. 8), and trypsin (data not shown) for 30 min to 16 h at 37°C. *B. subtilis* supernatant effected proteolysis of EG-A-L, yielding initial products with M_rs of 59,000 and 57,000, while prolonged incubation gave rise to an apparently pure and stable protein with an M_r of 57,000 corresponding to EG-A-S. Incubation with trypsin also resulted in a stable and pure 57,000-M_r protein (data not shown). Treatment with proteinase K resulted in initial production of a 57,000-M_r protein, but prolonged incubation gave total proteolysis (Fig. 8). Zymogram analysis demonstrated that all of the 57,000-M_r proteins were active against CMC (data not shown). Furthermore, EG-A-L was stable for several days at 37°C in the absence of proteases.

DISCUSSION

Nucleotide sequence analysis of *celA*, which encodes EG-A of *B. lautus* PL236, and its flanking regions revealed the following. (i) An open reading frame of 2,100 bp encodes a 700-amino-acid protein whose first 33 NH_2-terminal amino acids resemble a gram-positive signal peptide and is preceded by a possible ribosome-binding site. (ii) No obvious transcription-promoting or -terminating sequences are present in the flanking regions. (iii) An open reading frame of 435 bp, presumably constituting the 3′ end of a gene located upstream of *celA*, encodes a peptide with significant homology to the COOH terminus of EG-A.

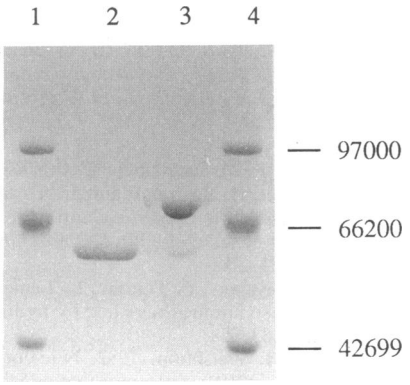


FIG. 6. Purified EG-A-S and EG-A-L (2 μg of each) were applied to lanes 2 and 3, respectively. Lanes 1 and 4 contained Bio-Rad molecular weight markers with the M_rs indicated.

TABLE 5. Substrate specificities of EG-A-L and EG-A-S

Substrate[a]	Sp act (U/mg)	
	EG-A-L	EG-A-S
CMC	2.2	2.4
ASC[b]	0.54	0.09
Avicel	0.028	0
Lichenan	7.3	6.9

[a] All substrates were used at 0.5%.
[b] ASC, phosphoric acid-swollen cellulose.

In combination, these observations suggest that *celA* is part of an operon which possibly contains at least one other gene that encodes a protein involved in cellulolysis.

Hydrophobic cluster analysis and conventional protein and DNA homology searches showed that EG-A-S (containing the active domain) is not homologous to any other protein. EG-A-S is therefore proposed to be the first member of a new family of cellulases and hemicellulases.

EG-A-L includes a small COOH-terminal cellulose-binding domain which is indeed homologous with a group of similar domains from several other bacterial cellulases (33), constituting a second family of bacterial cellulose-binding domains.

Analysis of the translational products of *celA* gave the following results. (i) Two proteins with activity against CMC, EG-A-L and EG-A-S, were observed on zymograms. (ii) Both forms, in various proportions, were seen when *celA* was expressed in *E. coli* and *B. subtilis*. (iii) Both forms were seen when protease-deficient mutants of *B. subtilis* and *B. megaterium* were used. (iv) Both forms were seen when the

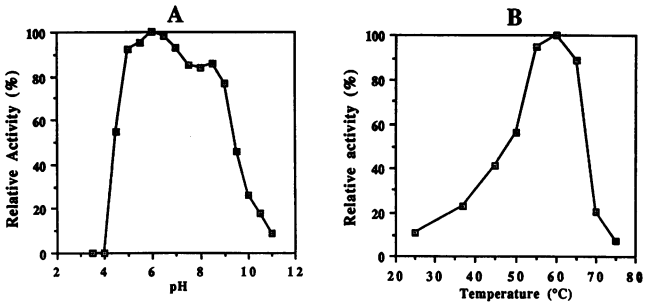


FIG. 7. pH (A) and temperature (B) optima of EG-A-L and EG-A-S. Only one curve is shown for each parameter, since the two enzymes had practically coinciding curves. Relative activity as a function of pH was determined by incubation of 1 μg of enzyme in 1-ml reactions at 37°C for 1 h with 4% low-viscosity CMC and a 100 mM concentration of one of the following buffers: acetic acid-sodium acetate (pH 3.5 to 5.5), potassium phosphate (pH 6 to 8), Tris HCl (pH 8.5 to 9.5), or glycine-sodium hydroxide (pH 10 to 11). Relative activity as a function of temperature was determined by incubation of 1 μg of enzyme in 1-ml reactions containing 1% high-viscosity CMC for 1 h in 100 mM potassium phosphate (pH 7) at temperatures from 25 to 80°C. Activity was quantified by determination of reducing sugars.

