APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Feb. 2003, p. 869–877 Vol. 69, No. 2
0099-2240/03/$08.00+0 DOI: 10.1128/AEM.69.2.869–877.2003

Production by *Clostridium acetobutylicum* ATCC 824 of CelG, a Cellulosomal Glycoside Hydrolase Belonging to Family 9

Ana M. López-Contreras,[1,2]* Aernout A. Martens,[1] Nora Szijarto,[2] Hans Mooibroek,[2] Pieternel A. M. Claassen,[2] John van der Oost,[1] and Willem M. de Vos[1]

Laboratory of Microbiology[1] and Agrotechnological Research Institute (ATO),[2] Wageningen University and Research Centre, Wageningen, The Netherlands

Received 17 July 2002/Accepted 31 October 2002

The genome sequence of *Clostridium acetobutylicum* ATCC 824, a noncellulolytic solvent-producing strain, predicts the production of various proteins with domains typical for cellulosomal subunits. Most of the genes coding for these proteins are grouped in a cluster similar to that found in cellulolytic clostridial species, such as *Clostridium cellulovorans*. CAC0916, one of the open reading frames present in the putative cellulosome gene cluster, codes for CelG, a putative endoglucanase belonging to family 9, and it was cloned and overexpressed in *Escherichia coli*. The overproduced CelG protein was purified by making use of its high affinity for cellulose and was characterized. The biochemical properties of the purified CelG were comparable to those of other known enzymes belonging to the same family. Expression of CelG by *C. acetobutylicum* grown on different substrates was studied by Western blotting by using antibodies raised against the purified *E. coli*-produced protein. Whereas the antibodies cross-reacted with CelG-like proteins secreted by cellobiose- or cellulose-grown *C. cellulovorans* cultures, CelG was not detectable in extracellular medium from *C. acetobutylicum* grown on cellobiose or glucose. However, notably, when lichenan-grown cultures were used, several bands corresponding to CelG or CelG-like proteins were present, and there was significantly increased extracellular endoglucanase activity.

Among the members of the bacterial genus *Clostridium* several species, including *Clostridium acetobutylicum*, have the capacity to produce organic solvents, as well as organic acids, during fermentation of a wide range of carbohydrate substrates. *Clostridium* strains have been used to develop industrial production of acetone, butanol, and ethanol, the so-called ABE fermentation. Although this process was abandoned several decades ago due to its inability to compete with the petrochemical synthesis of these solvents (9, 13), new possibilities for more sustainable solvent production via ABE fermentation of cheap substrates have been suggested (5). These substrates include lignocellulosic materials, such as domestic organic waste, that contain sugars which can be utilized during ABE fermentation (4). However, cellulose, the main component of lignocellulose, cannot be degraded by solvent-producing clostridia, and use of cellulose as a substrate for ABE fermentation requires addition of expensive cellulolytic enzymes (19). Hence, there is a need to develop cellulose-degrading clostridia that are able to produce solvents. One way to approach this is to clone and express genes involved in cellulose degradation in appropriate solvent-producing clostridia (20). However, this requires a basic understanding of why solventogenic clostridia do not ferment cellulose, while other, non-solvent-producing clostridial strains do so effectively (25).

Recently, the complete genome of *C. acetobutylicum* ATCC 824 has been sequenced; this organism is a solventogenic strain that has been extensively characterized at the physiological and genetic levels (21). This organism is able to utilize a broad range of mono- and disaccharides, starches, and other substrates, such as pectin, inulin, whey, and xylan. However, it does not ferment or bind to cellulose (10), although, like other solvent-producing strains, it produces extracellular cellulolytic enzymes during growth on different substrates (18).

Analysis of the *C. acetobutylicum* ATCC 824 genome revealed the presence of a number of genes coding for enzymes involved in cellulose degradation. Some of these genes appear to be grouped in a gene cluster similar to the cluster found in true cellulolytic clostridia, such as *Clostridium cellulovorans* and *Clostridium cellulolyticum*. The cellulosome is an extracellular high-molecular-mass complex consisting of a number of catalytic components (glycohydrolases) that tightly bind to one large noncatalytic scaffolding protein that attaches to cellulose and to the cell surface, thereby minimizing the diffusion of hydrolytic products (25, 29). The production of cellulosomes is typical of cellulolytic clostridial strains, and until now there was no evidence that any solvent-producing strain produced them or had the genes coding for cellulosomal components in their chromosomes.

Since *C. acetobutylicum* ATCC 824 does not grow on cellulose, the question arises whether the genes in the presumed cellulosomal gene cluster indeed code for secreted and cellulose-binding glycoside hydrolases and whether these genes are expressed under the appropriate conditions. In this study we focused on one of the genes that is present in the putative cellulosome gene cluster, the open reading frame (ORF) CAC916, which codes for a putative glycoside hydrolase belonging to family 9 (6, 7). This gene was cloned in *Escherichia coli*, and the overproduced purified protein was functionally characterized. Expression studies of this gene in *C. acetobuty-*

* Corresponding author. Present address: Bioconversion Group, Agrotechnological Research Institute (ATO), P.O. Box 17, 6700 AA Wageningen, The Netherlands. Phone: 31 317 478572. Fax: 31 317 475347. E-mail: a.m.lopezcontreras@ato.wag-ur.nl.

licum grown on different substrates were performed and revealed that in lichenan- and xylose-grown cultures CelG-like proteins were secreted.

MATERIALS AND METHODS

Bacterial strains, media, and growth conditions. *C. acetobutylicum* ATCC 824 was kindly supplied by P. Soucaille (INSA, Toulouse, France). Stock cultures were maintained as spore suspensions in sterile 10% (vol/vol) glycerol at −20°C. Spore suspensions were heat shocked for 10 min at 75°C in a water bath prior to inoculation. For production of precultures, vegetative cells were grown overnight at 37°C in clostridial growth medium (23, 32), which contained (per liter) 0.75 g of KH_2PO_4, 0.75 g of K_2HPO_4, 0.348 g of $MgSO_4$, 0.01 g of $MnSO_4 \cdot H_2O$, 0.01 g of $FeSO_4 \cdot 7H_2O$, 1 g of NaCl, 2 g of asparagine, 5 g of yeast extract, 2 g of $(NH_4)_2SO_4$, and 50 g of glucose. For growth experiments the same medium was used except that the glucose was replaced by one of the following carbon sources at a concentration of 2% (wt/vol): glucose (Merck, Darmstadt, Germany); or cellobiose, xylose, lichenan, or laminarin (all obtained from Sigma, Zwijndrecht, The Netherlands). All experiments were performed in an anaerobic chamber (Coy Laboratory Products, Grass lake, Mich.) under an atmosphere containing 20% CO_2, 4% H_2, and 76% N_2, unless indicated otherwise.