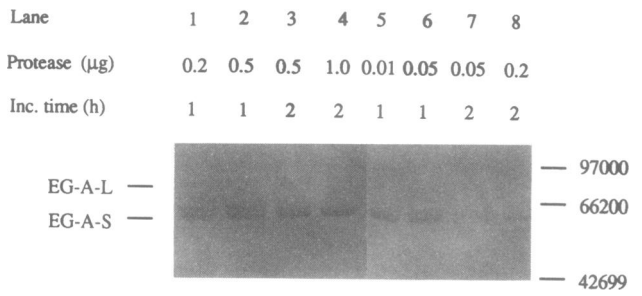


FIG. 8. Proteolysis of EG-A-L. EG-A-L (2 μg) was incubated (Inc.) with various amounts of *B. subtilis* stationary-phase supernatant protein (lanes 1 to 4) and proteinase K (lanes 5 to 8) for various times. To the left is marked the migration of pure EG-A-L and EG-A-S, and to the right is marked the migration of Bio-Rad low-molecular-weight markers. In the absence of added protease, EG-A-L was stable for 36 h (data not shown).

presumed signal peptide-encoding region and ribosome-binding site of *celA* were substituted with those of a gene encoding an α-amylase from *B. licheniformis*. (v) The NH_2 terminus of EG-A-S is very similar to the predicted NH_2 terminus of EG-A. (vi) The processed region of approximately 150 amino acids which is removed from EG-A-L enables it to bind to cellulose, an ability which EG-A-S does not have. Activity against CMC is independent of the presence of this region, while activity against insoluble substrates is enhanced by its presence. (vii) Purified EG-A-L can be processed in vitro by *B. subtilis* supernatant, proteinase K, and trypsin to yield an active 57,000-M_r protein that may correspond to EG-A-S.

We conclude that processing of EG-A-L is the result of exposition of an interdomain region to unspecific cleavage by host proteases, since purified EG-A-L is stable.

EGs from *B. subtilis* are apparently processed by a mechanism much like that of EG-A when expressed in *E. coli* and *B. subtilis*. They are initially translated as proteins with an M_r of 55,000, and in the case of the EGs from *B. subtilis* DLG and PAP115, mature proteins with M_rs of 36,000 and 32,000 were detected (26, 27, 42).

Removal of cellulose-binding domains has also been observed with CenA and Cex of *C. fimi* (13). These proteins contain homologous cellulose-binding domains at their NH_2 and COOH termini, respectively (24, 25, 39). Nonglycosylated, heterologously expressed versions of these proteins are cleaved at specific interdomain positions by a *C. fimi* protease. However, the glycosylated CenA and Cex produced by *C. fimi* are resistant to proteolytic degradation when bound to cellulose and only slowly processed in solution (13). It remains to be shown whether EG-A produced by *B. lautus* PL236 is protected against proteolytic attack by glycosylation.

EG-A from *B. lautus* PL236 is a further example of multidomain cellulases, typical of cellulases from organisms capable of crystalline-cellulose degradation (12). EG-A is also typical for its susceptibility to proteolytic attack, a major problem for large-scale industrial production and application of proteins constituted of various domains.

ACKNOWLEDGMENTS

We thank Pierre Béguin, Georges Rapoport, Bernard Henrissat, Martin Schülein, Søren Molin, and Flemming Hansen for valuable discussions and support. We thank Georges Rapoport, in whose laboratory much of the work was carried out, for reading and commenting on the manuscript. J. C. Huet and Torben Halkier are thanked for NH_2-terminal sequencing work, and Kjeld and Fanny Norris are thanked for oligodeoxynucleotide synthesis. Patricia Vary is thanked for supplying *B. megaterium* PV291. Gitte Bak Poulsen, Monica Moltke, Eva Gertman, and Carlos Belinchon are thanked for excellent technical assistance.