C. cellulovorans DSM 3052 (28) was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen culture collection and was cultivated in medium 320 with cellulose MN 300 (Macherey-Nagel, Düren, Germany) or cellobiose (Sigma) as the carbon source. For vector construction and protein production *E. coli* strains XL1 blue (Stratagene, Amsterdam, The Netherlands) and M15(pREP4) (Qiagen, Germantown, Md.), respectively, were used. These strains were grown in Luria-Bertani broth as described previously (24), and the medium was supplemented, when appropriate, with ampicillin (50 μg/ml), isopropyl-β-thiogalactopyranoside (IPTG) (50 μg/ml), 5-bromo-4-chloro-3-indolyl-β-galactoside (X-Gal) (40 μg/ml), or kanamycin (50 μg/ml).

Transformation procedures, DNA manipulation, and PCR. All general DNA manipulations in *E. coli* were carried out essentially as described previously (24). Restriction endonucleases and modification enzymes were purchased from Roche Diagnostics (Mijdrecht, The Netherlands), Eurogentec (Maastricht, The Netherlands), or Qiagen. DNA isolation from *E. coli* was performed with a Wizard Plus SV miniprep kit (Promega Inc., Leiden, The Netherlands). Genomic DNA from *C. acetobutylicum* ATCC 824 was isolated as described previously (22).

Oligonucleotides used for mutagenesis or PCR were purchased from Eurogentec. The DNA fragments containing the desired mutations were cloned first into pGEMT-Easy (Promega Inc.) and subsequently into the expression vector pQE60 (Qiagen). The mutations were verified by sequencing by using an automated laser fluorescent ALF DNA sequencer (Amersham Pharmacia Biotech, Roosendaal, The Netherlands) and fluorescently labeled M13 universal and reverse primers. The complete coding sequence of the *celG* gene from *C. acetobutylicum* ATCC 824 was obtained by PCR as an approximately 2.1-kb *Nco*I/*Bam*HI fragment. The primers used were based on the sequence of the *celG* gene (described with gi:15894203 at the National Center for Biotechnology Information web site) and included the forward primer 5′-CGCCATGGAGAAGTTGT TAGCTACTTTG-3′ and the reverse primer 5′-GGCGGATCCTTACTGTT TACCTGAAATAA-3′, to which *Nco*I and *Bam*HI sites (underlined) were added, respectively. The *celG* gene without a signal peptide sequence was amplified by using the same reverse primer and forward primer 5′-CGCCATG GCTAAGAGTACAGAGGATAAAAAC-3′ (*Nco*I site is underlined), which resulted in a PCR fragment that was approximately 2 kb long. The *Nco*I sites were added in such a way that they allowed in-frame fusion of the coding region to the pQE60 expression sequences.

Analytical methods. Reaction products were analyzed with a Dionex BioLC high-performance chromatography system (Dionex Corporation) by using a Carbopac PA-100 anion-exchange column (25 cm by 4 mm) equilibrated in 0.05 M NaOH and pulsed amperometric detection. The samples loaded (20 μl) were eluted isocratically with 0.05 M NaOH for 2 min, followed by a 20-min linear gradient of 0.05 to 0.95 M NaOH at a rate of 1 ml/min. Peaks were identified by analyzing a standard mixture containing 1 mg of a cellooligosaccharide mixture (Sigma) per ml by using the same chromatographic conditions.

Preparation of culture samples for enzyme assays and Western blotting. Cells were sedimented by centrifugation at 10,000 × *g* at 4°C for 15 min. The culture supernatant was collected, filtered through a 0.22-μm-pore-size filter, and concentrated approximately 30-fold by ultrafiltration through a PM10 membrane (Millipore, Etten-Leur, The Netherlands) at 4°C. Low-molecular-weight compounds and medium components were removed from the concentrated material by dilution with 4 volumes of ice-cold 50 mM sodium citrate buffer (pH 5.7), followed by reconcentration by ultrafiltration. This process was repeated three times, and the resulting fractions were used for enzymatic activity determinations or for Western blot assays.

For preparation of cell extracts of *C. acetobutylicum*, cells (1 ml) were harvested by centrifugation, resuspended in 0.1 ml of 50 mM Tris HCl (pH 8.0) buffer, and sonicated with a Branson Sonifier. Cell debris was removed by centrifugation (10,000 × *g* for 20 min). The insoluble fraction was resuspended in 0.1 ml of the same buffer.

Cellulose-bound proteins in *C. cellulovorans* cultures were harvested by centrifuging an aliquot of the culture and washing the pellet, consisting of cellulose, cellulose-bound proteins, and cells, briefly with a buffer containing 8 M urea. Proteins bound to the cellulose were denatured and became soluble, whereas the bacterial cells remained mostly intact. The protein-containing soluble fraction was obtained by centrifugation.

Cellulose-binding proteins in *C. acetobulylicum* concentrated extracellular medium samples were extracted twice with 0.3% (wt/vol) Avicel or once with 0.5% (wt/vol) Avicel as indicated below for 45 min at 37°C. Proteins bound to cellulose were eluted as described above.

Cellulase activity assays. *E. coli* transformants grown on agar plates were screened for endoglucanase activity by the Congo red staining method described previously (30), with the following modifications: carboxymethyl cellulose (CMC) (high viscosity; Sigma) was added to the medium at a concentration of 0.1% (wt/vol), and carboxymethyl cellulase (CMCase) activity was detected after 2 to 4 days of incubation at 37°C by washing the cells out of the plates with sterile demineralized water and staining the CMC with Congo red. Zymograms were prepared as described previously (26).

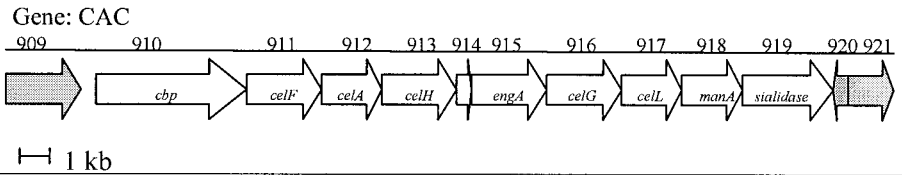
Glycolytic activities were determined as described previously (18). The following substrates (obtained from Sigma, unless indicated otherwise) were used (final concentrations): 0.5% (wt/vol) CMC (low viscosity), 0.2% (wt/vol) lichenan, 0.5% (wt/vol) laminarin, 0.5% (wt/vol) oat spelt xylan, and 0.5% (wt/vol) Avicel (Merck) in 50 mM citrate buffer (pH 5.7). Lichenan and oat spelt xylan were incubated with the enzyme samples in a water bath at 39°C for 10 to 60 min; Avicel, laminarin, and CMC were incubated for up to 30 h. The amounts of the reducing sugar formed were measured by the 3,5-dinitrosalicylic acid (Sigma) method (11). One unit of activity corresponded to the formation of 1 μmol of reducing sugar (D-glucose) per min, unless indicated otherwise. Protein concentrations in the samples were determined by the Bradford assay (Bio-Rad, Veenedaal, The Netherlands).