Without the generous support of G.D.-XII of the Commission of the European Communities, this study would not have been possible.

REFERENCES

1. **Béguin, P., P. Cornet, and J.-P. Aubert.** 1985. Sequence of a cellulase gene of the thermophilic bacterium *Clostridium thermocellum*. J. Bacteriol. **162:**102–105.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
3. **Bolivar, F., R. L. Rodriguez, P. J. Green, M. C. Betlach, H. L. Heyneker, H. M. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2:**95–113.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
5. **Carter, H. L., III, L.-F. Wang, R. H. Doi, and C. P. Moran, Jr.** 1988. Promoter in the *rpoO* operon used by σ^H-RNA polymerase in *Bacillus subtilis*. J. Bacteriol. **170:**1617–1621.
6. **Casadaban, J. M., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138:**179–207.
7. **Chang, S., and S. N. Cohen.** 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. Mol. Gen. Genet. **168:**111–115.
8. **Diderichsen, B., U. Wedsted, L. Hedegaard, B. R. Jensen, and C. Sjøholm.** 1990. Cloning of *aldB*, which encodes α-acetolactate decarboxylase, an exoenzyme from *Bacillus brevis*. J. Bacteriol. **172:**4315–4321.
9. **Dubnau, D., and R. Davidoff-Abelson.** 1971. Fate of transforming DNA following uptake by competent *B. subtilis*. Formation and properties of the donor recipient complex. J. Mol. Biol. **56:**209–221.
10. **Ehrlich, S. D.** 1977. Replication and expression of plasmids from *Staphylococcus aureus* in *B. subtilis*. Proc. Natl. Acad. Sci. USA **74:**1680–1682.
11. **Gaboriaud, C., V. Bissery, T. Benchetrit, and J. P. Mornon.** 1987. Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett. **224:** 149–155.
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:**303–315.
13. **Gilkes, N. 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 a homologous protease and the effect on catalysis. J. Biol. Chem. **263:**10401–10407.
14. **Gilman, M. Z., J. L. Wiggs, and M. J. Chamberlin.** 1981. Nucleotide sequence of two *Bacillus subtilis* promoters used by *Bacillus subtilis* sigma 28 RNA polymerase. Nucleic Acids Res. **9:**5991–6000.
15. **Hazlewood, G. P., M. P. M. Romaniec, K. Davidson, O. Grepinet, P. Béguin, J. Millet, O. Raynaud, and J.-P. Aubert.** 1988. A catalogue of *Clostridium thermocellum* endoglucanase, β-glucosidase and xylanase genes cloned in *Escherichia coli*. FEMS Microbiol. Lett. **51:**231–236.
16. **Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle, and J.-P. Mornon.** 1989. Cellulase families revealed by hydrophobic cluster analysis. Gene **81:**83–95.
17. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. **150:**815–825.
18. **Howard, G. T., and B. White.** 1988. Molecular cloning and expression of cellulase genes from *Ruminococcus albus* 8 in

Escherichia coli bacteriophage lambda. Appl. Environ. Microbiol. **54:**1752–1755.