Purification of recombinant CelG. *E. coli* M15(pREP4) harboring the overexpression plasmid pWUR65 was grown at 37°C in Luria-Bertani medium containing ampicillin and kanamycin. Induction was performed by adding 1 mM IPTG and incubating the culture overnight at 18°C. Subsequently, the cells from 100 ml of the culture were collected by centrifugation, resuspended in 10 ml of PC buffer (50 mM KH_2PO_4, 10 mM sodium citrate; pH 7.0), and sonified with a Branson Sonifier. Clear cell extracts were obtained by centrifugation at 10,000 × *g* for 20 min at 4°C. CelG was purified from the cell extracts by using a previously described cellulose-binding approach (12, 27), as follows. Cell extracts were stirred for 40 min at room temperature with 0.3 g of Avicel microcrystalline cellulose. The Avicel was removed from the suspension by centrifugation, and extraction with Avicel was repeated three times. The three pellets of Avicel were pooled and subsequently washed once with 5 ml of 1 M NaCl in PC buffer and twice with 5 ml of demineralized water. The bound proteins were eluted by stirring the Avicel overnight in 300 ml of demineralized water at 4°C and were concentrated by ultrafiltration by using an Omegacell device with a polyethersulfone membrane with a 10-kDa cutoff (Pall Corporation, East Hills, N.Y.). By using fast-performance liquid chromatography, the concentrated sample was applied to a Superdex 200 column (Amersham Pharmacia Biotech, Roosendaal, The Netherlands) equilibrated with 50 mM phosphate buffer (pH 7). Eluted fractions were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and used for assays.

Immunological procedures. Aliquots containing 100 μg of purified CelG were mixed with an equal amount of adjuvant (Specoll) and injected into a New Zealand rabbit. Antiserum was collected 9 weeks after the first immunization. Western blot analyses were carried out with antiserum diluted 1,000-fold by using a standard protocol (24), and the detection limit for purified CelG under the conditions assayed was approximately 3 ng.

RESULTS

Analysis of the sequence of the protein encoded by ORF CAC0916. Analysis of the complete genome sequence of *C. acetobutylicum* ATCC 824 indicated that there was a putative



Gene	Gene product	Predicted modular structure	Number of residues	Mol wt
CAC0910	CBP	CBM_3-SLH-(SLH-cohesin)$_5$	1483	154,4
CAC0911	CelF	**GH48**-dockerin	726	80,7
CAC0912	CelA	**GH5**-dockerin	606	54,1
CAC0913	CelH	**GH9**-CBM_3-dockerin	712	79,7
CAC0914	CCP	cohesin	181	19,4
CAC0915	EngA	**GH44**-dockerin	606	66,8
CAC0916	CelG	**GH9**-CBM_3-dockerin	692	76,6
CAC0917	CelL	**GH9**-dockerin	539	59,8
CAC0918	ManA	**GH5**-dockerin	425	47,3
CAC0919	sialidase	**GH74**-dockerin	839	90,8
CAC0561	CelE	CBM_4-Ig-**GH9**-dockerin	878	96,1
CAC3469		SLH-**GH5**-CA-dockerin	1012	110,1

FIG. 1. Schematic representation of the putative cellulosome gene cluster present in the chromosome of *C. acetobutylicum* ATCC 824 and description of the ORFs found in the genome that contain cohesin or dockerin domains. Abbreviations: CBP, cellulose-binding protein; CCP, cohesin-containing protein; GH, glycoside hydrolase; CBM, carbohydrate-binding module; Ig, immunoglobulin-like domain; CA, cell adhesion domain; SLH, surface layer homology domain. The catalytic modules are indicated by boldface type. The number of residues includes the signal sequence. The molecular weights were determined from the peptide sequences.

gene cluster containing 10 unidirectionally transcribed genes predicted to encode secreted proteins with cohesin or dockerin domains, which in many cases contain glycoside hydrolase catalytic domains (21) (Fig. 1). The dockerin domain is typical of cellulosomal components and is usually located at the C terminus of the protein; it consists of two duplicated sequences containing 22 amino acids each and is highly conserved in all the proteins that bind to the scaffolding protein to form active cellulosomes.

The gene cluster contains three ORFs, CAC0913, CAC0916, and CAC0917, which code for putative secreted glycoside hydrolases with a catalytic domain belonging to family 9; the products of these ORFs are designated CelH (previously designated CelG; designated CelH due to its high level of similarity to EngH from *C. cellulovorans*), CelG, and CelL, respectively (Fig. 1). CelG has an N-terminal signal peptide sequence consisting of 23 amino acids typical of gram-positive bacteria (31). The mature CelG consists of a family 9 catalytic domain, a carbohydrate-binding module belonging to family 3 (6), and a dockerin domain located at its C terminus. The overall amino acid sequence, including the domain structure, was homologous to the amino acid sequences of EngH from *C. cellulovorans* (40% identity and 55% similarity) and CelJ from *C. cellulolyticum* (39% identity and 53% similarity) (Fig. 2). Comparison of the catalytic domain of CelG with the catalytic domains of the other enzymes belonging to the same family present in the *C. acetobutylicum* cluster revealed a higher level of similarity to CelL (64% identity and 75% similarity) than to CelH (43% identity and 58% similarity). The catalytic domains of different glycoside hydrolases belonging to family 9 produced by *C. acetobutylicum*, *C. cellulolyticum*, and *C. cellulovorans* contain highly conserved regions, including active-site residues, as indicated by similarity to the region experimentally determined for CelD from *Clostridium thermocellum* (3, 14) (Fig. 3).

The carbohydrate-binding module between the catalytic and dockerin domains is involved in cellulose binding, and it is present in two other proteins that are predicted to be encoded by the gene cluster, the cellulose-binding protein (CAC0910) and CelH (CAC0913) (Fig. 1). This module is highly conserved in these three proteins and other cellulosomal proteins from *C. cellulovorans*, *C. cellulolyticum*, and *Clostridium josui* (data not shown).

Production of CelG by *E. coli* and characterization of CelG. The *C. acetobutylicum celG* gene was cloned into the expression vector pQE60 with and without the coding sequence for its predicted 22-residue signal peptide sequence, resulting in plasmids pWUR64 and pWUR65, respectively. Upon induction with IPTG, *E. coli* M15(pREP4) harboring pWUR64 and pWUR65 produced an additional protein of the expected size (approximately 76 and 74 kDa, respectively). However, the strain harboring pWUR64 produced a much smaller amount of CelG than the strain harboring pWUR65 (data not shown). For this reason we continued our work with *E. coli* M15(pREP4) harboring pWUR65 that coded for the mature CelG.

Upon induction with 1 mM IPTG for 3 h at 37°C, most of the CelG produced by *E. coli*(pREP4)(pWUR65) was found in the insoluble fraction of the cells due to formation of inclusion bodies. However, when the incubation temperature was decreased to 18°C and the induction time was extended to over-

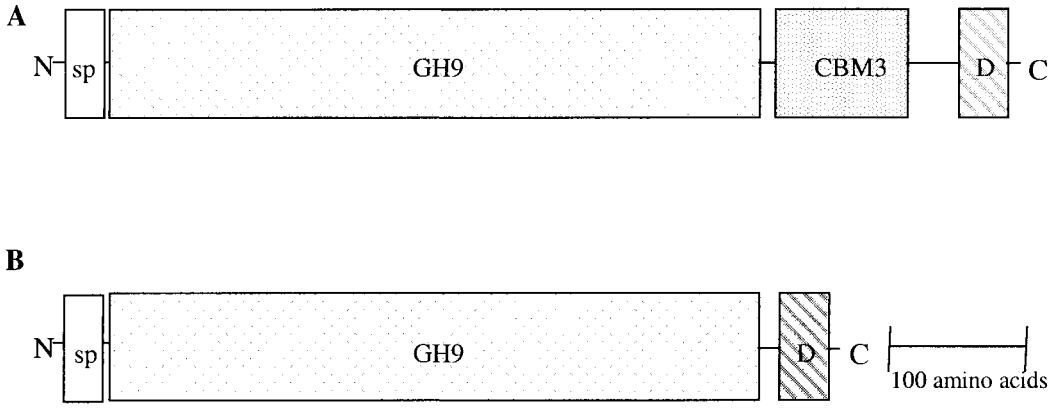


FIG. 2. Predicted domain structure comparison of glycoside hydrolases belonging to family 9 encoded in the gene clusters from *C. acetobutylicum*, *C. cellulovorans*, and *C. cellulolyticum*. (A) Structure found in *C. acetobutylicum* CelJ, *C. acetobutylicum* CelG, *C. cellulovorans* EngH, *C. cellulolyticum* CelG, and *C. cellulolyticum* CelJ. (B) Structure found in *C. acetobutylicum* CelL, *C. cellulovorans* EngL, and *C. cellulolyticum* CelM. sp, signal peptide; CBM3, carbohydrate-binding module family 3; D, dockerin domain.

night incubation, the major part of the mature CelG became soluble and could be purified from the cell extract by making use of its capacity to bind to cellulose. This was realized by incubation of the cell extract with microcrystalline cellulose, which resulted in binding of the major part of CelG, which could be recovered by extensive washing of the cellulose with demineralized water (Fig. 4). The recovered CelG was almost pure, and it was purified further to homogeneity by column chromatography by using Superdex 200 resin. Both in the cell extract and in the purified fractions the protein appeared as two bands of approximately the same size. Both proteins showed activity, as determined by zymogram analysis with CMC or lichenan as the substrate (results not shown). As expected from previous studies of similar enzymes (1), the major band, which was larger, corresponded to the full-length mature CelG and the smaller band may have corresponded to a truncated form in which the dockerin domain had been removed.

The purified CelG showed the highest CMCase activity at pH 6 and at 50°C (data not shown). The CMCase activities of related family 9 glycoside hydrolases from *C. cellulolyticum* (1) were also highest under similar conditions. CelG showed the maximum activity with CMC and less activity with lichenan, xylan, and Avicel (Table 1). The products obtained from the degradation of Avicel after 20 h of incubation with CelG were analyzed by high-performance liquid chromatography. A mixture of glucose, cellobiose, cellotriose, and cellotetraose was found to be the main product (results not shown).

Extracellular cellulase production by *C. acetobutylicum* ATCC 824 grown on different carbon sources. *C. acetobutylicum* ATCC 824 is able to utilize a wide range of carbon sources, including pentose sugars, such as xylose, and also insoluble sugar polymers, such as xylan (17). Lichenan is a polymer formed by β-1,4/1,3-linked glucose units, and although it has very limited solubility, it does not have a crystalline structure like that of cellulose. Laminarin, a polymer of β-1,3-linked glucose units, is easily soluble in water. Lichenase and laminarinase activities have been found in the extracellular medium of cultures of *Clostridium beijerinckii* (20) and hence were also used as substrates for growth of *C. acetobutylicum*. While growth on glucose, cellobiose, xylose, or lichenan was readily established (Fig. 4), on laminarin no growth was observed independent of preculture conditions. Since lichenan is mostly insoluble at the concentration used, growth was determined by measuring the amounts of protein in cell extracts of culture samples taken after different incubation times, while in all other cultures growth was monitored by measuring the optical density at 600 nm (Fig. 5).

At the late exponential growth phase cells were removed, and the extracellular medium was concentrated by ultrafiltration and dialyzed. The glycoside hydrolase activities present in the concentrated extracellular media were determined (Table 2). In the extracellular medium of cultures grown on lichenan the endoglucanase (CMCase) activity was significantly higher than the endoglucanase (CMCase) activities in the extracellular media of the cultures grown on the other substrates. The laminarinase and xylanase activities in cultures grown on lichenan or xylan were also higher than those in the cultures grown on glucose and cellobiose as the substrates. Under the assay conditions used, no reducing sugars were detected as degradation products of Avicel.

Expression of CelG by *C. acetobutylicum*. To determine the production and localization of CelG by *C. acetobutylicum*, Western blot analyses were performed by using polyclonal antibodies raised against the purified *E. coli*-produced CelG. The concentrated extracellular medium samples of cultures grown on glucose, cellobiose, xylose, and lichenan used for the activity assays shown in Table 2 were also used for the immunological assay (Fig. 6). In addition, extracellular medium samples from *C. cellulovorans* grown on cellulose and cellobiose were also prepared and used as controls, since this organism is able to grow on cellulose and its cellulolytic system has been extensively studied (8, 29). Moreover, CelG shows high levels of similarity to several of the glycoside hydrolases produced by *C. cellulovorans*, especially EngL and EngH (Fig. 2), which are components of the cellulosome complex that are encoded by the cellulosome gene cluster. EngH has a cellulose-binding module that exhibits a high level of similarity to the module present in CelG and is a similar size (approximately 79 kDa), while EngL does not have this binding module and is smaller