19. **Jauris, S., K. P. Rücknagel, W. H. Schwartz, P. Kratzsch, K. Bronnenmeier, 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. **223:**258–267.
20. **Jørgensen, P. L., and C. K. Hansen.** 1990. Multiple endo-β-1,4-glucanase-encoding genes from *Bacillus lautus* PL236 and characterization of the *celB* gene. Gene **93:**55–60.
21. **Jørgensen, P. L., C. K. Hansen, G. B. Poulsen, and B. Diderichsen.** 1990. In vivo genetic engineering: homologous recombination as a tool for plasmid construction. Gene **96:**37–41.
22. **Keggins, K. M., P. S. Lovett, and E. J. Duvall.** 1978. Molecular cloning of *Bacillus* DNA in *Bacillus subtilis* and properties of the vector plasmid pUB110. Proc. Natl. Acad. Sci. USA **75:** 1423–1427.
23. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
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. **Lo, A., R. M. Mackay, V. L. Seligy, and G. E. Willick.** 1988. *Bacillus subtilis* β-1,4-endoglucanase products from intact and truncated genes are secreted into the extracellular medium by *Escherichia coli*. Appl. Environ. Microbiol. **54:**2287–2292.
27. **MacKay, R. M., A. Lo, G. Willick, M. Zuker, S. Baird, M. Dove, F. Moranelli, and V. Seligy.** 1986. Structure of a *Bacillus subtilis* endo-1,4-β-glucanase. Nucleic Acids Res. **14:**9159–9170.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262:**10035–10038.
30. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
31. **McLaughlin, J. R., C. L. Murray, and J. C. Rabinowitz.** 1981. Unique features in the ribosome-binding site sequence of the gram-positive *Staphylococcus aureus* β-lactamase gene. J. Bacteriol. **256:**11283–11291.
32. **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.
33. **Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miler, 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.
34. **Miller, G. L.** 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugars. Anal. Chem. **31:**426–428.
35. **Moran, C. P., Jr., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonensheim, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186:**339–346.
36. **Nakamura, A., T. Uozomi, and T. Beppu.** 1987. Nucleotide sequence of a cellulase gene of *Bacillus subtilis*. Eur. J. Biochem. **164:**317–320.
37. **Nakamura, L. K.** 1984. *Bacillus amylolyticus* sp. nov., nom. rev., *Bacillus lautus* sp. nov., nom. rev., *Bacillus pabuli* sp. nov., nom. rev., and *Bacillus validus*. Int. J. Syst. Bacteriol. **34:**224–226.
38. **Navarro, A., M.-C. Chebrou, P. Béguin, and J.-P. Aubert.** 1991. Nucleotide sequence of the cellulase gene *celF* of *Clostridium thermocellum*. Res. Microbiol. **142:**927–936.
39. **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* gene for endoglucanase B. Appl. Environ. Microbiol. **54:**518–523.
40. **Palva, I., R. F. Petterson, N. Kalkkunen, P. Lehtovaara, M. Sarvas, H. Soederlund, K. Takkinen, and L. Kaariainen.** 1981. Nucleotide sequence of the promoter and NH_2-terminal signal peptide region of the α-amylase gene from *Bacillus amyloliquefaciens*. Gene **15:**43–51.
41. **Rather, P. N., R. E. Hay, G. L. Ray, W. G. Haldenwang, and C. P. Moran.** 1986. Nucleotide sequences that define promoters that are used by *Bacillus subtilis* 629 RNA polymerase. J. Mol. Biol. **192:**557–565.
42. **Robson, L. M., and G. H. Chambliss.** 1987. Endo-β-1,4-glucanase gene of *Bacillus subtilis* DLG. J. Bacteriol. **169:**2017–2025.
43. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
44. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
45. **Saul, D. J., L. C. Williams, D. R. Love, L. W. Chamley, and P. L. Bergquist.** 1988. Nucleotide sequence of a gene from *Caldocellum saccharolyticum* encoding for exoglucanase and endoglucanase activity. Nucleic Acids Res. **17:**439.
46. **Sloma, A., A. Ally, and J. Pero.** 1988. Gene encoding a minor extracellular protease in *Bacillus subtilis*. J. Bacteriol. **170:** 5557–5563.
47. **Sloma, A., C. F. Rudolph, G. A. Rufo, Jr., B. J. Sullivan, K. A. Theriault, D. Ally, and J. Pero.** 1990. Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*. J. Bacteriol. **172:**1024–1029.
48. **Sloma, A., G. A. Rufo, Jr., K. A. Theriault, M. Dwyer, S. W. Wilson, and J. Pero.** 1991. Cloning and characterization of the gene for an additional extracellular serine protease of *Bacillus subtilis*. J. Bacteriol. **173:**6889–6895.
49. **Teather, R. M., and P. J. Wood.** 1982. Use of Congo red-polysaccharide interactions in the enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. **43:**777–780.
50. **Vary, P. S., and Y.-P. Tao.** 1988. Development of genetic methods in *Bacillus megaterium*, p. 403–407. *In* A. T. Ganesan and J. A. Hoch (ed.), Genetics and biotechnology of bacilli. Vol. 2. Academic Press, Inc., San Diego.
51. **von Tersch, M. A., and B. C. Carlton.** 1983. Megacinogenic plasmids of *Bacillus megaterium*. J. Bacteriol. **155:**872–877.
52. **Wu, X.-C., S. Nathoo, A. S.-H. Pang, T. Carne, and S.-L. Wong.** 1990. Cloning, genetic organization, and characterization of a structural gene encoding bacillopeptidase F from *Bacillus subtilis*. J. Biol. Chem. **265:**6845–6850.
53. **Yang, M. Y., E. Ferrari, and D. J. Henner.** 1984. Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an in vitro-derived deletion mutation. J. Bacteriol. **160:**15–21.