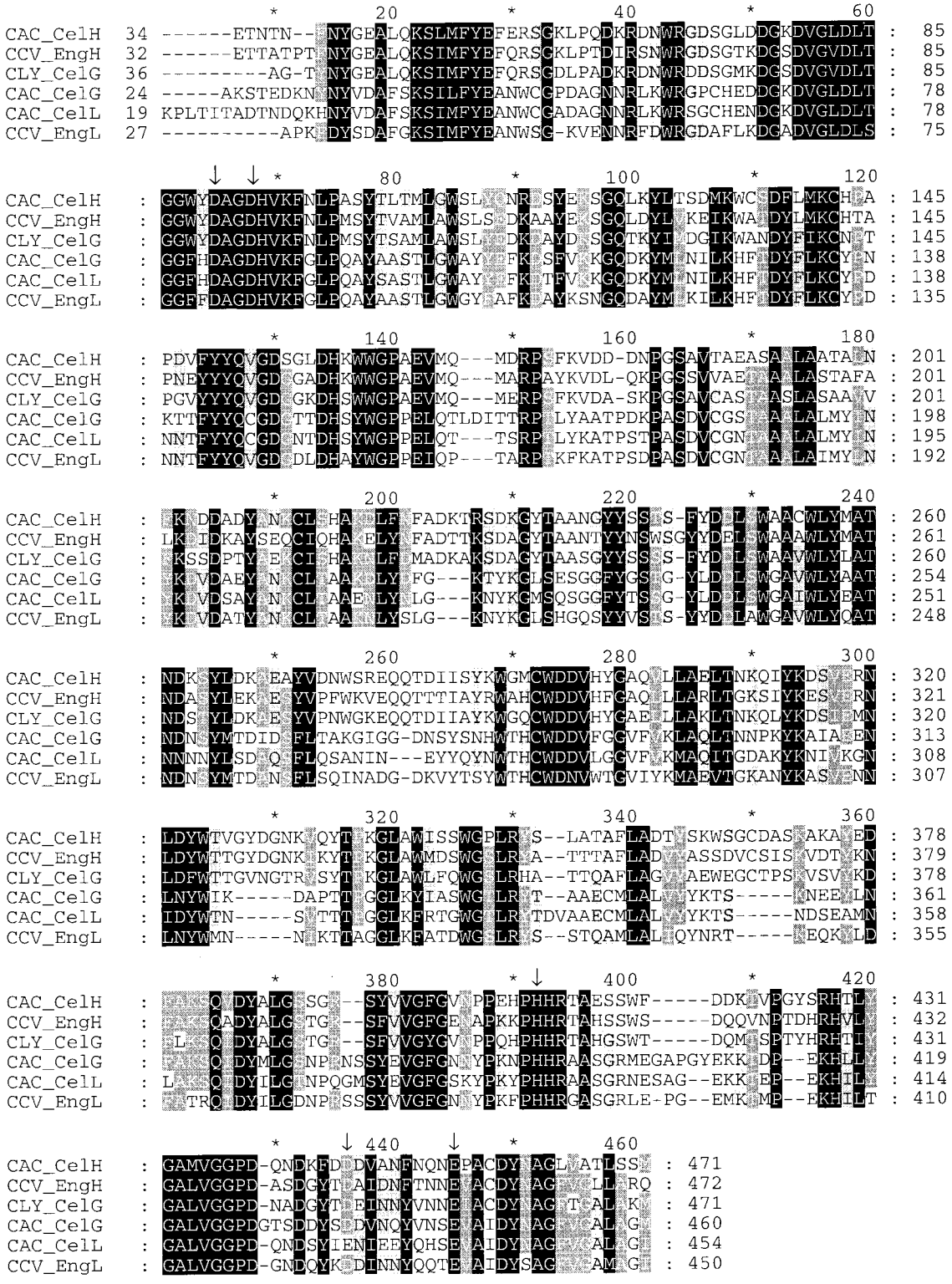


FIG. 3. Alignment of the catalytic domains of glycoside hydrolases belonging to family 9 encoded in the gene clusters of *C. acetobutylicum* (CAC), *C. cellulolyticum* (CLY), and *C. cellulovorans* (CCV) that contain a carbohydrate-binding module belonging to family 3. The arrows indicate possible active-site residues, as determined by sequence similarity with CelD from *C. thermocellum*, in which the residues were determined by site-directed mutagenesis.

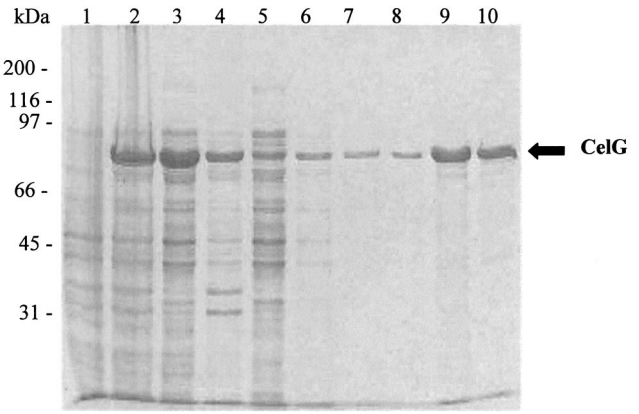


FIG. 4. Purification of CelG from *E. coli*(pREP4)(pWUR65). SDS-PAGE of noninduced cells (lane 1), induced cells (lane 2), cell extract of induced cells (lane 3), and the insoluble fraction (lane 4) showed that predominantly soluble CelG was produced. Further purification of CelG following extraction of the cell extract by Avicel was revealed by SDS-PAGE of the unbound fraction (lane 5), subsequent washes in 1 M NaCl (lane 6) or demineralized water (twice) (lanes 7 and 8), and preparations eluted with 8 M urea (lane 9) or demineralized water (lane 10).

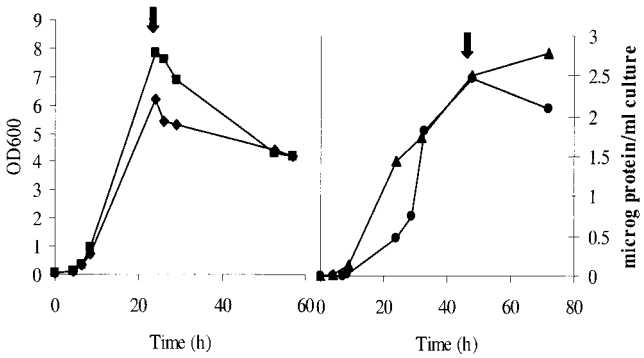


FIG. 5. Growth curves for cultures of *C. acetobutylicum* ATCC 824 on different substrates. The optical densities at 600 nm (OD600) were determined for cultures grown on cellobiose (■), glucose (◆), and xylose (●), while growth on lichenan (▲) was determined by monitoring the protein concentration in cell extracts. The concentration of protein in the cell extract from a lichenan-grown culture incubated for 48 h equaled that of a culture which was grown on glucose and had an optical density at 600 nm of approximately 2.0. The arrows indicate the times when cultures were collected to determine enzymatic activities (Table 2).

(approximately 50 kDa) (29). Due to the high levels of similarity among these proteins, particularly in their highly conserved catalytic domains (Fig. 3), cross-reactivity of the *C. acetobutylicum* anti-CelG antibodies with proteins present in extracellular medium from *C. cellulovorans* cultures could be expected.

C. cellulovorans grown on cellobiose or cellulose produced a larger amount of total extracellular protein (as calculated from concentrated samples, between 25 and 26 μg of protein/ml in both cultures) than the *C. acetobutylicum* cultures produced (up to 3 μg of protein per ml [Table 2]). Several extracellular proteins with affinity to anti-CelG were present in the supernatants of cultures of *C. cellulovorans*, and the amount was larger in the cellulose-grown culture (Fig. 6, lanes 1 and 2). The presence of multiple bands in the Western blot indicates that there were several proteins with similarity to CelG in the extracellular medium. In the *C. cellulovorans* cellulosome gene cluster there are four genes coding for proteins containing catalytic domains belonging to glycoside hydrolase family 9 (29), and these proteins or their degradation products are likely to cross-react with the anti-CelG antibodies. The cross-reacting protein with a molecular mass of approximately 75 kDa most likely corresponds to EngH, which is highly homologous to CelG (Fig. 2 and 3), and the band at approximately 56 kDa could correspond to EngL, which is highly homologous to the protein encoded by the gene CAC0917 (Fig. 1) (29). Interestingly, in the lane loaded with the cellulose-bound fraction of the *C. cellulovorans* cellulose-grown culture (Fig. 6, lane 3) there is only one major band corresponding to a protein with a molecular mass of approximately 75 kDa that could correspond to EngH (29).

Western blot analyses showed that cultures of *C. acetobutylicum* grown on glucose and cellobiose produced several proteins that cross-reacted with the anti-CelG antibodies, but all of them were considerably smaller than CelG (Fig. 6, lanes 4 and 5, respectively). No proteins that were the size of CelG were detected even when as much as 30 μg of protein was applied (data not shown). However, in the supernatants from the xylose- and lichenan-grown cultures, there were additional large proteins that reacted with the anti-CelG antibodies. In the medium from the lichenan-grown culture, one of anti-CelG-reacting proteins had a molecular mass (approximately 75 kDa) corresponding to that of the mature CelG produced

TABLE 1. Specific activities of purified *E. coli*-produced CelG on different substrates

Substrate	Sp act (U/mg of protein)[a]	%
CMC	7,400	100
Lichenan	2,300	31
Xylan	20	0.3
Avicel	7	0.1
Laminarin	<0.01	

[a] One unit of activity was defined as 1 nmol of reducing sugar released per min.

TABLE 2. Cellulolytic or hemicellulolytic activities in extracellular media of cultures of *C. acetobutylicum* ATCC 824 grown on different carbon sources

Carbon source[a]	Sp act (U/mg of protein)				
	Lichenase	Xylanase	CMCase	Laminarinase	Avicelase
Cellobiose	119.2	33.3	0.3	0.08	<0.02
Glucose	72.1	17.7	0.1	0.02	<0.01
Xylose	88.3	42.6	0.2	0.1	<0.01
Lichenan	90	43.4	4.5	0.2	<0.04

[a] Samples were taken at the end of the exponential growth phase. The average estimated protein concentrations in the extracellular medium calculated from the concentrations determined in the concentrated samples were 0.9 μg of protein/ml in cellobiose cultures, 3 μg of protein/ml in glucose cultures, 0.8 μg of protein/ml in xylose cultures, and 1.2 μg of protein in lichenan cultures. One unit of activity corresponded to the formation of 1 μmol of reducing sugar per min. The specific activity corresponded to units per milligram of protein in the concentrated medium sample.

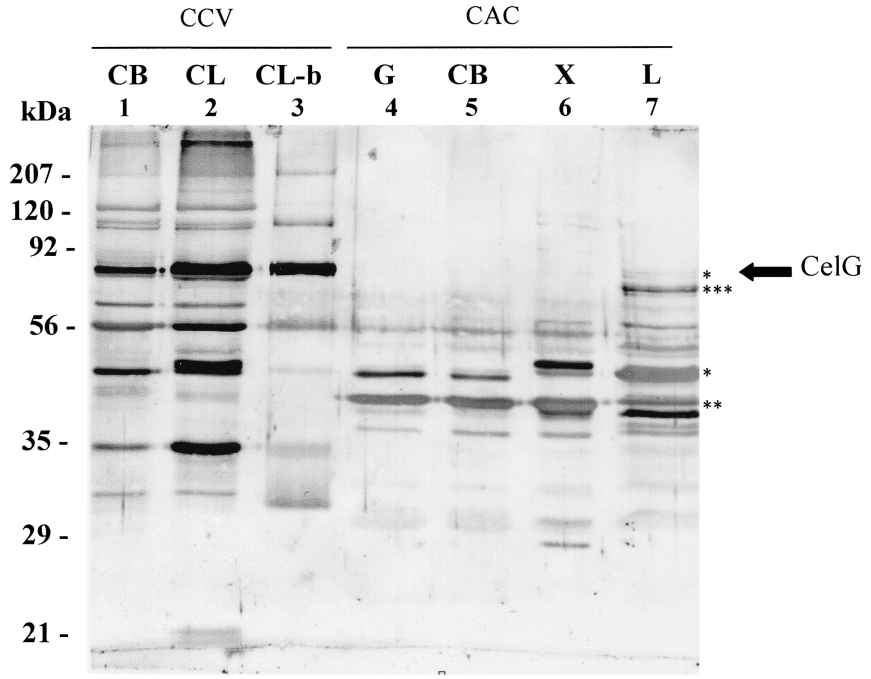


FIG. 6. Detection of CelG in extracellular growth medium of *C. acetobutylicum* (CAC) and *C. cellulovorans* (CCV) cultures. The Western blot was generated using polyclonal antibodies raised against purified CelG. Lane 1, cellobiose (CB)-grown *C. cellulovorans*; lane 2, cellulose (CL)-grown *C. cellulovorans*; lane 3, cellulose-bound proteins from cellulose-grown *C. cellulovorans*; lane 4, glucose (G)-grown *C. acetobutylicum*; lane 5, cellobiose (CB)-grown *C. acetobutylicum*; lane 6, xylose (X)-grown *C. acetobutylicum*; lane 7, lichenan (L)-grown *C. acetobutylicum*. In most lanes, 2.4 μg of protein was loaded; in lane 3 less than 1 μg of protein was applied. The arrow indicates the position of the mature *E. coli*-produced CelG. Bands that correspond to proteins in the extracellular medium of lichenan-grown cultures that bind to Avicel are indicated as follows: three asterisks indicate binding of the protein, two asterisks indicate binding of approximately 50% of the protein, and one asterisk indicates binding of approximately 10% of the protein.

by *E. coli*. Remarkably, a protein that was only slightly smaller than CelG was found to react strongly with the anti-CelG antibodies and could correspond to a degradation product of the secreted CelG or a further processed active form of this enzyme.

No protein corresponding to CelG could be detected in Western blots loaded with 10 μg of proteins from cytoplasmic and membrane fractions of *C. acetobutylicum* cells from the same cultures as the extracellular media used in the experiment whose results are shown in Fig. 6 (data not shown).

In an attempt to extract cellulose-binding proteins from concentrated extracellular media from cultures of *C. acetobutylicum* grown on glucose or cellobiose and cellobiose-grown *C. cellulovorans* cultures, Avicel (0.5%, wt/vol) was added to the samples. Following extraction for 1 h at 37°C, centrifugation, and subsequent resolubilization of the bound proteins in buffer with 8 M urea, no band corresponding to CelG from *C. acetobutylicum* was visible in the Western blots, while in samples from *C. cellulovorans* cultures several proteins (including one which was the size of EngH producing the major band) were found to cross-react with anti-CelG (data not shown).

To determine the binding capacity to Avicel of the anti-CelG-reacting proteins present in the extracellular medium of lichenan-grown cultures, an assay was carried out in which a concentrated sample was extracted with 0.3% (wt/vol) Avicel. The concentrated extracellular medium sample and the extracted sample were subjected to SDS-PAGE and Western blot analysis by using the anti-CelG antibodies. While some bands were not affected by the Avicel extraction, several bands in the Western blot disappeared completely (100%), or their intensities decreased in the extracted sample (approximately 50 or 10%) compared to the intensities in the concentrated medium sample (Fig. 6). The approximately 70-kDa anti-CelG-binding protein (Fig. 6, lane 7) bound significantly to Avicel and was not detectable in the extracted sample (data not shown). The protein having the molecular mass of mature CelG (approximately 74 kDa) was also extracted, although to a lesser extent (Fig. 6), indicating that there was binding to Avicel.

DISCUSSION

While *C. acetobutylicum* does not grow on cellulose, genomic analysis predicts the presence of genes coding for proteins that contain glycoside hydrolase as well as cohesin or dockerin domains (Fig. 1). A gene cluster containing 10 genes involved in cellulosome formation exhibits high levels of similarity, both in gene organization and predicted amino acid sequence, to the clusters found in true cellulolytic strains, such as *C. cellulolyticum*, *C. josui*, and *C. cellulovorans* (2, 15, 29). In these known cellulosome gene clusters the first gene codes for a nonenzymatic scaffolding protein (CBP), and this gene is followed by a gene coding for a glycoside hydrolase belonging to family 48 and then by genes coding for glycoside hydrolases

belonging to different families and other small cohesin-containing proteins. Another conserved feature is the presence of at least three genes coding for enzymes containing a catalytic domain belonging to family 9 glycoside hydrolases, and this suggests that these enzymes could play an important role in cellulosomal function. Outside the gene cluster there is a single ORF (CAC0561) present in the genome of *C. acetobutylicum* which codes for a putative cellulosomal glycoside hydrolase belonging to family 9 (Fig. 1). In order to study expression of the putative cellulosomal genes and to analyze glycolytic activity in *C. acetobutylicum*, we focused on CelG, the protein coded by the ORF CAC0916.

A *celG*-overexpressing *E. coli* strain was constructed that produced large amounts of intracellular CelG which could be easily purified by making use of the high affinity for cellulose of its cellulose-binding domain (CBM3). Purified CelG had characteristics similar to those of other enzymes belonging to family 9 of glycoside hydrolases (Table 1). It exhibited the highest activity with CMC and somewhat lower activity with lichenan, and its relative activities with xylan and Avicel were low, all of which are typical features of endoglucanases.

C. acetobutylicum was grown on different substrates, and the extracellular hydrolytic activities produced were determined. In concentrated extracellular media from lichenan-grown cultures there was significantly higher endoglucanase (CMCase) activity than in concentrated extracellular media from any of the other cultures. In xylan- or lichenan-grown cultures the extracellular xylanase activity was higher than the activity in cellobiose- or glucose-grown cultures (Table 2). In general, the extracellular glycoside hydrolase activities measured were lowest in glucose-grown cultures, suggesting that there is a catabolite repression mechanism. In all cases some residual laminarinase activity was observed; this activity was highest in lichenan- and xylose-grown cultures, but apparently it was too low to support growth on laminarin.

To determine expression of the *celG* gene, polyclonal antibodies raised against purified CelG from *E. coli* were used. The concentrated medium samples used in the activity assays were used to detect the presence of CelG by Western blot analysis (Fig. 6). Only in the sample from the culture grown on lichenan did a distinct protein which was the size of the mature CelG (approximately 74 kDa) react with the anti-CelG antibodies (Fig. 6). In the samples from lichenan- and xylose-grown cultures, additional proteins that were smaller than CelG and were not present in the glucose- or cellobiose-grown cultures reacted with the anti-CelG antibodies. These proteins were likely to be degradation products of CelG or other smaller proteins, indicating that xylose and lichenan induce the production of CelG-like proteins. In the case of *C. cellulovorans* it is known that although the composition of the cellulosomes can be regulated by the growth substrate, major cellulosomal subunits are constitutively expressed (16). This could explain the finding that in cellobiose- or cellulose-grown cultures of *C. cellulovorans* basically the same set of anti-CelG-binding proteins were present (Fig. 6, lanes 1 and 2). Strongly reacting bands appeared; one of these bands was approximately the same size as mature CelG (approximately 74 kDa) and probably represented EngH, and another had a molecular mass of approximately 56 kDa and could have corresponded to EngL, an enzyme with a catalytic domain of glycoside hydrolase family 9 and a dockerin domain but with no carbohydrate-binding module (Fig. 2) (29). In the cellulose-bound fraction of the *C. cellulovorans* cellulose-grown culture (Fig. 6, lane 3) the major anti-CelG-binding protein appeared to be the same size as CelG, and no strong reaction with smaller proteins was observed.

This study is the first functional analysis of a gene encoding a putative cellulosomal enzyme from *C. acetobutylicum*. Possible explanations for the lack of cellulolytic activity in *C. acetobutylicum* include (i) the absence of functional genes coding for necessary specific cellulosomal subunits, (ii) the presence of mutations in genes coding for cellulosomal components that could lead to the production of unstable or inactive cellulosomes, or (iii) deregulated expression of the cellulosome components or the transport systems needed for efficient export and assembly in the extracellular medium. In this study we showed that one of the genes in the putative cellulosome gene cluster codes for an active endoglucanase enzyme with a functional cellulose-binding domain. This enzyme was not detectable in the extracellular medium of glucose- or cellobiose-grown cultures of *C. acetobutylicum*. However, growth on xylose or on lichenan gave rise to enhanced production of CelG-related proteins by *C. acetobutylicum*. In the extracellular medium of lichenan-grown cultures, a weak band corresponding to a protein that was the same size as the purified mature *E.coli*-produced CelG was detectable by Western blotting, and it bound to a certain extent to Avicel (Fig. 6). In the same sample, a slightly smaller protein that reacted with anti-CelG bound to Avicel, suggesting that it possibly contains a carbohydrate-binding module with affinity for cellulose, which could indicate that this protein is a processed active form of CelG (its molecular mass could correspond to that of mature CelG without the dockerin domains, approximately 68 kDa) The extracellular endoglucanase activity found in lichenan-grown cultures was higher than that in the other cultures (Table 2), which is consistent with the presence of CelG in the medium, although probably the increased activity cannot be attributed to CelG alone. These induction results suggest that the regulation of the cellulosomal enzymes in *C. acetobutylicum* is different from that in *C. cellulovorans*, in which expression of CelG-like genes was observed in both cellulose- and cellobiose-grown cultures. Whether this difference in regulation can be related to the inability of *C. acetobutylicum* to utilize cellulose needs to be investigated. A detailed study of the genes involved in cellulosome formation and their expression patterns and characterization of their products are required to determine whether and under what conditions *C. acetobutylicum* is able to utilize cellulose or cellulosic substrates as an approach for cost-effective production of solvents.

ACKNOWLEDGMENTS

We thank Harry Kester for performing the analysis of oligosaccharides.

The work described in this paper was supported in part by an EU Madam Curie fellowship (grant FAIR-CT96-5047) to A. M. López-Contreras.

REFERENCES

1. **Belaich, A., G. Parsiegla, L. Gal, C. Villard, R. Hasser, and J.-P. Belaich.** 2002. Cel9M, a new family 9 cellulase of the *Clostridium cellulolyticum* cellulosome. J. Bacteriol. **184:**1378–1384.

2. **Belaich, J.-P., C. Tardif, A. Belaich, and C. Gaudin.** 1997. The cellulolytic system of *Clostridium cellulolyticum*. J. Biotechnol **57:**3–14.
3. **Chauvaux, S., P. Beguin, and J.-P. Aubert.** 1992. Site-directed mutagenesis of essential carboxylic residues in *C. thermocellum* endoglucanase CelD. J. Biol. Chem. **267:**4472–4478.
4. **Claassen, P. A. M., M. A. W. Budde, and A. M. López-Contreras.** 2000. Acetone, butanol and ethanol production from domestic organic waste by solventogenic clostridia. J. Mol. Microbiol. Biotechnol. **2:**39–44.
5. **Claassen, P. A. M., J. B. van Lier, A. M. López-Contreras, E. W. J. van Niel, L. Sijtsma, A. J. M. Stams, S. S. de Vries, and R. A. Weusthuis.** 1999. Utilisation of biomass for the supply of energy carriers. Appl. Microbiol. Biotechnol **52:**741–755.
6. **Coutinho, P. M., and B. Henrissat.** 18 December 2002, revision date. Carbohydrate-active enzymes server. [Online.] http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.
7. **Coutinho, P. M., and B. Henrissat.** 1999. Carbohydrate-active enzymes: an integrated database approach, p. 3–12. *In* H. J. Gilbert, G. Davies, B. Henrissat, and B. Svensson (ed.), Recent advances in carbohydrate bioengineering. The Royal Society of Chemistry, Cambridge, United Kingdom.
8. **Doi, R. H., J.-S. Park, C.-C. Liu, L. M. Malburg, Jr., Y. Tamaru, A. Ichiisi, and A. Ibrahim.** 1998. Cellolosome and noncellulosomal cellulases of *Clostridium cellulovorans*. Extremophiles **2:**53–60.
9. **Dürre, P.** 1998. New insights and novel developments in clostridial acetone/butanol/isopropanol fermentation. Appl. Microbiol. Biotechnol **49:**639–648.
10. **Gelhaye, E., H. Petitdemange, and R. Gay.** 1993. Adhesion and growth rate of *Clostridium cellulolyticum* ATCC 35319 on crystalline cellulose. J. Bacteriol. **175:**3452–3458.
11. **Ghose, T. K.** 1987. Measurement of cellulase activities. Pure Appl. Chem. **59:**257–268.
12. **Goldstein, M. A., M. Takagi, S. Hashida, O. Shoseyov, R. H. Doi, and I. H. Segel.** 1993. Characterization of the cellulose-binding domain of the *Clostridium cellulovorans* cellulose-binding protein A. J. Bacteriol. **175:**5762–5768.
13. **Jones, D. T., and D. R. Woods.** 1986. Acetone-butanol fermentation revisited. Microbiological Rev. **50:**484–524.
14. **Juy, M., A. G. Amit, P. M. Alzari, R. J. Poljack, M. Claeyssens, P. Beguin, and J.-P. Aubert.** 1992. Three-dimensional structure of a thermostable bacterial cellulase. Nature **357:**89–91.
15. **Kakiuchi, M., A. Isui, K. Suzuki, T. Fujino, E. Fujino, T. Kimura, S. Karita, K. Sakka, and K. Ohmiya.** 1998. Cloning and DNA sequencing of the genes encoding *Clostridium josui* scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome. J. Bacteriol. **180:**4303–4308.
16. **Kosugi, A., K. Murashima, and R. H. Doi.** 2001. Characterization of xylanolytic enzymes in *Clostridium cellulovorans*: expression of xylanase activity dependent on the growth substrate. J. Bacteriol. **183:**7037–7043.
17. **Lee, S. F., C. W. Forsberg, and L. N. Gibbings.** 1985. Xylanolytic activity of *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **50:**1068–1076.
18. **Lee, S. F., C. W. Forsberg, and L. N. Gibbins.** 1985. Cellulolytic activity of *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **50:**220–228.
19. **López-Contreras, A. M., P. A. M. Claassen, A. Mooibroek, and W. M. de Vos.** 2000. Utilisation of saccharides in extruded domestic organic waste by *Clostridium acetobutylicum* ATCC 824 to produce acetone, butanol and ethanol. Appl. Microbiol. Biotechnol **54:**162–167.
20. **López-Contreras, A. M., H. Smidt, J. van der Oost, P. A. M. Claassen, H. Mooibroek, and W. M. de Vos.** 2001. *Clostridium beijerinckii* cells expressing *Neocallimastix patriciarum* glycoside hydrolases show enhanced lichenan utilization and solvent production. Appl. Environ. Microbiol. **67:**5127–5133.
21. **Nölling, J., G. Breton, M. V. Omelchenko, K. S. Marakova, Q. Zeng, R. Gibson, H. M. Lee, J. Dubios, D. Qiu, J. Hitti, F. A. B. T. GTC Sequencing Center Production, Finishing, and Bioinformatics Teams, Y. I. Wolf, R. L. Tatusov, F. Sabathe, L. Doucette-Stamm, P. Soucaille, M. J. Daly, G. N. Bennett, E. V. Koonin, and D. R. Smith.** 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J. Bacteriol. **183:**4823–4838.
22. **Pospiech, A., and B. Neumann.** 1995. A versatile quick-prep of genomic DNA from Gram-positive bacteria. Trends Genet. **11:**217–218.
23. **Roos, J. W., J. K. McLaughlin, and E. T. Papoutsakis.** 1985. The effect of pH on nitrogen supply, cell lysis, and solvent production in fermentations of *C. acetobutylicum*. Biotechnol. Bioeng. **27:**681–694.
24. **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.
25. **Schwarz, W. H.** 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Appl. Microbiol. Biotechnol **56:**634–649.
26. **Schwarz, W. H., K. Bronnemeier, F. Grabnitz, and W. L. Staudenbauer.** 1987. Activity staining of cellulases in polyacrylamide gels containing mixed linkage beta-glucans. Anal. Biochem. **164:**72–77.
27. **Shoseyov, O., and R. H. Doi.** 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Biochemistry **87:**2192–2195.
28. **Sleat, R., A. A. Mah, and R. Robinson.** 1984. Isolation and characterization of an anaerobic, cellulolytic bacterium, *Clostridium cellulovorans* sp. nov. Appl. Environ. Microbiol. **48:**88–93.
29. **Tamaru, Y., S. Karita, A. Ibrahim, H. Chan, and R. H. Doi.** 2000. A large gene cluster for the *Clostridium cellulovorans* cellulosome. J. Bacteriol. **182:**5906–5910.
30. **Teather, R. M., and P. J. Wood.** 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. **43:**777–780.
31. **Von Heijne, G.** 1986. A new method for predicting signal peptide cleavage sites. Nucleic Acids Res. **14:**4683–4690.
32. **Wiesenborn, D. P., E. T. Papoutsakis, and F. B. Rudolph.** 1988. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. Appl. Environ. Microbiol. **54:**2717–2722.